

حاشا
الرحمن الرحيم
الرحمن الرحيم



دانشکده علوم ریاضی

پایان نامه کارشناسی ارشد آمار ریاضی

برآورد پارامترهای توزیع لاپلاس با استفاده از داده های سانسور فاصله ای

نگارنده: زهرا نادری

استاد راهنما

دکتر احمد نزاکتی رضا زاده

بهمن ۱۳۹۶

تقدیم نامہ

ماحصل آموختہ ہایم را تقدیم می‌کنم بہ آنان کہ مہر آسانی شان آرام بخش آلام زمینی ام است بہ استوارترین تکیہ گاہ آسانی ام، دستان پر مہر پدرم بہ سبزترین مگاہ زندگیم، چشمان سبز مادرم کہ ہرچہ آموختم در مکتب عشق شہ آموختم و ہرچہ بگو شتم قطرہ ای از دریای بی کران مہرتان را پاس توانم بگویم. امروز ہستی ام بہ امید شہاست و فردا کلید باغ بہشتم رضای شہا، رہ آور دی کردان سنگ تر از این ارزان نداشتہ تابہ خاک پایتان نہار کنم، باشد کہ حاصل تلاشم نسیم کوزہ غبار خشکی تان را بزوداید. بوسہ بردستان پر مہرتان.

سپاس گزارمی...

با سپاس از سه وجود مقدس:
آنان که ناتوان شدند تا ما به توانایی برسیم...
موهایشان سپید شد تا ما روسفید شویم...
و عاشقانه سوختند تا گرمابخش وجود ما و روشنگر راهمان باشند...
پدرانمان
مادرانمان
استادانمان

برخود واجب می دانم مراتب سپاس و قدردانی عمیق قلبی خود را به خدمت استاد فرزانه و گران قدرم جناب آقای دکتر احمد نزاکتی که در طول دوران تحصیل و ارائه پایان نامه از چشمه جوشان دانش و اخلاق والایشان بهره‌مند گشته‌ام، ابراز نمایم. همچنین از اساتید بزرگوار و دلسوز دکتر محمد آرشی و دکتر محمدرضا ربیعی که زحمت داوری این رساله را متقبل شدند؛ کمال تشکر و قدردانی را دارم. در انتها جا دارد از همه دوستان و عزیزانی که در این مدت مرا یاری و راهنمایی نمودند، قدردانی نمایم.

زهرا نادری
بهمن ۱۳۹۶

تعهد نامه

اینجانب زهرا نادری دانشجوی کارشناسی ارشد رشته آمار علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان برآورد پارامترهای توزیع لاپلاس با استفاده از داده های سانسور فاصله ای، تحت راهنمایی احمد نزاقتی رضا زاده متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های دیگر پژوهش گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام “دانشگاه صنعتی شاهرود” یا “Shahrood University of Technology” به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده اند، در مقالات مستخرج از پایان نامه رعایت می گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

زهرا نادری

بهمن ۱۳۹۶

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی باشد.

چکیده

در اکثر مطالعات پارامتری تحلیل بقا از مدل لوجستیک به دلیل انعطاف‌پذیر بودن استفاده می‌شود. در بیشتر موارد برای مدل لوجستیک فرض می‌شود که پارامتر شکل ثابت است، در حالیکه که این فرض برای برخی از داده‌ها نادرست است. در این مطالعه حالت سانسور فاصله‌ای برای توزیع لاپلاس، در حالتی که پارامتر شکل ثابت نیست ارائه می‌شود. و با انجام یک مطالعه شبیه‌سازی و به کارگیری روش مونت کارلویی، مزیت‌های آن نسبت به حالتی که پارامتر شکل ثابت باشد ارائه می‌شود. در پایان با یک مثال کاربردی نحوه محاسبه برآوردها نشان داده می‌شود. همچنین در این پایان نامه با استفاده از برآوردگر ماکزیمم درست‌نمایی، پارامترهای توزیع لاپلاس را برای داده‌های سانسور فاصله‌ای به دست آورده و با استفاده از شبیه‌سازی‌ها داده‌های واقعی نشان می‌دهیم که توزیع لاپلاس می‌تواند یک مدل بهتر برای داده‌های سانسور فاصله‌ای نسبت به مدل‌های دیگر از جمله لوجستیک، به رغم ساده بودن آن باشد.

کلمات کلیدی: توزیع لاپلاس، توزیع لوجستیک، برآوردگر ماکزیمم درست‌نمایی، سانسور، سانسور فاصله‌ای، قابلیت اعتماد، تابع بقا.

فهرست مطالب

م	فهرست تصاویر	
س	فهرست جداول	
۱	تعاریف و مقدمات	۱
۱	مقدمه	۱.۱
۲	آماره های مرتب	۲.۱
۲	خواص آماره های مرتب	۱.۲.۱
۲	کاربرد آماره های مرتب	۲.۲.۱
۴	سانسور	۳.۱
۵	سانسور نوع اول	۱.۳.۱
۶	سانسور نوع دوم	۲.۳.۱
۷	سانسور فاصله ای	۴.۱
۹	روش ماکسیمم درست‌نمایی	۵.۱
۱۰	فاصله اطمینان	۱.۵.۱
۱۱	شرایط نظم	۲.۵.۱
۱۱	اطلاع فیشر	۳.۵.۱
۱۲	توابع قابلیت اعتماد	۶.۱
۱۳	تابع بقاء	۱.۶.۱
۱۴	تابع نرخ خطر	۲.۶.۱
۱۴	میانگین زمان شکست	۳.۶.۱
۱۵	توزیع لاپلاس	۷.۱
۱۶	تعاریف و ویژگی های اساسی	۱.۷.۱
۲۱	برآورد پارامترها در توزیع لاپلاس با داده های سانسور فاصله ای	۲
۲۱	مقدمه	۱.۲

۲۲	برآوردگر ماکزیمم درست‌نمایی	۲.۲
۳۱	تحلیل پارامتری داده‌های بقای سانسور فاصله‌ای در توزیع لاپلاس	۳.۲
۳۲	توزیع لاپلاس	۱.۳.۲
۳۳	شبیه‌سازی	۳
۳۳	مقدمه	۱.۰.۳
۵۸	نتیجه‌گیری	۲.۰.۳
۵۸	کاربرد	۱.۳
۶۱	برنامه‌های کامپیوتری با نرم افزار R	۴
۶۹	مراجع	
۷۳	واژه‌نامه فارسی به انگلیسی	
۷۵	واژه‌نامه انگلیسی به فارسی	

فهرست تصاویر

۶	نمایش سانسور نوع اول	۱.۱
۷	نمایش سانسور نوع دوم	۲.۱
۸	انواع سانسور	۳.۱
۱۴	رابطه بین $f(t)$ ، $R(t)$ و $r(t)$	۴.۱
۱۷	نمودار چولگی	۵.۱
۱۷	توزیع چگالی توزیع‌های لاپلاس، نرمال و کوشی	۶.۱
۱۷	توزیع چگالی توزیع‌های لاپلاس و نمایی	۷.۱
۱۸	توزیع چگالی توزیع لاپلاس	۸.۱
۱۸	توزیع چگالی توزیع لاپلاس	۹.۱
۲۰	تابع چگالی توزیع لوجستیک	۱۰.۱
۲۰	تابع توزیع لوجستیک	۱۱.۱
۳۶	نمودار قدر مطلق اریبی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۱.۳
۳۶	نمودار قدر مطلق اریبی پارامتر σ برای دو مدل لاپلاس و لوجستیک	۲.۳
۳۷	نمودار MSE پارامتر μ برای دو مدل لاپلاس و لوجستیک	۳.۳
۳۷	نمودار MSE پارامتر σ برای دو مدل لاپلاس و لوجستیک	۴.۳
۳۸	نمودار چولگی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۵.۳
۳۸	نمودار چولگی پارامتر σ برای دو مدل لاپلاس و لوجستیک	۶.۳
۳۹	نمودار کشیدگی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۷.۳
۳۹	نمودار کشیدگی پارامتر σ برای دو مدل لاپلاس و لوجستیک	۸.۳
۴۲	نمودار قدر مطلق اریبی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۹.۳
۴۲	نمودار قدر مطلق اریبی پارامتر σ برای دو مدل لاپلاس و لوجستیک	۱۰.۳
۴۳	نمودار MSE پارامتر μ برای دو مدل لاپلاس و لوجستیک	۱۱.۳
۴۳	نمودار MSE پارامتر σ برای دو مدل لاپلاس و لوجستیک	۱۲.۳
۴۴	نمودار چولگی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۱۳.۳
۴۴	نمودار چولگی پارامتر σ برای دو مدل لاپلاس و لوجستیک	۱۴.۳

۴۵	نمودار کشیدگی پارامتر μ برای دومدل لاپلاس و لوجستیک	۱۵.۳
۴۵	نمودار کشیدگی پارامتر σ برای دومدل لاپلاس و لوجستیک	۱۶.۳
۴۸	نمودار قدر مطلق اربیبی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۱۷.۳
۴۸	نمودار قدر مطلق اربیبی پارامتر σ برای دومدل لاپلاس و لوجستیک	۱۸.۳
۴۹	نمودار MSE پارامتر μ برای دومدل لاپلاس و لوجستیک	۱۹.۳
۴۹	نمودار MSE پارامتر σ برای دومدل لاپلاس و لوجستیک	۲۰.۳
۵۰	نمودار چولگی پارامتر μ برای دومدل لاپلاس و لوجستیک	۲۱.۳
۵۰	نمودار چولگی پارامتر σ برای دومدل لاپلاس و لوجستیک	۲۲.۳
۵۱	نمودار کشیدگی پارامتر μ برای دومدل لاپلاس و لوجستیک	۲۳.۳
۵۱	نمودار کشیدگی پارامتر σ برای دومدل لاپلاس و لوجستیک	۲۴.۳
۵۴	نمودار قدر مطلق اربیبی پارامتر μ برای دو مدل لاپلاس و لوجستیک	۲۵.۳
۵۴	نمودار قدر مطلق اربیبی پارامتر σ برای دومدل لاپلاس و لوجستیک	۲۶.۳
۵۵	نمودار MSE پارامتر μ برای دومدل لاپلاس و لوجستیک	۲۷.۳
۵۵	نمودار MSE پارامتر σ برای دومدل لاپلاس و لوجستیک	۲۸.۳
۵۶	نمودار چولگی پارامتر μ برای دومدل لاپلاس و لوجستیک	۲۹.۳
۵۶	نمودار چولگی پارامتر σ برای دومدل لاپلاس و لوجستیک	۳۰.۳
۵۷	نمودار کشیدگی پارامتر μ برای دومدل لاپلاس و لوجستیک	۳۱.۳
۵۷	نمودار کشیدگی پارامتر σ برای دومدل لاپلاس و لوجستیک	۳۲.۳

فهرست جداول

۱۸	برخی از ویژگی‌های توزیع‌های لاپلاس، نرمال و کوشی	۱.۱
۱۹	مقایسه‌ای بین برآوردهای نمونه‌ای	۲.۱
۲۰	برخی از ویژگی‌های توزیع لوجستیک	۳.۱
۳۴	نتایج شبیه‌سازی در مدل لاپلاس	۱.۳
۳۵	نتایج شبیه‌سازی در مدل لوجستیک	۲.۳
۴۰	نتایج شبیه‌سازی در مدل لاپلاس	۳.۳
۴۱	نتایج شبیه‌سازی در مدل لوجستیک	۴.۳
۴۶	نتایج شبیه‌سازی در مدل لاپلاس	۵.۳
۴۷	نتایج شبیه‌سازی در مدل لوجستیک	۶.۳
۵۲	نتایج شبیه‌سازی در مدل لاپلاس	۷.۳
۵۳	نتایج شبیه‌سازی در مدل لوجستیک	۸.۳
۵۹	برآورد پارامترها	۹.۳
۵۹	برآورد پارامترها	۱۰.۳

فصل ۱

تعاریف و مقدمات

۱.۱ مقدمه

در این فصل، مفاهیم و مفاهیمی را که در این پایان‌نامه مورد استفاده قرار خواهد گرفت، معرفی می‌کنیم. در بخش اول، آماره‌های مرتب و کاربرد آن توضیح داده شده است. در بخش دوم، ضمن معرفی مفاهیم سانسورها و انواع آن‌ها، چند نوع خاص را بیشتر بررسی نموده ایم. در بخش سوم سانسور فاصله‌ای را مطرح کرده و در بخش چهارم برآوردگرهای ماکسیمم درست‌نمایی و برخی ویژگی‌های مهم آن بیان شده است و در بخش پنجم مفاهیم مربوط به توابع قابلیت اعتماد بررسی کرده و در بخش ششم به معرفی فرآیند هندسی پرداخته ایم و در بخش‌های بعد توزیع‌های لاپلاس و لوجستیک را مورد بررسی قرار داده ایم. مطالب این فصل عمدتاً از مقاله‌های چن و همکاران^۱ [۱]، دینگ و همکاران^۲ [۶]، هورریمن^۳ [۱۰]، کوزوبوسکی و همکاران^۴ [۱۱]، لین و همکاران^۵ [۱۳]، لیندسی و همکاران^۶ [۱۴]، نگ و

^۱Chen

^۲Ding

^۳Hurlimann

^۴Kozubowski

^۵Lin

^۶Lindsey

همکاران^۷ [۱۸] و ... استفاده شده است.

۲.۱ آماره های مرتب

فرض کنید T_1, \dots, T_n متغیرهای تصادفی باشند. اگر آن‌ها را به ترتیب صعودی مرتب کنیم و مقادیر مرتب شده را به صورت $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(n)}$ ، نشان دهیم. در این صورت $T_{(j)}$ را j -امین آماره مرتب در نمونه‌ای به حجم n گویند. کوچکترین و بزرگترین آماره مرتب، یعنی $T_{(1)}$ و $T_{(n)}$ به مقادیر فرین معروف هستند و در مطالعات آماری بیشتر مورد توجه می‌باشند.

۱.۲.۱ خواص آماره های مرتب

آماره‌های مرتب تشکیل یک زنجیر مارکف می‌دهند، زیرا هر آماره‌ی مرتب فقط به آماره مرتب ما قبل خود ربط دارد و از دیگر آماره‌ها مستقل است. فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی از توزیع پیوسته $F(x)$ باشد، آن‌گاه:

۱- تابع چگالی توام آماره‌های مرتب

$$f_{X_{1:n}, \dots, X_{n:n}}(x_1, x_2, \dots, x_n) = \begin{cases} n! \prod_{i=1}^n f_X(x_i) & x_1, \dots, x_n \\ 0 & \text{در غیر این صورت} \end{cases}$$

۲- تابع چگالی حاشیه‌ای یک آماره مرتب

$$f_{X_{i:n}}(x_i) = [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x), \quad -\infty < x_i < \infty.$$

۳- تابع چگالی توام دو آماره مرتب

$$f_{X_{i:j:n}}(x_i, x_j) = \begin{cases} C(n, i, j) [F(x_i)]^{i-1} [F(x_i) - F(x_j)]^{j-i-1} \\ \times [1 - F(x)]^{n-j} f(x_i) f(x_j) & -\infty < x_i < x_j < \infty. \\ 0 & x_i \geq x_j \end{cases}$$

۲.۲.۱ کاربرد آماره‌های مرتب

آماره‌های مرتب در قسمت‌های مختلف توصیفی و استنباطی علم آمار، دارای کاربرد می‌باشند. یکی از کاربردهای این آماره‌ها در آمار توصیفی، کشف مشاهدات پرت داده‌های خیلی کوچک یا خیلی بزرگ در یک مجموعه از داده‌ها می‌باشد همان گونه که می‌دانیم یکی از ضعف‌های

استفاده از میانگین به عنوان معیاری جهت تمرکز داده‌ها، میزان حساسیت بالای آن نسبت به داده‌های دور افتاده و تغییرات الگو است. عوض میانه یا میانگین اصلاح شده، که این معیارها نیز جزء معیارهای تمرکز می‌باشند، نسبت به تغییرات الگو از حساسیت کمتری برخوردار هستند و در محاسبه آن‌ها، آماره‌های مرتب نقش عمده‌ای ایفا می‌کنند.

در آمار استنباطی نیز می‌توان به مثال‌های زیر در خصوص کاربرد آماره‌های مرتب اشاره نمود.

۱. در محاسبه تابع توزیع تجربی نیازمند استفاده از آماره‌های مرتب هستیم و از آنجا که در آزمون‌های نیکویی برازش، اغلب تمرکز روی تغییرات بین تابع توزیع تجربی و تابع توزیع فرضی است، آماره‌های مرتب در این آزمون‌ها نقش اساسی دارند.

۲. در موضوعات کنترل کیفیت برای بررسی در کنترل بودن تولیدات، اغلب از نمودار میانگین و دامنه تغییرات و یا از نمودار میانه و دامنه تغییرات استفاده می‌شود که برای محاسبه میانه و دامنه تغییرات ناچار به استفاده از آماره‌های مرتب می‌باشیم.

۳. آماره‌های مرتب، در بسیاری از موارد به عنوان آماره‌های بسنده معرفی می‌شوند و برآوردهای ناریب با کمترین واریانس، فواصل اطمینان و پرتوان ترین آزمون برای پارامترهای مجهول را فراهم می‌کنند.

۴. در آزمایشات طول عمر، اغلب به دلیل فاکتورهایی از قبیل زمان یا اعتبار مالی، امکان شرکت دادن تمام واحدها در آزمایش وجود ندارد. بنابراین مشاهده طول عمر تمام واحدها امکان پذیر نیست، به همین دلیل از روش‌های سانسور استفاده می‌کنیم و آماره‌های مرتب نقش مهمی در سانسورها دارند.

برای جزئیات بیشتر در مورد خواص آماره‌های مرتب و کاربرد آن‌ها، می‌توان به دیوید و ناگاراچا (۲۰۰۳) یا آرنولد و همکاران (۲۰۰۸) مراجعه نمود.

همان طور که در بخش قبل توضیح داده شد آماره‌های مرتب نقش به سزایی در مطالعات طول عمر و تحلیل داده‌های قابلیت اعتماد دارند. در این بخش این آماره‌ها را در حالتی که توزیع تحت بررسی نمایی است مورد مطالعه قرار می‌دهیم. فرض کنید T_1, T_2, \dots, T_n نمونه تصادفی به حجم n از توزیعی پیوسته با تابع توزیع F و تابع چگالی f باشد. آماره‌های ترتیبی نمونه را به صورت $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(n)}$ نمایش می‌دهیم. در بین دلایل مختلف اهمیت آماره‌های ترتیبی در قابلیت اعتماد، به دو دلیل زیر اشاره می‌کنیم.

(الف) در یک سیستم $(n - k - 1)$ از n ، طول عمر سیستم برابر با $T_{(k)}$ است.

(ب) در برآورد متوسط طول عمر مثلا یک قطعه الکتریکی، هنگامی که قرار است n داده تولید کنیم n قطعه را در زمان $t = 0$ وارد آزمایش می‌کنیم و زمان شکست آن‌ها را ثبت می‌کنیم. واضح است که زمان شکست قطعات مقادیر آماره ترتیبی $T_{(1)}, T_{(2)}, \dots, T_{(n)}$ خواهند بود.

۳.۱ سانسور

تحلیل بقاء از نظر علم آمار شامل فنون مختلف آماری در تحلیل متغیرهای نوعاً نامنفی، زمان خرابی یک مؤلفه فیزیکی (مکانیکی یا الکتریکی) یا زمان مرگ یک واحد زنده (سلول، بیمار و...) می باشد. ممکن است این متغیر زمان یادگیری یک مهارت باشد، یا حتی امکان دارد به زمان هیچ ارتباطی نداشته باشد. برای مثال این متغیر می تواند مبلغ پرداختی یک شرکت بیمه در وضعیت خاص باشد. در برخی موارد بیمار بهبود یافته و مبلغ کل پرداختی بیمه او معین است. در موارد دیگر، ممکن است بیماری همچنان ادامه داشته باشد و تنها مبلغ پرداختی تا آن زمان معلوم باشد. آنچه تحلیل بقاء را از دیگر مباحث آماری جدا می کند، عمل سانسور کردن است، به طور کلی یک مشاهده سانسور شده فقط شامل قسمتی از اطلاعات مربوط به متغیر تصادفی مورد نظر است. در آزمون های طول عمر، آزمایش های کلینیکی، تحقیقات زیست شناسی، تحلیل بقاء و دیگر زمینه های کاربردی علم آمار با نمونه هایی مواجه هستیم که در قسمت هایی از مقادیر ممکن متغیر تصادفی مورد مطالعه، محدود شده و یا همه مشاهدات نمونه ثبت نشده اند. یا همه واحدهای نمونه به نتیجه نرسیده اند. این محدودیت ممکن است به صورت اختیاری توسط آمارگر اعمال شود یا ماهیت آزمایش طوری باشد که خود به خود محدودیتی را در مشاهدات به وجود آورد.

بعضی از محدودیت ها عبارتند از:

۱. هزینه بر بودن آزمایش
۲. عدم دسترسی به همه واحدها یا مایوس شدن از نتیجه دادن همه واحدها
۳. اشکالات فنی
۴. طولانی شدن مدت آزمایش
۵. فرصت کم برای اعلام نتایج

این محدودیت به وجود آمده در نمونه را سانسور گویند. هر چند با سانسور کردن مشاهدات، مقداری از اطلاعات موجود را از دست می دهیم، با این وجود گاهی اوقات به دلایلی که بیان شد مایل یا مجبور به سانسور کردن مشاهدات هستیم. برخی از انواع سانسورها عبارتند از:

۱. سانسور نوع یک
۲. سانسور نوع دو
۳. سانسور فاصله ای
۴. سانسور فزاینده نوع یک

۵. سانسور فزاینده نوع دو

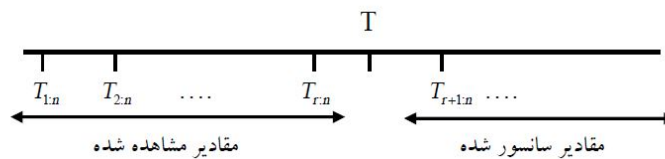
در این بخش، سانسور نوع یک و سانسور نوع دو و سانسور فاصله‌ای توضیح داده می‌شوند. برای جزئیات بیشتر در مورد انواع سانسورها و کاربرد آن‌ها به کوهن، بالاکریشن و کوهن ۱۹۹۱، لاولس ۱۹۸۲ و همچنین نلسن ۲۰۰۴ مراجعه نمایید.

۱.۳.۱ سانسور نوع اول

در مطالعات مربوط به قابلیت اعتماد و آزمایش‌های طول عمر ممکن است با نمونه‌هایی مواجه شویم که در آنها طول عمر واحدهای آزمایشی به طور نادقیق گزارش شده‌اند. روش‌های ارائه شده در آمار برای برآورد پارامتر توزیع‌های طول عمر در این موارد مناسب نیستند. در مطالعات طول عمر، اگر آزمایشگر بخواهد زمان شکست همه واحدهای آزمایشی را مشاهده و ثبت کند باید مدت آزمایش را به اندازه ماکزیمم طول عمر واحدها در نظر بگیرد که این امر عملاً در بیشتر موارد ممکن نیست و منطقی نمی‌باشد. بنابراین با نمونه‌هایی مواجه می‌شویم که اطلاعات کامل بعضی از واحدهای نمونه به دلایلی ثبت نشده‌اند. این محدودیت به وجود آمده در نمونه را سانسور تعریف می‌کنند. انواع طرح‌های سانسور معروف در مطالعات عملی عبارتند از: سانسور نوع اول، سانسور نوع دوم، سانسور فاصله‌ای، سانسور دوطرفه نوع دوم، سانسور فزاینده نوع اول، سانسور فزاینده نوع دوم، سانسور ترکیبی نوع اول و دوم و سانسور ترکیبی فزاینده نوع دوم. روش‌های استنباطی با استفاده از داده‌های حاصل از طرح‌های مختلف سانسور مورد توجه آماردانان زیادی قرار گرفته است. نتایج ارایه شده در پژوهش‌های انجام شده بر این فرض استوار است که داده‌های طول عمر مشاهده شده مقادیر دقیقی باشند، در حالی که در اغلب کاربردهای عملی ممکن است با داده‌های نادقیق مواجه شویم. در حقیقت می‌توان گفت معرفی روش‌های مقایسه، بخش قابل توجهی از مباحث آماری را به خود اختصاص داده است. این موضوع در تحلیل بقا و به ویژه در نظریه قابلیت اعتماد از اهمیت بیشتری برخوردار بوده و روش‌های مقایسه‌ی متنوعی مورد مطالعه‌ی پژوهشگران قرار گرفته است. استفاده از مدل احتمالی تنش مقاومت، یک شکل مقایسه‌ی بین دو متغیر است که در دهه‌های اخیر بسیار مورد توجه بوده است.

فرض کنید n آزمودنی را در زمان $t = 0$ وارد آزمایش می‌کنیم. تا زمان ثابت T می‌مانیم و سپس آزمایش را خاتمه می‌دهیم. در این صورت نتیجه آزمایش این است که r محصول $r = 0, 1, \dots, n$ قبل از زمان T از کار افتاده‌اند و زمان دقیق شکست آن‌ها معلوم است و تنها اطلاعاتی که در مورد $n - r$ محصول باقیمانده داریم آن است که طول عمرشان از T بیشتر است و اصطلاحاً سانسور شده‌اند (در این جا T ثابت و r متغیر تصادفی است). شکل (۱.۱) این نوع سانسور را نشان می‌دهد.

اگر مشاهدات به دست آمده از آزمایش انجام شده بر اساس سانسور نوع اول را با $t_{1:n}, t_{2:n}, \dots, t_{n:n}$ نشان دهیم و طول عمر واحدها دارای تابع چگالی احتمال $f(t; \theta)$ و تابع توزیع تجمعی $F(t; \theta)$



شکل ۱.۱: نمایش سانسور نوع اول

باشند، آنگاه تابع درست‌نمایی بصورت زیر خواهد شد

$$L(\theta) = \prod_{i=1}^r f(t_{i:n}; \theta) [1 - F(t; \theta)]^{n-r}.$$

توجه داریم که متغیر تصادفی r در سانسور نوع اول دارای توزیع دوجمله ای با پارامترهای $(n; F(t; \theta))$ است.

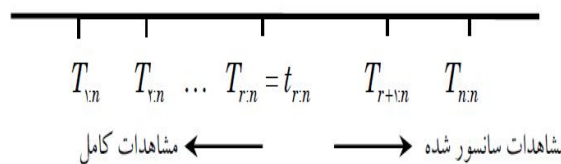
مثال ۱.۳.۱. یک مجموعه شامل n قطعه الکترونیکی، زمانی که در حالت روشن قرار دارند، در نظر بگیرید. آزمایش تا 20° روز ادامه پیدا می کند. در طی آزمایش مدت زمان تا خرابی هر قطعه یادداشت می شود. بعد از 20° روز (زمان اتمام آزمایش) برای قطعاتی که تا این زمان سالم هستند، فقط می دانیم که زمان لازم تا خرابی آنها بیشتر از 20° روز است. این مثال، نمونه‌ای از داده‌های سانسور نوع یک است که نقطه سانسور آن 20° روز است.

۲.۳.۱ سانسور نوع دوم

یکی از چالش‌های موجود در تحلیل بقا وجود سانسور در داده‌ها است. در بسیاری از تحقیقات از جمله آزمون‌های طول عمر و مطالعات تحلیل بقا با نمونه‌هایی مواجه هستیم که اطلاعات کامل بعضی از واحدهای نمونه به دلایلی ثبت نشده‌اند، یا به عبارت دیگر آزمایشگر ممکن است قادر نباشد اطلاعات کاملی از زمان شکست همه واحدهای مورد آزمایش به دست آورد. به عنوان مثال اگر زمان مرگ یا از کارافتادگی واحد مورد بررسی، خارج از دوره‌ی زمانی مورد مطالعه باشد در این صورت تعداد دقیق مرگ یا از کارافتادگی که در یک بازه اتفاق می افتد معلوم نیست. به عنوان مثالی دیگر، در یک آزمایش بالینی ممکن است افراد از ادامه آزمایش منصرف شوند و یا به علت کمبود بودجه، آزمایش متوقف شود. همچنین در آزمایش‌های صنعتی ممکن است واحدهای آزمایش به منظور صرفه‌جویی در زمان و هزینه، قبل از شکست و به صورت از پیش تعیین شده از آزمایش کنار گذاشته شوند. در چنین شرایطی داده‌هایی که تحلیل گر به دست می‌آورد، داده‌های سانسور شده می‌باشند و گوییم سانسور رخ داده است. اگر در آزمایشی با n واحد بخواهیم دقیقاً m ($m < n$) شکست را مشاهده کنیم و بعد از مشاهده m - امین شکست آزمایش را متوقف کنیم، در این صورت طول عمر $(n - m)$ واحد باقیمانده سانسور می‌شوند، به این روش از سانسور، سانسور نوع دو یک مرحله‌ای از راست گویند.

در این روش از سانسور در حقیقت زمان لازم برای رسیدن به m - امین شکست متغیر تصادفی مورد مطالعه است. این روش از سانسور می‌تواند به سانسور نوع دوم دو مرحله‌ای (راست و چپ) تعمیم داده شود. این حالت زمانی رخ می‌دهد که زمان‌های r شکست اول ثبت نشده باشد و مشاهدات از زمان $(r + 1)$ امین شکست شروع شوند و در زمان مشاهده m - امین شکست، آزمایش به اتمام می‌رسد. (در این جا r ثابت و $t_{r:n}$ مقدار یک متغیر تصادفی است). شکل ۲.۱ این نوع سانسور را نشان می‌دهد.

با این فرض که طول عمر واحدها دارای تابع چگالی احتمال $f(t; \theta)$ و تابع توزیع تجمعی



شکل ۲.۱: نمایش سانسور نوع دوم

$F(t; \theta)$ باشند، آنگاه تابع درست‌نمایی در این حالت به صورت زیر خواهد شد.

$$L(\theta) = \prod_{i=1}^r f(t_{i:n}; \theta) [1 - F(t; \theta)]^{n-r}.$$

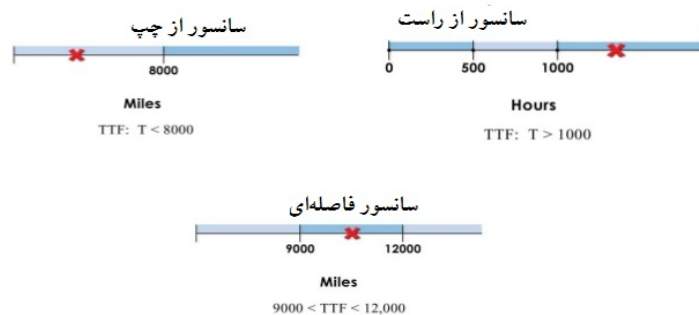
مثال ۲.۳.۱. فرض کنید هدف برآورد میانگین طول عمر یک قطعه تولید شده باشد. آزمایش استفاده از قطعه‌های تولید شده را تا زمان خرابی قطعه 50° - ام ادامه می‌دهیم. بعد از اینکه 50° قطعه خراب شد، آزمایش را متوقف می‌کنیم. در اینجا زمان خرابی سایر قطعات مشخص نیست. در این سانسور تعداد واحدهایی که خراب می‌شوند از قبل مشخص و برابر مقدار ثابتی است که توسط محقق تعیین می‌شود، ولی زمان اتمام آزمایش یک متغیر تصادفی است. برای جزئیات بیشتر در مورد سانسور نوع دو، می‌توان به کوهن ۱۹۹۱ و بالا کریشن ۱۹۹۶ مراجعه نمود.

۴.۱ سانسور فاصله‌ای

سانسور فاصله‌ای معمولاً برای نشان دادن یک نوع طرح نمونه‌گیری برای توصیف نوعی از اطلاعات ناقص به کار می‌رود. داده‌های سانسور فاصله‌ای، به معنی آن است که یک متغیر تصادفی فقط در فاصله‌ای که قرار گرفته، مورد توجه است. در اکثر برنامه‌های تجزیه و تحلیل بقا، متغیر تصادفی زمان برخی از رویدادهاست، مانند مرگ و یا بیماری، برای مثال داده‌های بقا با سانسور فاصله‌ای زمانی در مطالعات پزشکی یا سلامتی رخ می‌دهد که مستلزم پیگیری پیوسته است.

بسیاری از مطالعات بالینی و مطالعات طولی سانسور فاصله‌ای می‌باشند. در این وضعیت، فردی که برای مشاهدات برنامه ریزی شده برای تغییرات بالینی، با تغییر وضعیت بیماری ممکن است برخی از مشاهدات نیز تغییر کند. به این ترتیب، یک دوره سانسور شده چرا که با فاصله زمانی ایجاد شده تغییراتی اعمال شده است. یک مورد خاص مهم در فاصله سانسور فاصله‌ای، داده‌ها اطلاعات و وضعیت فعلی است. در این وضعیت، هر یک از این موارد (وضعیت وقوع رویداد مورد علاقه در زمان مشاهده) تنها یکبار مشاهده می‌شوند به عبارت دیگر مشاهدات زمان وقوع این رویداد، سانسور چپ یا راست است (زمان زنده ماندن کمتر یا بیشتر از زمان مشاهده است).

یک مورد خاص دیگر از داده‌های سانسور شده که ممکن است اتفاق بیفتد داده‌ها باقی مانده و سانسور شده، که در آن زمان به رویداد مورد علاقه یا سزاوار سقوط است یا دقیقاً شناخته شده است. یک دلیل پشت سر گذاشتن داده‌های چپ سانسور شده، عدم توانایی اندازه‌گیری متغیر مورد علاقه هنگامی در یک سطح خاص است. برای مثال عوارض جانبی مربوط به یک دارو که گاهی اوقات مشخص می‌شود. برای تجزیه و تحلیل زمان شکست سانسور شده داده‌ها، در زیر، ما ابتدا برآورد غیر پارامتری توابع توزیع به عنوان عملکرد خطر برای زمان بقا را بررسی کرده و سپس تجزیه و تحلیل رگرسیون با سانسور فاصله‌ای را، تحت مدل‌های گوناگون رگرسیون مورد بررسی قرار می‌دهیم. طرح سانسور فاصله‌ای را می‌توان به شرح زیر توصیف کرد: فرض کنید n آزمودنی با طول عمر T_1, \dots, T_n وجود دارد. که در آن (L_i, R_i) - امین سانسور تصادفی از توزیع $F(\theta)$ باشد. در اینجا L_i و R_i نقطه ابتدا و انتهای سانسور فاصله‌ای را نشان می‌دهند. اگر T_i طول عمر سانسور در فاصله‌ی i ام باشد T_i قابل مشاهده است اگر T_i در فاصله زمانی (L_i, R_i) نباشد، در غیر این صورت T_i قابل مشاهده نیست. حال δ_i را تعریف کنید $\delta_i = I(T_i)$. شکل زیر انواع سانسورها را نشان می‌دهد. حال فرض کنید که داده‌های مشاهده



شکل ۳.۱: انواع سانسور

شده را با $\{I_i\}_{i=1}^n$ که $I_i = (L_i, R_i)$ ، نشان دهیم. برای i -ام نوع رخداد اگر $L_i = 0$ یعنی مشاهدات از چپ سانسور شده‌اند و اگر $R_i = \infty$ یعنی مشاهدات از راست سانسور شده‌اند. فرض کنید $\{\tau_j\}_{j=0}^{m+1}$ عناصر مرتب و α_{ij} یک نشانگر (متغیر ساختگی) از رویداد I_i در (τ_{j-1}, τ_j)

باشد، بنابراین لگاریتم تابع درستنمایی متناسب است با

$$\log L(\theta) = \alpha_0 + \sum_{i=1}^{m+1} n_i \log (F(\tau_i, \theta) - F(\tau_{i-1}, \theta)) \quad (1.1)$$

$$n_i = \sum_j \alpha_{ij}$$

۵.۱ روش ماکسیمم درستنمایی

در تحلیل داده های قابلیت اعتماد روش ماکسیمم درستنمایی ML و استنباط آماری بر مبنای ML نقش محوری ایفا می کند. فرض کنید توزیع یک متغیر طول عمر T دارای تابع چگای احتمال $f(t, \theta)$ باشد که در آن θ یک بردار k مجهول است. فرض کنید یک نمونه تصادفی از $f(t, \theta)$ استخراج کنیم و نمونه منتج به مشاهدات مستقل t_1, t_2, \dots, t_n شود (t_i ها زمان دقیق شکست را نشان می دهد). طبق تعریف تابع درستنمایی که با $L(\theta)$ نمایش می دهیم به صورت زیر خواهد بود.

$$L(\theta; t_1, t_2, \dots, t_n) = \prod_{i=1}^n f(t_i, \theta) \quad (2.1)$$

لازم به ذکر است که

(الف) تابع درستنمایی $L(\theta)$ لزوماً نسبت به θ مشتق پذیر نیست.

(ب) توجه داشته باشید که پس از مشاهده مقادیر t_i ، تابع درستنمایی تنها تابعی از پارامتر θ است. از نقطه نظر محاسباتی معمولاً بهتر است از لگاریتم تابع درستنمایی استفاده کنیم. در این صورت تساوی (۲.۱) به تساوی زیر تبدیل می شود.

$$l(\theta; t_1, t_2, \dots, t_n) = \sum_{i=1}^n \ln f(t_i, \theta)$$

که در آن $l(\theta) = \ln L(\theta)$ لگ تابع درستنمایی است. بنا به تعریف، برآوردگر درستنمایی ماکسیمم MLE پارامتر θ از حل معادلات زیر به دست می آید که به آن معادله درستنمایی می گویند.

$$\frac{\partial}{\partial \theta_i} l(\theta) = 0, \quad i = 1, 2, \dots, k$$

در اغلب موارد به ویژه در آزمون های طول عمر، برآورد $\hat{\theta}$ به صورت تحلیلی قابل حصول نیست و باید آن ها را با استفاده از روش های عددی محاسبه کرد. بسیاری از نرم افزارهای موجود توانایی حل معادلات درستنمایی به روش عددی را دارند.

MLE نه تنها از پشتوانه شهودی بسیار قوی برخوردار است، بلکه ویژگی های منحصر به فرد

دیگری نیز دارد که آن‌ها را از دیگر برآوردگرها متمایز می‌کند. از جمله ویژگی‌های برجسته MLE ها عبارت است از:

۱. روش ماکزیمم درست‌نمایی پایاست. بدین معنی که اگر علاقه مند باشیم MLE تابعی مانند $\psi(\theta)$ را بیابیم آنگاه $\hat{\psi} = \psi(\hat{\theta})$ ، MLE $\psi(\theta)$ است که در آن $\hat{\theta}$ برآورد ماکزیمم درست‌نمایی θ است.

۲. تحت بعضی شرایط نظم، برای حجم نمونه‌های بزرگ، توزیع برآوردگرهای MLE نرمال است.

۳. تحت شرایط نظم، برآوردگرهای MLE در حد دارای خاصیت سازگاری هستند. بدین معنی که برای حجم نمونه‌های بزرگ MLE تقریباً ناریب و دارای کمترین واریانس می‌باشند.

۴. روش درست‌نمایی را می‌توان برای شیوه‌های مختلف جمع آوری داده‌ها، مانند انواع روش‌های سانسور به کار برد. لذا برآورد MLE را می‌توان براساس مکانیزم‌های مختلف جمع آوری داده‌ها به دست آورد.

ساختار تابع درست‌نمایی داده‌های سانسور شده در کتاب‌ها و مقالات زیادی مورد بحث قرار گرفته‌اند. در سانسور نوع یک و نوع دو، n واحد آزمودنی را در لحظه $t = 0$ وارد آزمون می‌کنیم و آزمون را تا لحظه t ادامه می‌دهیم و زمان شکست آن‌ها ثبت می‌کنیم. فرض کنید $f(x)$ و $F(x)$ به ترتیب تابع چگالی احتمال و تابع توزیع تجمعی از توزیع عمر انتخاب شده با پارامتر θ و تعداد $(n - r)$ واحد بدون شکست در آزمایش باقی مانده باشند. آن‌گاه تابع درست‌نمایی آن‌ها را به صورت زیر خواهیم داشت:

$$L(\theta) = \frac{n!}{(n-r)!} \left(\prod_{i=1}^r f(x_{(i)}) \right) (1 - F(x_T))^{n-r} \quad (3.1)$$

توجه کنید که در سانسور نوع یک $x_T = t$ و در سانسور نوع دو $x_T = x(r)$ است.

۱.۵.۱ فاصله اطمینان

اگر X_1, \dots, X_n یک نمونه تصادفی از تابع چگالی $f(., \theta)$ باشد و T_1 و T_2 دو آماره باشند به طوری که $(T_1 \leq T_2)$ و $P_\theta [T_1 < \theta < T_2] = \gamma$ که γ به θ بستگی نداشته و برابر با $1 - \alpha$ باشد، در این صورت بازه تصادفی $[T_1, T_2]$ را یک فاصله اطمینان γ درصدی برای θ و γ را ضریب اطمینان می‌نامیم که به صورت دلخواه توسط پژوهشگر انتخاب می‌شود. اغلب مقادیر ۵ درصد و ۱ درصد که متناظر با سطوح اطمینان ۹۵ درصد و ۹۹ درصد می‌باشد برای α در نظر گرفته می‌شود.

لازم به ذکر است که تعداد کمی فاصله اطمینان دقیق شناخته شده است. بنابراین، در اکثر

مواقع، نیازمند جستجوی راهی برای دستیابی به فواصل اطمینان تقریبی با دقتی در سطح قابل قبول باشیم. یک راه برای رسیدن به فواصل اطمینان تقریبی، فاصله اطمینان مجانبی است که به صورت زیر تعریف می‌شود.

تعریف ۱.۵.۱. بازه $[T_1, T_2]$ را فاصله اطمینان مجانبی گویند هرگاه گزاره

$$P[T_1 < \theta < T_2] = \gamma$$

به صورت حدی وقتی n به سمت بی‌نهایت میل کند، درست باشد.

۲.۵.۱ شرایط نظم

فرض کنید X یک متغیر تصادفی حقیقی مقدار باشد که مطلقاً پیوسته و دارای تابع توزیع تجمعی $F_\theta \in \Omega_\theta = F(x; \theta) : \theta \in \Theta$ باشد. در این پایان‌نامه خانواده توزیع‌هایی را مورد بررسی قرار می‌دهیم که دارای شرایط نظم به شرح زیر باشند:

(۱) Θ یک زیر بازه از مجموعه اعداد حقیقی است ($\Theta \subseteq R$) و دامنه تغییرات X به θ بستگی ندارد.

(۲) مشتق تابع چگالی نسبت به θ وجود دارد. به عبارت دیگر $\frac{\partial}{\partial \theta} f(x; \theta)$ معنی‌دار است. در نتیجه $\frac{\partial}{\partial \theta} \log f(x; \theta)$ برای هر $\theta \in \Theta$ وجود دارد.

(۳) اگر $A \subset X$ یک مجموعه اندازه‌پذیر نسبت به m باشد، آن‌گاه:

$$\frac{\partial}{\partial \theta} \int_A f(x; \theta) dx = \int_A \frac{\partial}{\partial \theta} f(x; \theta) dx$$

(۴) مشتق دوم $F_\theta \in \Omega_\theta$ نسبت به x وجود داشته باشد و رابطه زیر برقرار باشد:

$$\frac{\partial}{\partial \theta} f(x; \theta) = \frac{\partial^2}{\partial \theta \partial x} F(x; \theta) = \frac{\partial^2}{\partial x \partial \theta} F(x; \theta).$$

(۵) برای هر برآوردگر T پارامتر θ ، رابطه زیر برقرار باشد:

$$\frac{\partial}{\partial \theta} \int t(x) f(x; \theta) dx = \int \frac{\partial}{\partial \theta} t(x) f(x; \theta) dx$$

۳.۵.۱ اطلاع‌فیش

فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال $f(x; \theta)$ باشد. در این صورت تحت شرایط نظم، میزان تغییر نسبی چگالی، زمانی که θ به $\theta + \Delta\theta$ تغییر می‌کند عبارت است از:

$$\frac{f(x; \theta + \Delta\theta) - f(x; \theta)}{\Delta\theta f(x; \theta)}.$$

حد کسر فوق وقتی که $\Delta\theta$ به سمت صفر میل می کند را با $S(x; \theta)$ نشان می دهیم و داریم:

$$\begin{aligned} S(x; \theta) &= \frac{\frac{\partial f(x; \theta)}{\partial \theta}}{f(x; \theta)} = \frac{f'_\theta(x; \theta)}{f(x; \theta)} \\ &= \frac{\partial}{\partial \theta} \log f(x; \theta) \end{aligned} \quad (۴.۱)$$

که در آن $f'_\theta(x; \theta)$ مشتق تابع $f(x; \theta)$ نسبت به θ را نشان می دهد. با تبدیل x به X ، متغیر تصادفی

$$S(X; \theta) = \frac{f'_\theta(X; \theta)}{f(X; \theta)}$$

به دست می آید که به متغیر نمره (یا امتیاز) معروف است. به سادگی ثابت می شود که تحت شرایط نظم $E[S(X; \theta)] = 0$ و در نتیجه

$$Var[S(X; \theta)] = E^2[S(X; \theta)] = E\left[\left(\frac{\partial}{\partial \theta} \log f(x; \theta)\right)^2\right] > 0.$$

هر اندازه $Var[S(X; \theta)]$ بیشتر باشد، متوسط مربع تغییر نسبی چگالی X در θ بیشتر می باشد. از این رو $Var[S(X; \theta)]$ را مقدار اطلاع درباره پارامتر θ در متغیر تصادفی X می نامند و آن را با $I_X(\theta)$ نشان می دهند. این نوع اطلاع که به اطلاع فیشر مشهور است، برای اولین بار توسط فیشر (۱۹۲۵) آماردان انگلیسی، معرفی گردید. اینک اطلاع فیشر را به صورت زیر تعریف می کنیم.

تعریف ۱.۵.۱. فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال $f(x; \theta)$ باشد. تحت شرایط نظم، میزان اطلاع فیشر نهفته در متغیر تصادفی X در خصوص پارامتر θ را می توان به صورت زیر محاسبه نمود:

$$I_X(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \log f(X; \theta)\right].$$

یکی از مهم ترین کاربرد اطلاع فیشر، محاسبه آن در نمونه های سانسور فزاینده می باشد که در سال های اخیر توجه بسیاری از محققان را به خود جلب کرده است از جمله می توان به زنگ و پارک (۲۰۰۴)، بالاکریشن و همکاران (۲۰۰۸) و ابوالنین (۲۰۰۸) اشاره کرد.

۶.۱ توابع قابلیت اعتماد

توابع متفاوتی در قابلیت اعتماد مورد استفاده قرار می گیرند که در این بخش با برخی از آنها مانند تابع بقاء، تابع نرخ خطر، میانگین زمان شکست و میانگین طول عمر فقط در حالت پیوسته آشنا می شویم.

تعریف ۱.۶.۱. تابع $f(x)$ در بازه $[a, b]$ که $0 \leq a \leq b \leq \infty$ ، مطلقاً پیوسته می‌باشد هرگاه برای هر مقدار مثبت ϵ ، $\delta > 0$ ای وجود داشته باشد به طوری که برای زیر دنباله‌های مجزای

$$[x_k, y_k] \quad k = 1, 2, \dots, n$$

که در شرط

$$\sum_{k=1}^n |y_k - x_k| < \delta$$

صدق می‌کنند، داشته باشیم:

$$\sum_{k=1}^n |g(y_k) - g(x_k)| < \epsilon$$

۱.۶.۱ تابع بقاء

اولین تعریفی که در قابلیت اعتماد با آن مواجه می‌شویم، تابع قابلیت اعتماد است. بسیاری از توابعی که در ادامه معرفی می‌گردند را می‌توان بر اساس این تابع بیان نمود فرض کنید متغیر تصادفی نامنفی پیوسته X طول عمر یک قطعه باشد که دارای تابع توزیع F و تابع چگالی f باشد که در این صورت می‌توان نوشت.

$$F(x) = \begin{cases} P(X \leq x) = \int_0^x f(t)dt & \text{اگر } x \geq 0 \\ 0 & \text{اگر } x < 0 \end{cases} \quad (5.1)$$

تعریف ۲.۶.۱. برای متغیر تصادفی نامنفی پیوسته X تابع نامنفی S را که در آن

$$S(x) = 1 - F(x) = \int_x^{\infty} f(t)dt \quad (6.1)$$

تابع قابلیت اعتماد یا تابع بقا گوئیم. اگر تابع $F_X(\cdot)$ مطلقاً پیوسته باشد، آن‌گاه تابع چگالی متغیر تصادفی X وجود دارد و $f_X(x) = -S'_X(x)$. در واقع $S(x)$ احتمال عدم وقوع خرابی در بازه $(0, x]$ را نشان می‌دهد. همانند تابع توزیع می‌توان خواص چهارگانه تابع بقاء متغیر تصادفی X را به صورت زیر بیان نمود:

$$0 \leq S(x) \leq 1; \quad \forall x \geq 0 \quad -1$$

$$S(0) = 1, S(\infty) = \lim_{x \rightarrow \infty} S(x) = 0 \quad -2$$

۳- $S(x)$ تابعی یکنوا و غیر صعودی است.

۴- $S(x)$ تابعی از چپ پیوسته است.

۲.۶.۱ تابع نرخ خطر

تابع نرخ خطر به دلیل تعبیر احتمالی که در تعریف آن نهفته است و در ادامه آن را بیان می‌کنیم، از اهمیت بالایی در قابلیت اعتماد برخوردار است. به عنوان مثال، استیفنسون (۱۹۳۰) در علم بیمه از این تابع تحت عنوان شدت میرایی نام برده است. فاصله زمانی $(x, x + \Delta x]$ را در نظر می‌گیریم. احتمال خرابی قطعه در این فاصله به شرط اینکه در بازه $[\circ, x]$ خرابی اتفاق نیفتاده باشد عبارتست از:

$$\begin{aligned} P(x < X < x + \Delta x | X > x) &= \frac{P(x < X < x + \Delta x)}{P(X > x)} \\ &= \frac{F(x + \Delta x) - F(x)}{S(x)} \end{aligned} \quad (7.1)$$

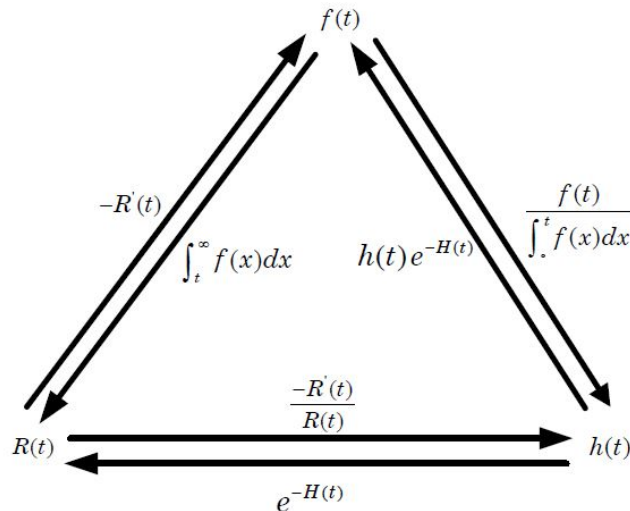
تعریف ۱.۶.۱. با تقسیم رابطه (۷.۱) بر فاصله زمانی به طول $\Delta x > \circ$ وقتی که $\Delta x \rightarrow \circ$ ، تابع نرخ خطر حاصل می‌شود و آن را با $h(x)$ نمایش می‌دهیم. به عبارت دیگر،

$$h(x) = \lim_{\Delta x \rightarrow \circ} \frac{1}{\Delta x} \frac{F(x + \Delta x) - F(x)}{S(x)} = \frac{f(x)}{S(x)} \quad (8.1)$$

تابع مخاطره تجمعی با انتگرال گیری از رابطه (۸.۱) حاصل می‌شود.

$$H(x) = \int_{\circ}^x h(t) dt = -\log S(x) \quad (9.1)$$

شکل ۴.۱ رابطه بین $f(t)$ ، $S(t)$ و $h(t)$ را نشان می‌دهد.



شکل ۴.۱: رابطه بین $f(t)$ ، $R(t)$ و $r(t)$

۳.۶.۱ میانگین زمان شکست

مقایسه دو متغیر تصادفی طول عمر بر اساس توابع نرخ خطر آن‌ها امکان‌پذیر نیست چرا که نمودار این توابع ممکن است همدیگر را قطع کنند. از این رو به دنبال معیاری هستیم تا بتوانیم

دو متغیر طول عمر را مورد مقایسه قرار دهیم. یکی از این معیارها می‌تواند میانگین باقیمانده عمر و یا در حالت خاص میانگین زمان شکست باشد. از سوی دیگر این سوال اساسی ممکن است پیش بیاید که یک دستگاه که تا زمان x عمر کرده است، به طور متوسط تا چه زمانی به کارش ادامه می‌دهد؟ میانگین باقیمانده عمر پاسخگوی این سوال می‌باشد.

تعریف ۲.۶.۱. میانگین زمان شکست m برای متغیر تصادفی طول عمر T ، به صورت زیر تعریف می‌شود:

$$m = E(T) = \int_0^{\infty} xf(x)dx.$$

قضیه زیر نشان می‌دهد که چگونه با استفاده از تابع قابلیت اعتماد می‌توان $E(T)$ را محاسبه کرد.

قضیه ۱.۶.۱. اگر T متغیر تصادفی پیوسته با تابع قابلیت اعتماد $S(t)$ و تابع چگالی احتمال $f(t)$ باشد آن‌گاه در صورت وجود $E(T)$ ،

$$E(T) = \int_0^{\infty} S(X)dx$$

۷.۱ توزیع لاپلاس

کیفیت روش‌های استفاده شده در تجزیه و تحلیل آماری به شدت بستگی به فرض مدل یا توزیع احتمالی دارد. به خاطر این، سالهاست که در توسعه توزیع‌های استاندارد با روشهای آماری معتبر، که طراحی می‌شوند به عنوان مدل برای طیف گسترده‌ای از پدیده‌های دنیای واقعی تلاش قابل توجهی انجام شده است. با این حال، هنوز بسیاری از مهمات وجود دارد که در آن داده‌های واقعی هیچ کدام از مدل‌های کلاسیک یا استاندارد را دنبال نمی‌کنند. پدیده دنیای واقعی که نیاز به مطالعه آماری دارد، بسیار متداول است. بدین ترتیب، مدل عادی محبوب، یک مدل مفید برای مطالعه هر پدیده نیست. مدل‌های سخت و دشوار که دارای درجه نامتقارن مختلف هستند، اجزای ضروری ابزاری برای ایجاد یک مدل هستند. هدف از مطالعه حاضر بررسی یک توزیع احتمالی است که می‌تواند مورد استفاده قرار گیرد برای مدل‌های مختلف در دنیای واقعی. توزیع احتمالی لاپلاس نشان داده، که این مدل بهتر از مدل‌های موجود برای مدل سازی بعضی از مشکلات دنیای واقعی است. در اینجا یک طرح کلی از این توزیع را مورد بررسی قرار می‌دهیم.

در بخش بعد ما توزیع احتمال لاپلاس و خواص اساسی آن را مطالعه خواهیم کرد. همچنین این مدل را با مدل‌های گاوسی از جمله توزیع نرمال و توزیع کوشی مقایسه خواهیم کرد.

۱.۷.۱ تعاریف و ویژگی‌های اساسی

توزیع لاپلاس

توزیع احتمالی لاپلاس که با $L(\alpha, \lambda)$ نشان می‌دهیم، به صورت زیر تعریف می‌شود:

$$f(x, \alpha, \lambda) = \frac{1}{2\lambda} \exp\left(-\frac{|x - \alpha|}{\lambda}\right); \quad -\infty < x < \infty \quad (10.1)$$

که در آن $\alpha \in (-\infty, \infty)$ پارامتر مکان و $\lambda > 0$ پارامتر مقیاس می‌باشد. این توزیع زمانی که پارامتر مکان برابر میانه باشد حداکثر مقدار خود را خواهد گرفت. توزیع لاپلاس یک توزیع متقارن است که دم‌های آن به شدت و سریعتر از توزیع نرمال و کوشی کاهش می‌یابند. همان طور که می‌دانید توزیع نرمال به صورت زیر است:

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad -\infty < x < \infty \quad (11.1)$$

که در آن $\mu \in (-\infty, \infty)$ و $\sigma^2 > 0$ است. و تابع چگالی توزیع کوشی به فرم زیر می‌باشد.

$$f(x, x_0, \gamma) = \frac{1}{\pi} \left(\frac{\gamma}{\gamma^2 + (x - x_0)^2} \right) \quad -\infty < x < \infty \quad (12.1)$$

که در آن $x_0 \in (-\infty, \infty)$ و $\gamma > 0$ است.

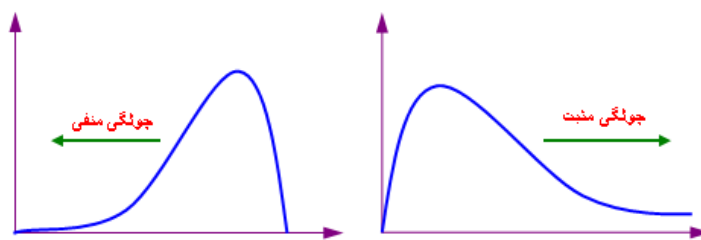
تعریف ۱.۷.۱. در آمار و نظریه احتمالات برجستگی یا کشیدگی توصیف کننده‌ی میزان قله‌ای بودن یک توزیع احتمالی است. هرچقدر شکل تابع چگالی احتمال قله‌ای تر و دارای دنباله پهن تر باشد میزان شاخص کشیدگی برای آن بیشتر است. کشیدگی برابر با گشتاور چهارم نرمال می‌باشد، به عبارت دیگر کشیدگی معیاری از تیزی منحنی در نقطه ماکزیمم است. مقدار کشیدگی برای توزیع نرمال برابر ۳ می‌باشد. و به صورت زیر به دست می‌آید.

$$\gamma_1 = \frac{\mu_4}{\sigma^4}$$

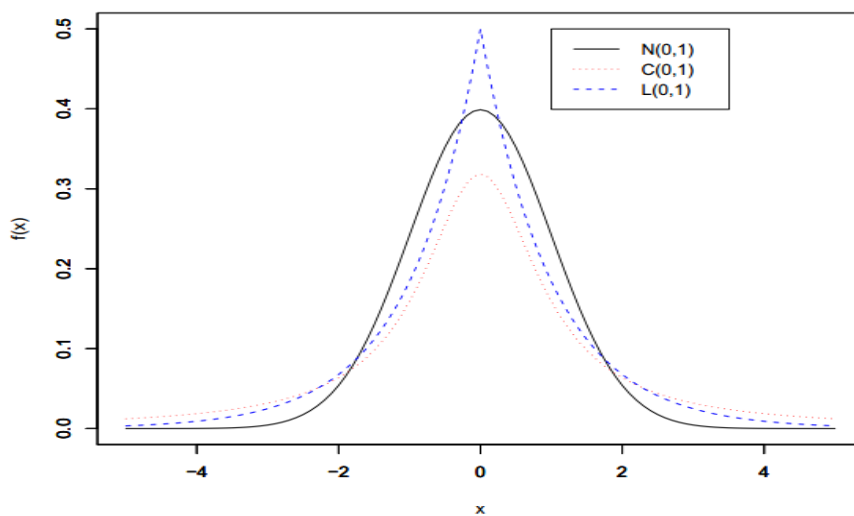
تعریف ۲.۷.۱. چولگی نشان دهنده میزان عدم تقارن توزیع احتمالی است. اگر داده‌ها نسبت به میانگین متقارن باشند، چولگی برابر صفر خواهد بود. به عبارت دیگر چولگی برابر با گشتاور سوم نرمال شده است. چولگی در حقیقت معیاری از وجود یا عدم تقارن تابع توزیع می‌باشد. برای یک توزیع کاملاً متقارن چولگی صفر و برای یک توزیع نامتقارن با کشیدگی به سمت مقادیر بالاتر چولگی مثبت و برای توزیع نامتقارن با کشیدگی به سمت مقادیر کوچکتر مقدار چولگی منفی است.

$$\gamma_2 = \frac{\mu_3}{\sigma^3}$$

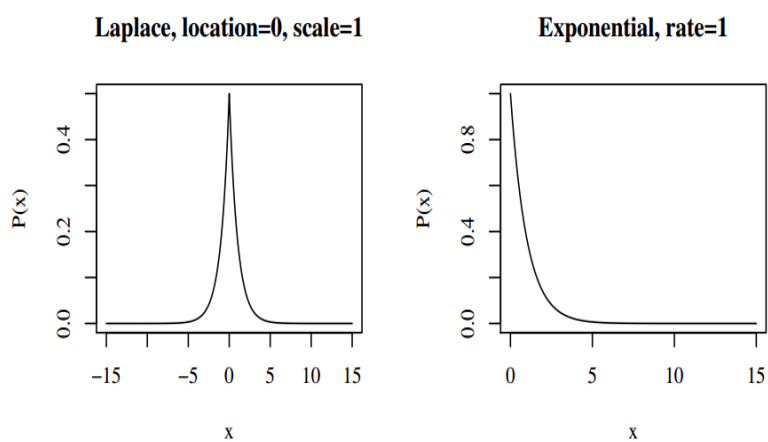
شکل ۶.۱ توزیع چگالی استاندارد توزیع‌های لاپلاس، نرمال و کوشی رو نشان می‌دهد.



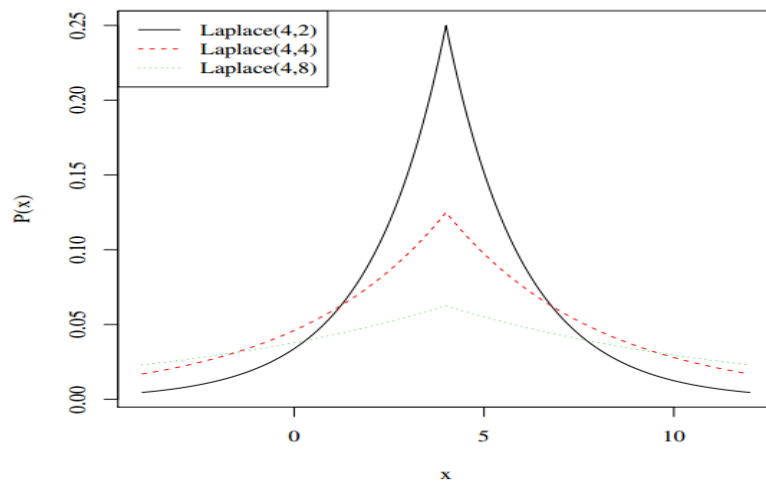
شکل ۵.۱: نمودار چولگی



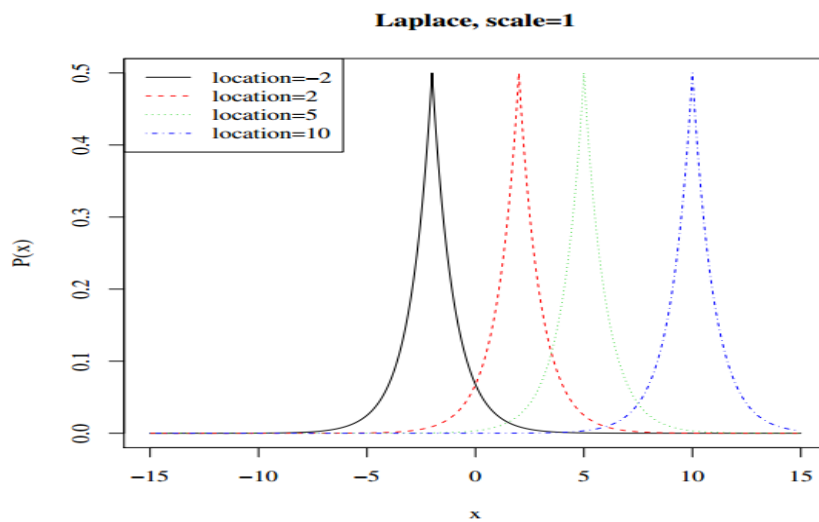
شکل ۶.۱: توزیع چگالی توزیع‌های لاپلاس، نرمال و کوشی



شکل ۷.۱: توزیع چگالی توزیع‌های لاپلاس و نمایی



شکل ۸.۱: توزیع چگالی توزیع لاپلاس



شکل ۹.۱: توزیع چگالی توزیع لاپلاس

جدول ۱.۱ برخی از خواص سه توزیع معرفی شده را نشان می‌دهد.

جدول ۱.۱: برخی از ویژگی‌های توزیع‌های لاپلاس، نرمال و کوشی

$Kur(X)$	$Sk(X)$	$V(x)$	$E(X)$	$f(x)$
۳	۰	σ^2	μ	نرمال
۶	۰	$2\lambda^2$	α	لاپلاس
∞	تعریف نشده	∞	تعریف نشده	کوشی

جدول ۲.۱ مقایسه‌ای از برآورد نمونه‌ای میانگین، میانه‌ی نمونه، $S = \frac{1}{4}(Q_3 - Q_1)$ که در آن Q_1 چارک اول و Q_3 چارک سوم می‌باشد. و همچنین $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ را نشان می‌دهد که در آن $\beta = 1 + \frac{\sigma^2}{n-1}$ است و m میانه را نشان می‌دهد.

جدول ۲.۱: مقایسه‌ای بین برآوردهای نمونه‌ای

$V(S)$	$E(S)$	$V(s^2)$	$E(s^2)$	$V(m)$	$E(m)$	$V(\bar{X})$	$E(\bar{X})$	$f(x)$
$\frac{1}{16nf(Q_1)^2}$	0.6745σ	$\frac{2\sigma^2}{n-1}$	σ^2	$\frac{\pi\sigma^2}{4n}$	μ	$\frac{\sigma^2}{n}$	μ	نرمال
$\frac{\lambda^2}{n}$	$\lambda \ln 2$	$\frac{2\sigma\lambda^2}{n}\beta$	$2\lambda^2$	$\frac{\lambda^2}{n}$	α	$\frac{2\lambda^2}{n}$	α	لاپلاس
$\frac{\pi^2\gamma^2}{4n}$	∞	∞	γ	$\frac{\pi^2\gamma^2}{4n}$	x_0	∞	تعریف نشده	کوشی

تابع توزیع لاپلاس به صورت زیر می‌باشد.

$$F(x, \alpha, \lambda) = \begin{cases} \frac{1}{4} \exp\left(\frac{x-\alpha}{\lambda}\right) & x \leq \alpha \\ 1 - \frac{1}{4} \exp\left(-\frac{x-\alpha}{\lambda}\right) & x \geq \alpha \end{cases} \quad (13.1)$$

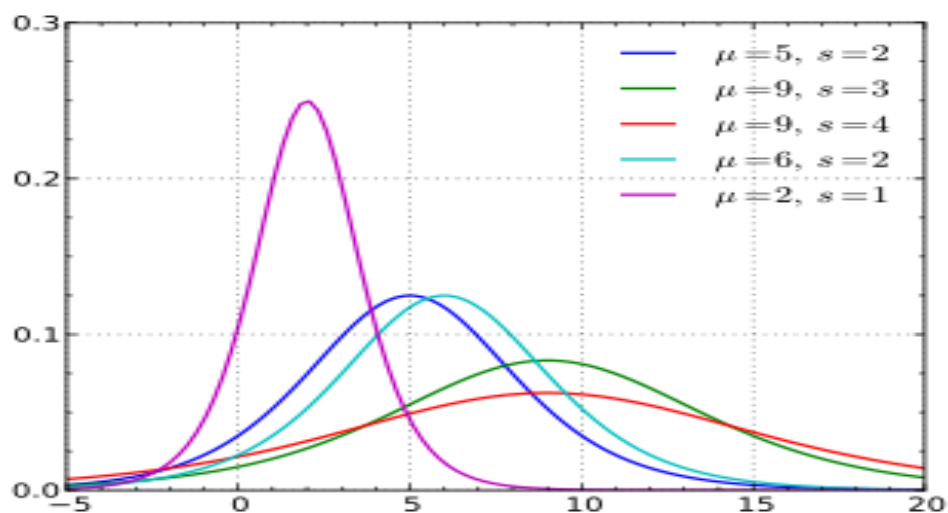
توزیع لوجستیک

توزیع احتمالی لوجستیک به صورت زیر تعریف می‌شود:

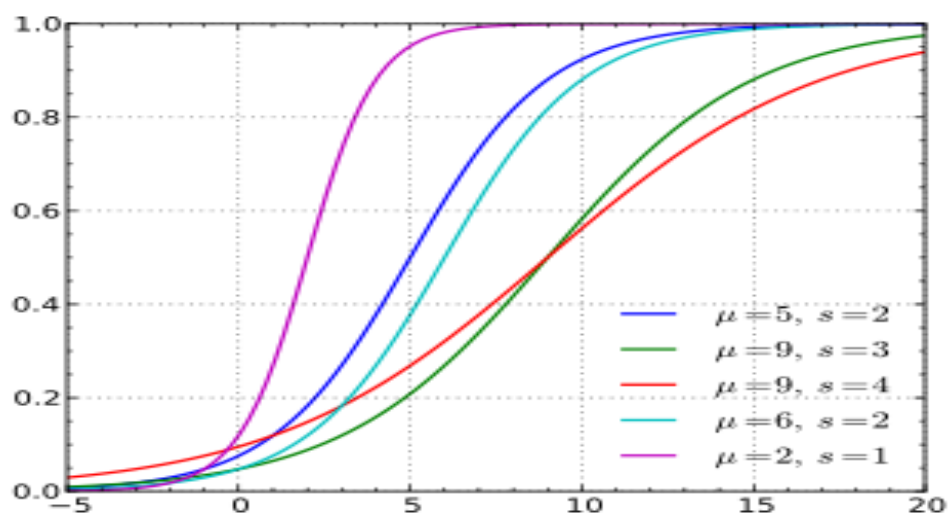
$$f(x, \mu, s) = \frac{\exp\left(-\frac{x-\mu}{s}\right)}{s \left(1 + \exp\left(-\frac{x-\mu}{s}\right)\right)^2} \quad (14.1)$$

که در آن $\alpha \in (-\infty, \infty)$ پارامتر مکان و $\lambda > 0$ پارامتر مقیاس می‌باشد. تابع توزیع لوجستیک به صورت زیر می‌باشد.

$$F(x, \mu, s) = \frac{1}{\left(1 + \exp\left(-\frac{x-\mu}{s}\right)\right)^2} \quad (15.1)$$



شکل ۱۰.۱: تابع چگالی توزیع لوجستیک



شکل ۱۱.۱: تابع توزیع لوجستیک

جدول ۳.۱ برخی از خواص توزیع لوجستیک را نشان می‌دهد.

جدول ۳.۱: برخی از ویژگی‌های توزیع لوجستیک

$Kur(X)$	$Sk(X)$	$V(x)$	$E(X)$	$f(x)$
۱.۲	۰	$\frac{s^2 \pi^2}{3}$	μ	لوجستیک

فصل ۲

برآورد پارامترها در توزیع لاپلاس با داده‌های سانسور فاصله‌ای

۱.۲ مقدمه

در این فصل ابتدا برآورد ماکزیمم درست‌نمایی در توزیع لاپلاس و خواص آنرا مورد بررسی قرار داده‌ایم و سپس در بخش دوم برآورد پارامترهای سانسور فاصله‌ای در توزیع لاپلاس بحث کرده‌ایم. مطالب این بخش عمدتاً از مقاله‌های تومازولا و همکاران^۱ [۲۸]، زو و همکاران^۲ [۳۰]، تیموری و همکاران^۳ [۲۵]، سان^۴ [۲۴]، اسکالن^۵ [۲۲]، نگ و همکاران^۶ [۱۸] و ... گرفته شده است.

^۱Tomazella

^۲Zo

^۳Teimoouri

^۴Sun

^۵Scallan

^۶Ng

۲.۲ برآوردگر ماکزیمم درست‌نمایی

فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی گروه‌بندی شده با $k+1$ فاصله و گامهای $[\tau_{i-1}, \tau_i]$ که $i = 1, 2, \dots, k+1$ است، و $-\infty = \tau_0 < \tau_1 < \tau_2 < \dots < \tau_{k+1} = \infty$ ثابتهای از پیش تعیین شده می‌باشند. همچنین $\tau = \{\tau_0, \tau_1, \tau_2, \dots, \tau_{k+1}\}$ مجموعه‌ای از نقاط پراکنده را نشان می‌دهد و n_i تعداد مشاهدات در فواصل $[\tau_{i-1}, \tau_i]$ که $i = 1, 2, \dots, k+1$ را مشخص می‌نماید. بنابراین لگاریتم تابع درست‌نمایی با توجه به ۱.۱ به صورت زیر می‌باشد

$$\log L(\mu, \sigma) = a_0 + \sum_{i=1}^{k+1} n_i \log [F(\tau_i; \mu, \sigma) - F(\tau_{i-1}; \mu, \sigma)] \quad (1.2)$$

که در آن $F(\cdot; \mu, \sigma)$ تابع توزیع تجمعی، a_0 ثابت و همچنین μ ($-\infty < \mu < \infty$) پارامتر موقعیت و $\sigma > 0$ پارامتر مقیاس می‌باشند. با در نظر گرفتن $\theta_1 = \frac{\mu}{\sigma}$ و $\theta_2 = \frac{1}{\sigma}$ لگاریتم تابع درست‌نمایی در رابطه (۱.۲) را مجدداً به صورت زیر باز پارامتری می‌کنیم.

$$\log L(\theta_1, \theta_2) = a_0 + \sum_{i=1}^{k+1} n_i \log [F(\theta_2 \tau_i - \theta_1) - F(\theta_2 \tau_{i-1} - \theta_1)] \quad (2.2)$$

که در آن $-\infty < \theta_1 < \infty$ و $\theta_2 > 0$ است. همچنین منظور از $F_0(x) = F(x; 0, 1)$ تابع توزیع تجمعی استاندارد $F(\cdot; \mu, \sigma)$ می‌باشد. تابع توزیع تجمعی لاپلاس استاندارد به صورت زیر است.

$$F_0(y) = \begin{cases} 1 - \frac{1}{2} \exp(-y), & y \geq 0 \\ \frac{1}{2} \exp(y), & y \leq 0 \end{cases} \quad (3.2)$$

و تابع چگالی احتمال لاپلاس استاندارد به صورت زیر می‌باشد.

$$f_0(y) = \frac{1}{2} \exp(-|y|) \quad (4.2)$$

قضیه ۱.۲.۲ نشان می‌دهد که برآوردهای ماکزیمم درست‌نمایی (θ_1, θ_2) با ماکزیمم کردن رابطه (۲.۲) وجود دارند و یکتا هستند.

قضیه ۱.۲.۲. فرض کنید که داده‌های گروهی n_1, \dots, n_{k+1} به طوری که $n_1 + n_{k+1} < n$ و $n_{j-1} + n_j < n$ که $2 \leq j \leq k+1$ وجود داشته باشند. آن‌گاه برآوردهای ماکزیمم درست‌نمایی (θ_1, θ_2) ، در نتیجه برآوردهای ماکزیمم درست‌نمایی (μ, σ) وجود دارند و یکتا می‌باشند.

برهان. به منظور نشان دادن وجود و منحصر به فرد بودن برآوردهای (θ_1, θ_2) ، کافی است نشان دهیم.

۱- ماتریس هسیان یک ماتریس نیمه منفی برای هر نقطه‌ی

$$(\theta_1, \theta_2) \in (-\infty, \infty) \times (0, \infty)$$

است و یا حداقل در نقاط مربوط برآوردگرهای ماکزیمم درستنمایی قطعاً منفی است، که درستی این موضوع در لم‌های زیر نشان داده شده است.

۲- همچنین برای هر $\eta > 0$ یک زیر مجموعه فشرده

$$K = K(\eta) \subset (-\infty, \infty) \times (0, \infty)$$

وجود دارد به طوری که $\{(\theta_1, \theta_2) : \log L(\theta_1, \theta_2) \geq -\eta\} \subset K$ و این موضوع در لم ۱.۲.۲ نشان داده شده است.

□

بنابراین اثبات کامل است.

تعریف ۱.۲.۲. در ریاضیات و در بحث ماتریس‌ها و توابع چند متغیره، **ماتریس هسیان** عبارت است از ماتریسی مربعی که شامل مشتقات جزئی مرتبه دوم است. این ماتریس بیانگر میزان انحنای موضعی تابع مورد نظر به ازای متغیرهای آن است. این ماتریس در قرن نوزدهم توسط ریاضیدان آلمانی مطرح و به نام او نامگذاری شد.

لم ۱.۲.۲. فرض کنید ماتریس‌های H_1 و H_{k+1} به صورت زیر تعریف شده باشند.

$$H_1(\theta_1, \theta_2) = \begin{bmatrix} \frac{\partial^2}{\partial \theta_1^2} \log F_0(\theta_2 \tau_1 - \theta_1) & \frac{\partial^2}{\partial \theta_1 \partial \theta_2} \log F_0(\theta_2 \tau_1 - \theta_1) \\ \frac{\partial^2}{\partial \theta_1 \partial \theta_2} \log F_0(\theta_2 \tau_1 - \theta_1) & \frac{\partial^2}{\partial \theta_2^2} \log F_0(\theta_2 \tau_1 - \theta_1) \end{bmatrix}$$

$$H_{k+1}(\theta_1, \theta_2) = \begin{bmatrix} \frac{\partial^2}{\partial \theta_1^2} \log \bar{F}_0(\theta_2 \tau_k - \theta_1) & \frac{\partial^2}{\partial \theta_1 \partial \theta_2} \log \bar{F}_0(\theta_2 \tau_k - \theta_1) \\ \frac{\partial^2}{\partial \theta_1 \partial \theta_2} \log \bar{F}_0(\theta_2 \tau_k - \theta_1) & \frac{\partial^2}{\partial \theta_2^2} \log \bar{F}_0(\theta_2 \tau_k - \theta_1) \end{bmatrix}$$

که در آن $\bar{F}_0(x) = 1 - F_0(x)$ ، بنابراین آن‌ها نیمه منفی هستند.

برهان. برای اثبات نیمه منفی بودن H کافی است دو شرط زیر را نشان دهیم.

۱- درایه‌های روی قطر اصلی ماتریس H منفی باشند.

۲- دترمینان ماتریس H نامنفی است.

فرض کنید $g_1(x) = \log F_0(x)$ باشد. اگر $x \geq 0$

$$g_1(x) = \log F_0(x) = \log \left(1 - \frac{1}{\tau} \exp(-x) \right)$$

$$\frac{dg_1(x)}{dx} = \frac{\frac{1}{\tau} \exp(-x)}{1 - \frac{1}{\tau} \exp(-x)}$$

$$\begin{aligned} \frac{d^2 g_1(x)}{dx^2} &= \frac{(-\frac{1}{\tau} \exp(-x)) \left(1 - \frac{1}{\tau} \exp(-x)\right) - \left(\frac{1}{\tau} \exp(-x)\right) \left(\frac{1}{\tau} \exp(-x)\right)}{\left(1 - \frac{1}{\tau} \exp(-x)\right)^2} \\ &= -2 \exp(-x) (\exp(-x) - \tau)^{-2} < 0 \end{aligned}$$

پس

$$H_1(\theta_1, \theta_2) = \begin{bmatrix} -\frac{2 \exp(\theta_2 \tau_1 - \theta_1)}{[\tau \exp(\theta_2 \tau_1 - \theta_1) - 1]^2} & \frac{2 \tau_1 \exp(\theta_2 \tau_1 - \theta_1)}{[\tau \exp(\theta_2 \tau_1 - \theta_1) - 1]^2} \\ \frac{2 \tau_1 \exp(\theta_2 \tau_1 - \theta_1)}{[\tau \exp(\theta_2 \tau_1 - \theta_1) - 1]^2} & -\frac{2 \tau_1^2 \exp(\theta_2 \tau_1 - \theta_1)}{[\tau \exp(\theta_2 \tau_1 - \theta_1) - 1]^2} \end{bmatrix}$$

بنابراین

$$H_1(\theta_1, \theta_2) = \left[\frac{d^2 g_1(x)}{dx^2} \right]_{x=\theta_2 \tau_1 - \theta_1} \times \begin{pmatrix} 1 & -\tau_1 \\ -\tau_1 & \tau_1^2 \end{pmatrix}$$

بنابراین مشاهده می‌شود درایه‌های قطر اصلی منفی و دترمینان ماتریس H نامنفی است، پس ماتریس H ماتریسی نیمه منفی می‌باشد. حال اگر $x < 0$ آن‌گاه

$$g_1(x) = \log F_0(x) = \log \left(\frac{1}{\tau} \exp(x) \right)$$

$$\frac{dg_1(x)}{dx} = \frac{\frac{1}{\tau} \exp(x)}{\frac{1}{\tau} \exp(x)} = 1$$

$$\frac{d^2 g_1(x)}{dx^2} = 0$$

بنابراین H_1 ماتریس صفر است پس بنابراین $H_1 \leq 0$ و لذا H_1 ماتریس نیمه منفی می‌باشد. به طور مشابه با H_1 می‌توان نشان داد که H_{k+1} نیز ماتریسی نیمه منفی است، زیرا

$$\bar{F}_0(x) = 1 - F_0(x) = \begin{cases} \frac{1}{\tau} \exp(-x) & , x \geq 0 \\ 1 - \frac{1}{\tau} \exp(x) & , x \leq 0 \end{cases}$$

حال فرض کنید $g_2(x) = \log \bar{F}_0(x)$ باشد. اگر $x \geq 0$ ، آنگاه

$$g_2(x) = \log \bar{F}_0(x) = \log \left(\frac{1}{\tau} \exp(x) \right)$$

$$\frac{dg_2(x)}{dx} = \frac{\frac{1}{\tau} \exp(x)}{\frac{1}{\tau} \exp(x)} = 1$$

$$\frac{d^2 g_2(x)}{dx^2} = 0$$

پس H_{k+1} ماتریس صفر است. حال اگر $x < 0$ آن‌گاه

$$g_2(x) = \log \bar{F}_0(x) = \log \left(1 - \frac{1}{\tau} \exp(x) \right)$$

$$\frac{dg_{\tau}(x)}{dx} = \frac{-\frac{1}{\tau} \exp(x)}{1 - \frac{1}{\tau} \exp(x)}$$

$$\begin{aligned} \frac{d^2 g_{\tau}(x)}{dx^2} &= \frac{-(\frac{1}{\tau} \exp(x)) \left(1 - \frac{1}{\tau} \exp(x)\right) - (\frac{1}{\tau} \exp(x))(\frac{1}{\tau} \exp(x))}{\left(1 - \frac{1}{\tau} \exp(x)\right)^2} \\ &= -2 \exp(x) (\tau \exp(x) - 1)^{-2} < 0 \end{aligned}$$

پس

$$H_{k+1}(\theta_1, \theta_{\tau}) = \begin{bmatrix} -\frac{\tau \exp(\theta_{\tau} \tau_k - \theta_1)}{[\tau \exp(\theta_{\tau} \tau_k - \theta_1) - 1]^{\tau}} & \frac{\tau \tau_1 \exp(\theta_{\tau} \tau_k - \theta_1)}{[\tau \exp(\theta_{\tau} \tau_k - \theta_1) - 1]^{\tau}} \\ \frac{\tau \tau_1 \exp(\theta_{\tau} \tau_k - \theta_1)}{[\tau \exp(\theta_{\tau} \tau_k - \theta_1) - 1]^{\tau}} & -\frac{\tau \tau_1^2 \exp(\theta_{\tau} \tau_k - \theta_1)}{[\tau \exp(\theta_{\tau} \tau_k - \theta_1) - 1]^{\tau}} \end{bmatrix}$$

بنابراین

$$H_{k+1}(\theta_1, \theta_{\tau}) = \left[\frac{d^2 g_{\tau}(x)}{dx^2} \right]_{x=\theta_{\tau} \tau_1 - \theta_1} \times \begin{pmatrix} 1 & -\tau_1 \\ -\tau_1 & \tau_1^2 \end{pmatrix}$$

بنابراین مشاهده می‌شود درایه‌های قطر اصلی منفی و دترمینان ماتریس H_{k+1} نامنفی است، پس ماتریس H_{k+1} ماتریسی نیمه منفی می‌باشد. \square

لم ۲.۲.۲. فرض کنید $v < u$ و $g(u, v) = \log[F_{\circ}(u) - F_{\circ}(v)]$ آنگاه ماتریس هسیان

$$H(u, v) = \begin{bmatrix} \frac{\partial^2 g(u, v)}{\partial u^2} & \frac{\partial^2 g(u, v)}{\partial u \partial v} \\ \frac{\partial^2 g(u, v)}{\partial u \partial v} & \frac{\partial^2 g(u, v)}{\partial v^2} \end{bmatrix}$$

ماتریس نیمه منفی است.

برهان. برای اثبات نیمه منفی بودن $H(u, v)$ کافی است دو شرط زیر را نشان دهیم.

$$1 - \frac{\partial^2 g(u, v)}{\partial v^2} \text{ و } \frac{\partial^2 g(u, v)}{\partial u^2} < 0$$

۲- دترمینان ماتریس $H(u, v)$ نامنفی است.

توجه داشته باشید که

$$g(u, v) = \log[F_{\circ}(u) - F_{\circ}(v)] = \begin{cases} \log \left[\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right] & u \geq 0, v \geq 0 \\ \log \left[1 - \frac{1}{\tau} \exp(-u) - \frac{1}{\tau} \exp(v) \right] & u \geq 0, v < 0 \\ \log \left[\frac{1}{\tau} \exp(u) - \frac{1}{\tau} \exp(v) \right] & u < 0, v < 0 \end{cases}$$

در حالت $u \geq 0, v \geq 0$

$$g(u, v) = \log[F_{\circ}(u) - F_{\circ}(v)] = \log \left[\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right]$$

$$\frac{\partial g(u, v)}{\partial u} = \frac{\frac{1}{\tau} \exp(-u)}{\frac{1}{\tau} \exp(-u) - \frac{1}{\tau} \exp(-v)}$$

$$\frac{\partial^2 g(u, v)}{\partial u^2} = \frac{\frac{1}{\tau} \exp(-u) \left(\frac{1}{\tau} \exp(-u) - \frac{1}{\tau} \exp(-v) \right) - \frac{1}{\tau} \exp(-u) \frac{1}{\tau} \exp(-u)}{\left(\frac{1}{\tau} \exp(-u) - \frac{1}{\tau} \exp(-v) \right)^2}$$

$$\begin{aligned} \frac{\partial^2 g(u, v)}{\partial u^2} &= \frac{-\frac{1}{\tau} \exp(-u) \frac{1}{\tau} \exp(-v)}{\left(\frac{1}{\tau} \exp(-u) - \frac{1}{\tau} \exp(-v) \right)^2} \\ &= \frac{-\exp(-u-v)}{(\exp(-u) - \exp(-v))^2} \end{aligned}$$

$$\frac{\partial^2 g(u, v)}{\partial u \partial v} = \frac{\frac{1}{\tau} \exp(-u) \left(\frac{1}{\tau} \exp(-v) \right)}{\left(\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right)^2}$$

$$\begin{aligned} \frac{\partial^2 g(u, v)}{\partial u \partial v} &= \frac{\frac{1}{\tau} \exp(-u) \frac{1}{\tau} \exp(-v)}{\left(\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right)^2} \\ &= \frac{\exp(-u-v)}{(\exp(-v) - \exp(-u))^2} \end{aligned}$$

$$\frac{\partial^2 g(u, v)}{\partial v \partial u} = \frac{\exp(-u-v)}{(\exp(-v) - \exp(-u))^2}$$

$$\frac{\partial g(u, v)}{\partial v} = \frac{-\frac{1}{\tau} \exp(-v)}{\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u)}$$

$$\frac{\partial^2 g(u, v)}{\partial v^2} = \frac{\frac{1}{\tau} \exp(-v) \left(\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right) - \frac{1}{\tau} \exp(-v) \frac{1}{\tau} \exp(-v)}{\left(\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right)^2}$$

$$\begin{aligned} \frac{\partial^2 g(u, v)}{\partial v^2} &= \frac{-\frac{1}{\tau} \exp(-u) \frac{1}{\tau} \exp(-v)}{\left(\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right)^2} \\ &= \frac{-\exp(-u-v)}{(\exp(-v) - \exp(-u))^2} \end{aligned}$$

$$H(u, v) = \begin{bmatrix} -\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} & \frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} \\ \frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} & -\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} \end{bmatrix} = H_{\tau}(u, v)$$

بنابراین داریم $\frac{\partial^2 g(u,v)}{\partial u^2} < 0$ و $\frac{\partial^2 g(u,v)}{\partial v^2} < 0$ و همچنین دترمینان ماتریس $H_{\Psi}(u,v)$ صفر است زیرا

$$|H_{\Psi}(u,v)| = \left(-\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} \right)^2 - \left(\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^2} \right)^2 = 0$$

در حالت $u \geq 0, v < 0$

$$g(u,v) = \log[F_0(u) - F_0(v)] = \log \left[1 - \frac{1}{\Psi} \exp(-u) - \frac{1}{\Psi} \exp(v) \right]$$

$$\frac{\partial g(u,v)}{\partial u} = \frac{\frac{1}{\Psi} \exp(-u)}{1 - \frac{1}{\Psi} \exp(-u) - \frac{1}{\Psi} \exp(v)}$$

$$\frac{\partial g(u,v)}{\partial v} = \frac{-\frac{1}{\Psi} \exp(v)}{1 - \frac{1}{\Psi} \exp(-u) - \frac{1}{\Psi} \exp(v)}$$

و با گرفتن مشتقات جزئی مرتبه دوم داریم:

$$H(u,v) = \begin{bmatrix} \frac{\exp(u)[-2+\exp(v)]}{\{1+\exp(u)[-2+\exp(v)]\}^2} & \frac{\exp(u+v)}{\{1+\exp(u)[-2+\exp(v)]\}^2} \\ \frac{\exp(u+v)}{\{1+\exp(u)[-2+\exp(v)]\}^2} & \frac{\exp(u+v)[1-2\exp(u)]}{\{1+\exp(u)[-2+\exp(v)]\}^2} \end{bmatrix} = H_{\Psi}(u,v)$$

پس $\frac{\partial^2 g(u,v)}{\partial u^2} < 0$ و $\frac{\partial^2 g(u,v)}{\partial v^2} < 0$ و دترمینان ماتریس $H_{\Psi}(u,v) > 0$ چون

$$|H_{\Psi}(u,v)| = \left(\frac{\exp(u)[-2+\exp(v)]}{\{1+\exp(u)[-2+\exp(v)]\}^2} \frac{\exp(u+v)[1-2\exp(u)]}{\{1+\exp(u)[-2+\exp(v)]\}^2} \right) - \left(\frac{\exp(u+v)}{\{1+\exp(u)[-2+\exp(v)]\}^2} \frac{\exp(u+v)}{\{1+\exp(u)[-2+\exp(v)]\}^2} \right)$$

$$= \frac{\exp(u)[-2+\exp(v)] \times \exp(u+v)[1-2\exp(u)] - \exp(u+v)^2}{\{1+\exp(u)[-2+\exp(v)]\}^4} > 0$$

و در حالت $u < 0, v < 0$

$$H(u,v) = \begin{bmatrix} -\frac{\exp(u+v)}{[\exp(v)-\exp(u)]^2} & \frac{\exp(u+v)}{[\exp(v)-\exp(u)]^2} \\ \frac{\exp(u+v)}{[\exp(v)-\exp(u)]^2} & -\frac{\exp(u+v)}{[\exp(v)-\exp(u)]^2} \end{bmatrix} = H_{\Phi}(u,v)$$

لذا $\frac{\partial^2 g(u,v)}{\partial v^2} < 0$ و $\frac{\partial^2 g(u,v)}{\partial u^2} < 0$ و دترمینان ماتریس $H_{\Phi}(u,v)$ برابر صفر است، زیرا

$$|H_{\Phi}(u,v)| = \left(-\frac{\exp(u+v)}{[\exp(v)-\exp(u)]^2} \right)^2 - \left(\frac{\exp(u+v)}{[\exp(v)-\exp(u)]^2} \right)^2 = 0$$

□ اثبات کامل است.

لم ۳.۲.۲. فرض کنید $g_i(\theta_1, \theta_2) = \log[F_0(\theta_2 \tau_i - \theta_1) - F_0(\theta_2 \tau_{i-1} - \theta_1)]$ و H_i برای هر $2 \leq i \leq k$ ماتریس هسیان باشد، آن‌گاه H_i ماتریس نیمه معین منفی است.

برهان. می‌دانیم

$$g_i(u, v) = \log[F_\circ(u) - F_\circ(v)] = \begin{cases} \log \left[\frac{1}{\tau} \exp(-v) - \frac{1}{\tau} \exp(-u) \right] & u \geq \circ, v \geq \circ \\ \log \left[1 - \frac{1}{\tau} \exp(-u) - \frac{1}{\tau} \exp(v) \right] & u \geq \circ, v < \circ \\ \log \left[\frac{1}{\tau} \exp(u) - \frac{1}{\tau} \exp(v) \right] & u < \circ, v < \circ \end{cases}$$

که در آن $u = \theta_\tau \tau_i - \theta_1$ و $v = \theta_\tau \tau_{i-1} - \theta_1$ می‌باشد. حال با مشتق‌گیری نسبت به θ_1 و θ_τ ، به آسانی می‌توان نشان داد

$$H_\tau(u, v) = \begin{pmatrix} 1 - \tau_i & -\tau_i + \tau_i \tau_{i-1} \\ 1 - \tau_{i-1} & -\tau_i + (\tau_{i-1})^2 \end{pmatrix} \begin{pmatrix} -\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} & \frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} \\ \frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} & -\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} \end{pmatrix}$$

$$H_\tau(u, v) = \begin{bmatrix} -1 & \tau_i \\ -1 & \tau_{i-1} \end{bmatrix}' \begin{pmatrix} -\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} & \frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} \\ \frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} & -\frac{\exp(-u-v)}{[\exp(-v) - \exp(-u)]^\tau} \end{pmatrix} \begin{bmatrix} -1 & \tau_i \\ -1 & \tau_{i-1} \end{bmatrix}$$

از آنجایی که $\frac{\partial^\tau g(u, v)}{\partial u^\tau} < \circ$ و $\frac{\partial^\tau g(u, v)}{\partial v^\tau} < \circ$ و همچنین دترمینان ماتریس $H_\tau(u, v)$ صفر است پس $H_i(u, v) \equiv A' H(u, v) A$ می‌توان نشان داد برای هر i که H_i با H_τ مشابه است. و به طور مشابه با H_τ برای هر i می‌توان نشان داد

$$A = \begin{bmatrix} -1 & \tau_i \\ -1 & \tau_{i-1} \end{bmatrix}$$

□

و مشابه لم ۲.۲.۲ H_i ماتریس نیمه منفی است.

لم ۴.۲.۲. فرض کنید $n_1 + n_{k+1} < n$ آن‌گاه

$$\lim_{\theta_\tau \rightarrow \circ^+} \sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_\tau) = -\infty \quad (5.2)$$

برهان. با توجه به اینکه $n_1 + n_{k+1} < n$ وجود دارد $2 \leq i \leq k+1$ به طوری که $n_i > \circ$ با توجه به معادله (۴.۲) داریم $1 < f_\circ(t) \leq \circ/5$ برای همه $t \in (-\infty, \infty)$ بنابراین زمانی که $\theta_\tau \rightarrow \circ^+$ و $\theta_1 \rightarrow +\infty$ داریم

$$\log \left[\int_{\theta_\tau \tau_{i-1} - \theta_1}^{\theta_\tau \tau_i - \theta_1} f_\circ(t) dt \right] = \log \left[\frac{1}{\tau} \exp(\theta_\tau \tau_{i-1} - \theta_1) - \frac{1}{\tau} \exp(\theta_\tau \tau_i - \theta_1) \right]$$

9

$$\log L(\theta_1, \theta_\tau) \leq n_i \log \int_{\theta_\tau \tau_{i-1} - \theta_1}^{\theta_\tau \tau_i - \theta_1} f_\circ(t) dt$$

$$\leq n_i \log \left[\frac{1}{\tau} \exp(\theta_\tau \tau_{i-1} - \theta_1) - \frac{1}{\tau} \exp(\theta_\tau \tau_i - \theta_1) \right]$$

بنابراین

$$\begin{aligned} \sup_{\theta_1 \rightarrow \infty} L(\theta_1, \theta_2) &\leq \sup_{\theta_1 \rightarrow \infty} n_i \log \left[\frac{1}{\varphi} \exp(\theta_2 \tau_{i-1} - \theta_1) - \frac{1}{\varphi} \exp(\theta_2 \tau_i - \theta_1) \right] \\ &= \sup_{\theta_1 \rightarrow \infty} n_i \log \left(\exp(-\theta_1) \left[\frac{1}{\varphi} \exp(\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(\theta_2 \tau_i) \right] \right) \\ &= \sup_{\theta_1 \rightarrow \infty} \left(n_i \log \exp(-\theta_1) + \log \left[\frac{1}{\varphi} \exp(\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(\theta_2 \tau_i) \right] \right) \\ &\leq \sup_{\theta_1 \rightarrow \infty} (-n_i \theta_1) + n_i \log \left[\frac{1}{\varphi} \exp(\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(\theta_2 \tau_i) \right] \end{aligned}$$

بنابراین برای هر $\theta_1 \rightarrow \infty$ و $\theta_2 \rightarrow \circ^+$ داریم.

$$\begin{aligned} &\sup_{\substack{\theta_1 \rightarrow \infty \\ \theta_2 \rightarrow \circ^+}} (-n_i \theta_1) + n_i \log \left[\frac{1}{\varphi} \exp(\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(\theta_2 \tau_i) \right] \\ &= \sup_{\theta_1 \rightarrow \infty} (-n_i \theta_1) + n_i \log \left[\frac{1}{\varphi} \exp(\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(\theta_2 \tau_i) \right] = -\infty \end{aligned}$$

و به طور مشابه زمانی که $\theta_1 \rightarrow -\infty$ و $\theta_2 \rightarrow \circ^+$ داریم.

$$\log \left[\int_{\theta_2 \tau_{i-1} - \theta_1}^{\theta_2 \tau_i - \theta_1} f_{\circ}(t) dt \right] = \log \left[\frac{1}{\varphi} \exp(-[\theta_2 \tau_{i-1} - \theta_1]) - \frac{1}{\varphi} \exp(-[\theta_2 \tau_i - \theta_1]) \right]$$

و

$$\begin{aligned} \log L(\theta_1, \theta_2) &\leq n_i \log \int_{\theta_2 \tau_{i-1} - \theta_1}^{\theta_2 \tau_i - \theta_1} f_{\circ}(t) dt \\ &\leq n_i \log \left[\frac{1}{\varphi} \exp(-[\theta_2 \tau_{i-1} - \theta_1]) - \frac{1}{\varphi} \exp(-[\theta_2 \tau_i - \theta_1]) \right] \end{aligned}$$

بنابراین

$$\begin{aligned} \sup_{\theta_1 \rightarrow \infty} L(\theta_1, \theta_2) &\leq \sup_{\theta_1 \rightarrow \infty} n_i \log \left[\frac{1}{\varphi} \exp(-[\theta_2 \tau_{i-1} - \theta_1]) - \frac{1}{\varphi} \exp(-[\theta_2 \tau_i - \theta_1]) \right] \\ &= \sup_{\theta_1 \rightarrow \infty} n_i \log \left(\exp(\theta_1) \left[\frac{1}{\varphi} \exp(-\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(-\theta_2 \tau_i) \right] \right) \\ &= \sup_{\theta_1 \rightarrow \infty} \left(n_i \log \exp(\theta_1) + \log \left[\frac{1}{\varphi} \exp(-\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(-\theta_2 \tau_i) \right] \right) \\ &\leq \sup_{\theta_1 \rightarrow \infty} (n_i \theta_1) + n_i \log \left[\frac{1}{\varphi} \exp(-\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(-\theta_2 \tau_i) \right] \end{aligned}$$

بنابراین برای هر $\theta_1 \rightarrow -\infty$ و $\theta_2 \rightarrow \circ^+$ داریم.

$$\begin{aligned} &\sup_{\substack{\theta_1 \rightarrow -\infty \\ \theta_2 \rightarrow \circ^+}} (n_i \theta_1) + n_i \log \left[\frac{1}{\varphi} \exp(-\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(-\theta_2 \tau_i) \right] \\ &= \sup_{\theta_1 \rightarrow -\infty} (n_i \theta_1) + n_i \log \left[\frac{1}{\varphi} \exp(-\theta_2 \tau_{i-1}) - \frac{1}{\varphi} \exp(-\theta_2 \tau_i) \right] = -\infty \end{aligned}$$

□

معادله (۵.۲) برقرار است.

لم ۵.۲.۲. فرض کنید $n_1 + n_{k+1} < n$ و $n_{j-1} + n_j < n$ برای تمام $1 \leq j \leq k+1$. آنگاه

$$\lim_{\theta_2 \rightarrow \infty} \sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_2) = -\infty \quad (۶.۲)$$

و

$$\lim_{|\theta_1| \rightarrow \infty} \sup_{\theta_2 > 0} \log L(\theta_1, \theta_2) = -\infty \quad (۷.۲)$$

برهان. فرض کنید $I = \{1 \leq j \leq k+1; n_j > 0\}$. برای هر $\theta_2 > 0$ ثابت واضح است که

$$\log L(\theta_1, \theta_2) \leq n_i \log \int_{\theta_2 \tau_{i-1} - \theta_1}^{\theta_2 \tau_i - \theta_1} f_0(t) dt \equiv M(\theta_1)$$

با استفاده از لم ۴.۲.۲ داریم

$$\sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_2) = \sup_{-\infty < \theta_1 < \infty} M(\theta_1) = -\infty$$

بنابراین برای هر θ_2 ثابت وجود دارد $\theta_1^* = \theta_1^*(\theta_2)$ به طوری که

$$\sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_2) = \log L(\theta_1^*, \theta_2) = \sum_{i \in I} n_i \log \int_{\theta_2 \tau_{i-1} - \theta_1^*}^{\theta_2 \tau_i - \theta_1^*} f_0(t) dt$$

فرض کنید $2C = \min_{1 \leq j \leq k+1} (\tau_j - \tau_{j-1}) > 0$ و همچنین $A > 0$ ثابت باشد. برای هر A حداقل یک i از مجموعه I وجود دارد به طوری که

$$1 - \theta_2 \tau_{i-1} - \theta_1^* = \theta_2 (\tau_{i-1} - \theta_1^*/\theta_2) \text{ و } \theta_2 \tau_i - \theta_1^* = \theta_2 (\tau_i - \theta_1^*/\theta_2) \text{ هم علامتند.}$$

$$2 - |\tau_{i-1} - \theta_1^*/\theta_2| \geq C \text{ و } |\tau_i - \theta_1^*/\theta_2| \geq C$$

اگر $i \in I$ باشد پس $\theta_2 \tau_i - \theta_1^* > 0$ و $\theta_2 \tau_{i-1} - \theta_1^* > 0$ می‌باشند، آنگاه داریم.

$$\int_{\theta_2 \tau_{i-1} - \theta_1^*}^{\theta_2 \tau_i - \theta_1^*} f_0(t) dt = \left[\frac{1}{\theta_2} \exp(\theta_2 \tau_i - \theta_1^*) - \frac{1}{\theta_2} \exp(\theta_2 \tau_{i-1} - \theta_1^*) \right]$$

و

$$\int_{\theta_2 \tau_{i-1} - \theta_1^*}^{\theta_2 \tau_i - \theta_1^*} \exp(-t) dt = \exp[-(\theta_2 \tau_{i-1} - \theta_1^*)] - \exp[-(\theta_2 \tau_i - \theta_1^*)]$$

بنابراین

$$\int_{\theta_2 \tau_{i-1} - \theta_1^*}^{\theta_2 \tau_i - \theta_1^*} f_0(t) dt < \int_{\theta_2 \tau_{i-1} - \theta_1^*}^{\theta_2 \tau_i - \theta_1^*} \exp(-t) dt < \exp[-(\theta_2 \tau_{i-1} - \theta_1^*)]$$

پس برای هر $A > 0$ و تمام $\theta_2 > AC$ داریم $-(\theta_2 \tau_{i-1} - \theta_1^*) < -CA/C = A$ و

$$n_i \log \int_{\theta_2 \tau_{i-1} - \theta_1}^{\theta_2 \tau_i - \theta_1} f_0(t) dt < -n_i A \leq A$$

بنابراین

$$\log L(\theta_1^*, \theta_2) = \sum_{i \in I} n_i \log \int_{\theta_2 \tau_{i-1} - \theta_1}^{\theta_2 \tau_i - \theta_1} f_{\circ}(t) dt < -A \quad (8.2)$$

برای تمام $\theta_2 > AC$ در نامعادله ۸.۲ به طور مشابه می‌توان نشان داد که

$$(\theta_2 \tau_{i-1} - \theta_1^*) < -C$$

و

$$(\theta_2 \tau_i - \theta_1^*) < -C$$

بنابراین برای هر A به اندازه کافی بزرگ داده شده داریم

$$\sup_{-\infty < \theta_1 < \infty} \log L(\theta_1, \theta_2) = \log L(\theta_1^*, \theta_2) < -A$$

و این همان معادله‌ی ۶.۲ می‌باشد. و به طور مشابه می‌توان نشان داد $\sup_{\theta_2 > 0} \log L(\theta_1, \theta_2) < -A$ برای هر A به اندازه کافی بزرگ و $|\theta_1|$ بزرگ، معادله (۷.۲) برقرار است. □

۶.۲.۲. فرض کنید که $n_1 + n_{k+1} < n$ و $n_{j-1} + n_j < n$ و $2 \leq j \leq k+1$ برای هر η ثابت ($\eta > 0$) مجموعه فشرده $k \equiv k(\eta) \subset (-\infty, \infty) \times (0, \infty)$ وجود دارد به طوری که

$$\{(\theta_1, \theta_2) : \log L(\theta_1, \theta_2) \leq -\eta\} \subset K$$

برهان. این لم نتیجه لم‌های ۵.۲.۲ و ۴.۲.۲ است. □

۳.۲ تحلیل پارامتری داده‌های بقای سانسور فاصله‌ای در توزیع لاپلاس

یکی از مدل‌های آماری که در دهه‌های اخیر بسیار مورد توجه قرار گرفته و کاربرد بسیاری در مطالعات پزشکی، مهندسی و سایر علوم دارد، تحلیل بقاست. تحلیل بقا شامل کلیه روش‌های آماری برای تحلیل داده‌های بقاست. ویژگی اصلی داده‌های بقا در مقایسه با داده‌های آماری وجود داده‌های سانسور می‌باشد. یکی از انواع پیچیده سانسور حالتی است که زمان دقیق رخداد پیشامد مشخص نیست و فقط این اطلاع موجود است که پیشامد در یک بازه زمانی اتفاق افتاده است که به آن سانسور فاصله‌ای می‌گویند. در تحلیل داده‌های بقا به طور کلی از دو مدل رگرسیونی استفاده مدل مخاطرات متناسب کاکس به عنوان یک مدل نیمه پارامتری و مدل‌های زمان شکست شتابنده به عنوان مدل

پارامتری. اگرچه رگرسیون کاکس کاربردی‌ترین مدل در در تحلیل بقاست، و اما مدل‌های پارامتری در برخی شرایط می‌توانند مناسب تر باشند. یکی از شرایط انعطاف پذیر بودن این مدل‌ها این است که پارامتر مقیاس و مکان هر دو می‌توانند تابعی از متغیرهای کمکی باشند. به دست آوردن توزیع پارامتری زمان بقا برخلاف مدل کاکس از اهمیت ویژه‌ای برخوردار است به طوری که می‌توان احتمال بقا و احتمال خطر را در تمام زمان‌ها محاسبه نمود و در برآورد پارامترها بر خلاف مدل کاکس از روش حداکثر درست‌نمایی استفاده کرد. همچنین نسبت به رگرسیون کاکس پارامترهای کمتری دارند و مدل ساده‌تری نتیجه می‌دهد. اکثر توزیع‌های پارامتری مانند لاپلاس دارای دو پارامتر مکان و مقیاس می‌باشند، در حالت استاندارد فرض می‌شود پارامتر مکان ثابت و پارامتر مقیاس تابعی از متغیرهای کمکی است. در این بخش برآورد توزیع لاپلاس در حالت سانسور فاصله‌ای هنگامی که هر دو پارامتر ثابت نیستند ارائه شده است.

۱.۳.۲ توزیع لاپلاس

توزیع لاپلاس دارای دو پارامتر α مکان و λ مقیاس می‌باشد. تابع چگالی توزیع لاپلاس با پارامترهای (α, λ) به صورت زیر بیان می‌شود:

$$f(t, \alpha, \lambda) = \frac{1}{2\lambda} \exp\left(-\frac{|t-\alpha|}{\lambda}\right); \quad t \in (-\infty, \infty), \alpha \in (-\infty, \infty), \lambda > 0 \quad (9.2)$$

همچنین توابع خطر $(h(t))$ و بقاء $(s(t))$ برای این توزیع به صورت زیر است

$$s(t) = \begin{cases} 1 - \frac{1}{2} \exp\left(\frac{t-\alpha}{\lambda}\right) & t \leq \alpha \\ \frac{1}{2} \exp\left(-\frac{t-\alpha}{\lambda}\right) & t \geq \alpha \end{cases}$$

$$h(t) = \begin{cases} \frac{\frac{1}{2\lambda} \exp\left(\frac{t-\alpha}{\lambda}\right)}{1 - \frac{1}{2} \exp\left(\frac{t-\alpha}{\lambda}\right)} & t \leq \alpha \\ \frac{1}{\lambda} & t \geq \alpha \end{cases}$$

در تحلیل پارامتری برای استنباط در مورد متغیرهای کمکی و برازش دقیق‌تر، فرض می‌شود پارامتر مقیاس با متغیرهای کمکی مرتبط است. برای این منظور پارامتر مقیاس را به صورت زیر با متغیرهای کمکی مرتبط می‌کنند.

فصل ۳

شبیه سازی

۱.۰.۳ مقدمه

در این فصل با استفاده از یک فرایند شبیه سازی مونت کارلو به بررسی و ارزیابی عملکرد و مقایسه برآوردهای ماکزیمم درستنمایی پارامترهای μ و σ در دو مدل لاپلاس و لوجستیک با داده های فاصله ای خواهیم پرداخت. ابتدا هزار نمونه تصادفی با اندازه $n = 20, 25, 30, 35, \dots, 100$ از توزیع لاپلاس استاندارد ($\alpha = 0, \lambda = 1$) شبیه سازی کرده ایم. و سپس برای هر ۱۰۰۰ نمونه برآورد ماکزیمم درستنمایی پارامترهای α و λ در توزیع لاپلاس محاسبه نموده و معیارهای اریبی، میانگین مربعات خطا، چولگی و کشیدگی را بدست آورده ایم. همچنین موارد بالا را برای توزیع لوجستیک زمانی که $\mu = 0$ و $\sigma = 6/\sqrt{\pi}$ مورد بررسی قرار داده ایم. همچنین مقدار τ را به صورت $(-\infty, 0, 0/5, 1, 1/5, 2, \infty)$ ، $(-\infty, -0/5, 0/25, 1, 1/75, 2/5, \infty)$ ، $(-\infty, -0/25, 0, 1, 2, 2/5, \infty)$ و $(-\infty, -2, -1, 0/5, 1, 3, \infty)$ انتخاب کرده ایم. پارامترهای توزیع لاپلاس و لوجستیک به گونه ای ساخته شده اند که آنها دارای میانگین و واریانس برابر هستند. در ادامه در قالب چند مثال این موارد را مورد بررسی قرار می دهیم. مطالب این فصل عمدتاً از مقاله های تومازولا و همکاران^۱ [۲۸]، سایت [۲۱]، پتلو و همکاران^۲ [۱۹]، لیو و همکاران^۳ [۱۶] و ... گرفته شده است.

^۱Tomazella

^۲Patel

^۳Lio

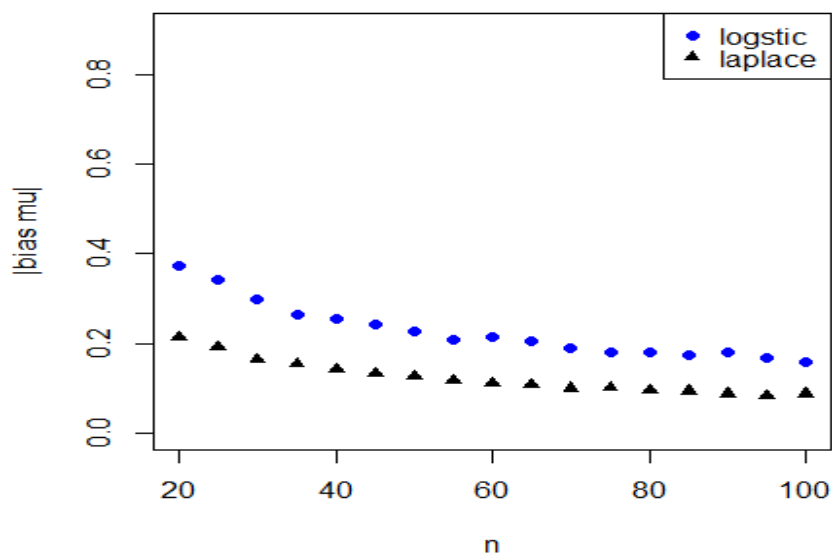
مثال ۱.۰.۳. فرض کنید که مقدار $\tau = (-\infty, 0, 0.5, 1, 1.5, 2, \infty)$ باشد. معیارهای قدرمطلق اریبی، میانگین مربعات خطا، چولگی و کشیدگی برای دو مدل لاپلاس و لوجستیک به صورت زیر می‌باشند.

جدول ۱.۳: نتایج شبیه سازی در مدل لاپلاس

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۶/۶۷	۴/۱۶	۱/۴۹	-۰/۰۱	۰/۲۳	۰/۰۸	۰/۳۴	۰/۲۱	۱
۶/۳۷	۳/۵۸	۱/۲۹	-۰/۱۲	۰/۱۵	۰/۰۶	۰/۲۸	۰/۱۹	۲
۱۹/۲۲	۴/۳۹	۲/۳۲	-۰/۲۷	۰/۱۴	۰/۰۵	۰/۲۶	۰/۱۶	۳
۴/۲۸	۳/۶۴	۰/۸۰	-۰/۰۱	۰/۰۹	۰/۰۴	۰/۲۳	۰/۱۵	۴
۷/۱۵	۳/۷۳	۱/۲۷	-۰/۲۱	۰/۰۸	۰/۰۳	۰/۲۱	۰/۱۴	۵
۵/۹۶	۳/۵۳	۱/۰۱	-۰/۰۱	۰/۰۷	۰/۰۳	۰/۲۰	۰/۱۳	۶
۴/۹۵	۳/۰۴	۰/۸۶	۰/۰۴	۰/۰۶	۰/۰۳	۰/۱۹	۰/۱۳	۷
۳/۸۹	۳/۷۴	۰/۶۴	-۰/۰۸	۰/۰۵	۰/۰۲	۰/۱۷	۰/۱۲	۸
۳/۵۲	۳/۴۵	۰/۵۷	۰/۰۶	۰/۰۵	۰/۰۲	۰/۱۸	۰/۱۱	۹
۴/۰۶	۳/۷۸	۰/۶۸	-۰/۱۷	۰/۰۴	۰/۰۲	۰/۱۶	۰/۱۱	۱۰
۴/۰۲	۴/۰۸	۰/۷۲	۰/۱۷	۰/۰۴	۰/۰۲	۰/۱۵	۰/۱۰	۱۱
۳/۵۵	۳/۵۸	۰/۵۷	۰/۱۰	۰/۰۴	۰/۰۲	۰/۱۵	۰/۱۰	۱۲
۳/۴۸	۳/۵۹	۰/۵۶	۰/۰۱	۰/۰۳	۰/۰۲	۰/۱۴	۰/۱۰	۱۳
۳/۹۶	۳/۰۵	۰/۴۴	-۰/۰۲	۰/۰۳	۰/۰۱	۰/۱۳	۰/۰۹	۱۴
۳/۵۲	۳/۵۱	۰/۵۲	۰/۱۱	۰/۰۳	۰/۰۱	۰/۱۳	۰/۰۹	۱۵
۳/۸۵	۳/۷۰	۰/۵۸	-۰/۱۶	۰/۰۳	۰/۰۱	۰/۱۳	۰/۰۸	۱۶
۳/۲۷	۳/۰۵	۰/۴۷	-۰/۱۵	۰/۰۳	۰/۰۱	۰/۱۳	۰/۰۹	۱۷

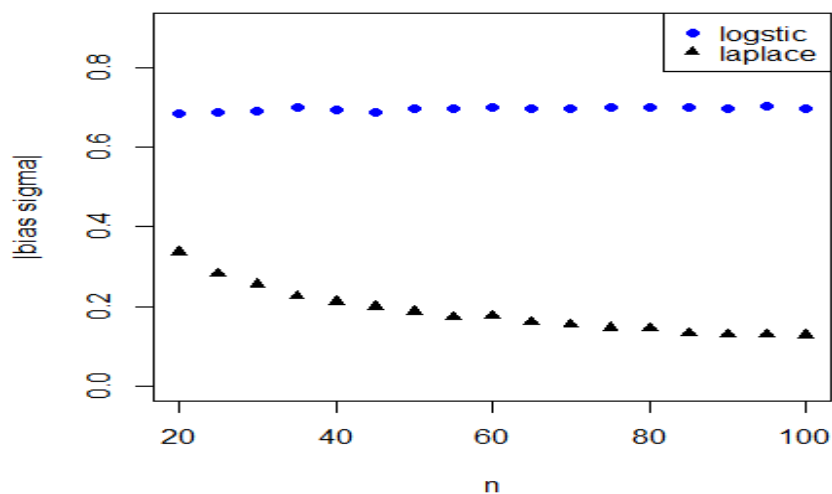
جدول ۲.۳: نتایج شبیه سازی در مدل لوجستیک

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۳/۳۲	۳/۵۳	۰/۵۵	۰/۱۱	۰/۵۰	۰/۲۳	۰/۶۸	۰/۳۷	۱
۴/۰۳	۲/۹۹	۰/۶۴	-۰/۰۱	۰/۵۰	۰/۱۹	۰/۶۹	۰/۳۴	۲
۳/۰۰	۳/۳۰	۰/۴۰	۰/۰۱	۰/۵۰	۰/۱۵	۰/۶۹	۰/۳۰	۳
۳/۱۶	۳/۳۳	۰/۴۳	-۰/۱۶	۰/۵۱	۰/۱۱	۰/۷۰	۰/۲۶	۴
۲/۹۱	۳/۳۳	۰/۳۶	-۰/۱۶	۰/۵۰	۰/۱۱	۰/۶۹	۰/۲۵	۵
۳/۱۱	۳/۲۶	۰/۳۲	-۰/۱۲	۰/۴۹	۰/۱۱	۰/۶۹	۰/۲۴	۶
۲/۹۶	۳/۰۲	۰/۲۷	-۰/۰۸	۰/۵۰	۰/۰۸	۰/۷۰	۰/۲۳	۷
۳/۱۹	۳/۲۰	۰/۳۹	۰/۰۶	۰/۵۰	۰/۰۷	۰/۷۰	۰/۲۱	۸
۳/۰۷	۳/۰۱	۰/۲۶	-۰/۰۸	۰/۵۰	۰/۰۷	۰/۷۰	۰/۲۱	۹
۳/۱۲	۳/۱۰	۰/۳۸	۰/۰۹	۰/۴۹	۰/۰۷	۰/۷۰	۰/۲۱	۱۰
۳/۴۰	۳/۰۹	۰/۳۳	-۰/۰۸	۰/۵۰	۰/۰۶	۰/۷۰	۰/۱۹	۱۱
۳/۳۵	۳/۰۶	۰/۴۴	۰/۰۴	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۲
۳/۱۷	۳/۰۰	۰/۲۴	۰/۰۸	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۳
۳/۳۰	۳/۲۰	۰/۲۹	-۰/۰۲	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۴
۳/۱۸	۲/۹۶	۰/۳۸	-۰/۰۲	۰/۴۹	۰/۰۵	۰/۷۰	۰/۱۸	۱۵
۳/۰۹	۲/۹۸	۰/۲۴	-۰/۱۲	۰/۵۰	۰/۰۴	۰/۷۰	۰/۱۷	۱۶
۲/۸۴	۳/۰۵	۰/۱۲	-۰/۰۳	۰/۴۹	۰/۰۴	۰/۷۰	۰/۱۶	۱۷



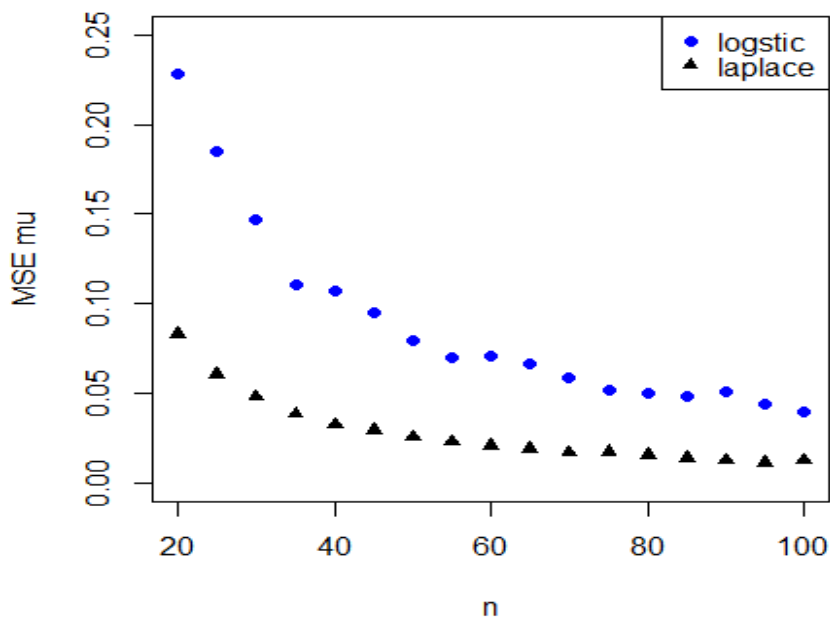
شکل ۱.۳: نمودار قدر مطلق آریبی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۱.۳ نشان می‌دهد که مقدار قدر مطلق آریبی میانگین μ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



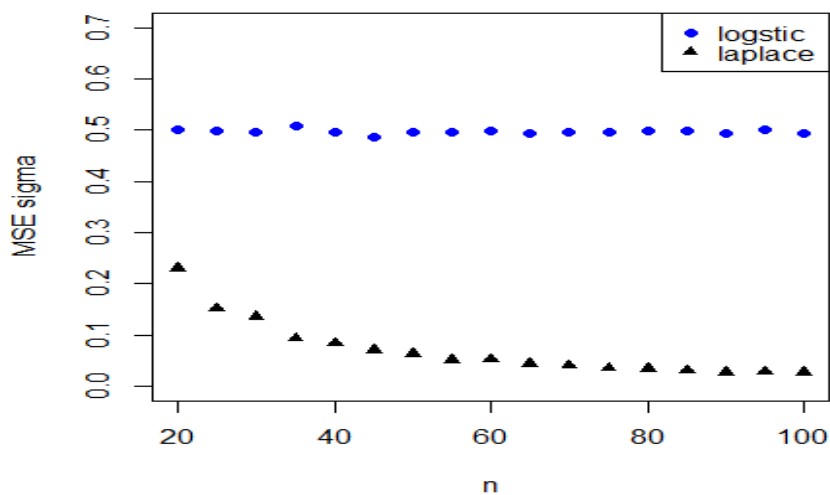
شکل ۲.۳: نمودار قدر مطلق آریبی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۲.۳ نشان می‌دهد که مقدار قدر مطلق آریبی σ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



شکل ۳.۳: نمودار MSE پارامتر μ برای دو مدل لاپلاس و لوجستیک

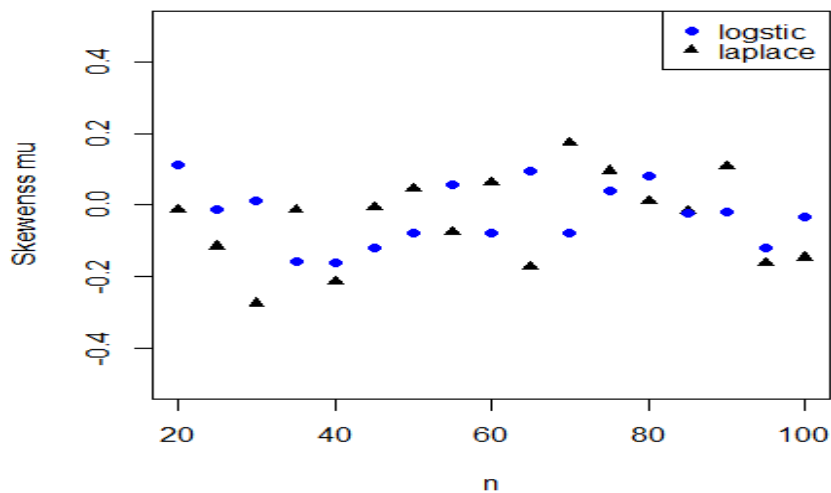
شکل ۳.۳ MSE پارامتر μ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم خطای توزیع لاپلاس کمتر است.



شکل ۴.۳: نمودار MSE پارامتر σ برای دو مدل لاپلاس و لوجستیک

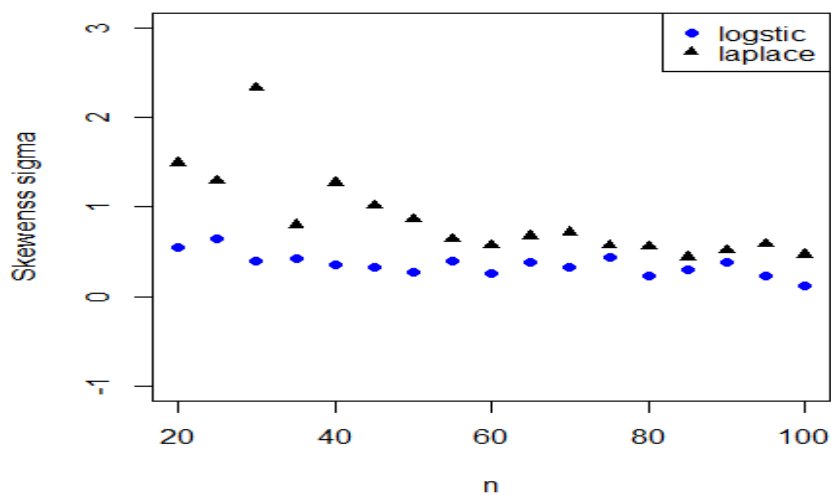
شکل ۴.۳ MSE پارامتر σ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم خطای توزیع

لاپلاس کمتر است.



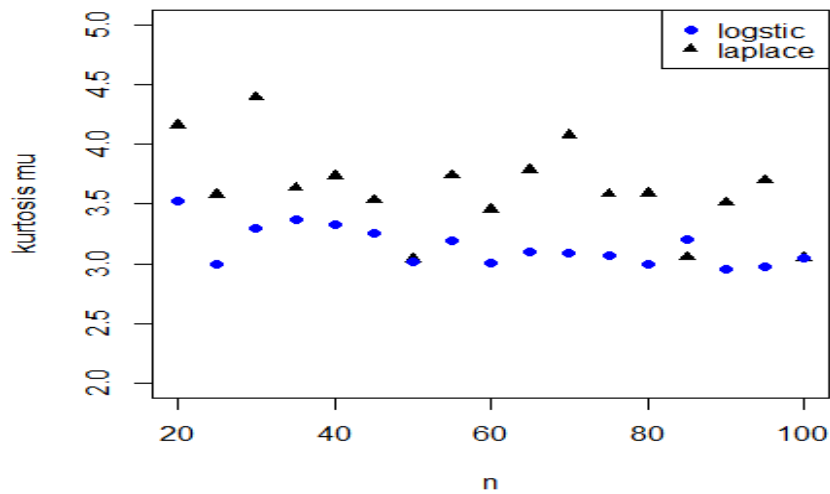
شکل ۵.۳: نمودار چولگی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۵.۳ چولگی پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی ذو توزیع مشابه می باشد.



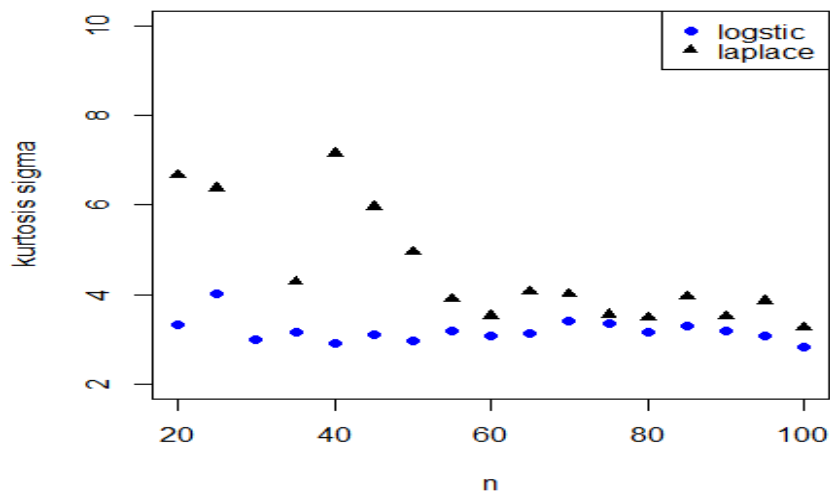
شکل ۶.۳: نمودار چولگی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۶.۳ چولگی پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی برای توزیع لوجستیک کمتر است.



شکل ۷.۳: نمودار کشیدگی پارامتر μ برای دو مدل لاپلاس و لجستیک

شکل ۷.۳ کشیدگی پارامتر μ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم کشیدگی برای توزیع لجستیک کمتر است.



شکل ۸.۳: نمودار کشیدگی پارامتر σ برای دو مدل لاپلاس و لجستیک

شکل ۸.۳ کشیدگی پارامتر σ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم کشیدگی برای توزیع لجستیک کمتر است.

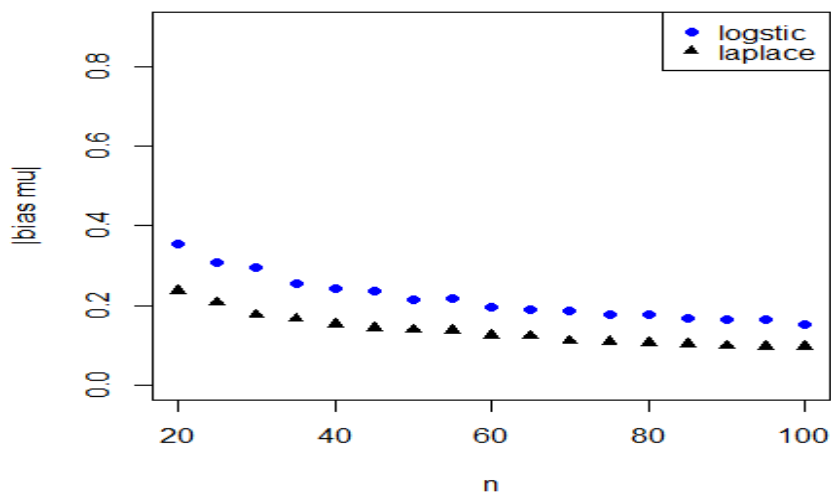
مثال ۲.۰.۳. فرض کنید که مقدار $\tau = (-\infty, -0.5, 0.25, 1, 1.75, 2.5, \infty)$ باشد. معیارهای قدرمطلق اریبی، میانگین مربعات خطا، چولگی و کشیدگی برای دو مدل لاپلاس و لوجستیک به صورت زیر می باشند.

جدول ۳.۳: نتایج شبیه سازی در مدل لاپلاس

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۳/۴۷	۳/۶۰	۰/۷۰	-۰/۱۸	۰/۱۲	۰/۰۹	۰/۲۶	۰/۲۴	۱
۷/۰۱	۴/۴۸	۱/۱۸	-۰/۲۱	۰/۱۰	۰/۰۷	۰/۲۳	۰/۲۱	۲
۳/۶۵	۳/۳۷	۰/۶۳	-۰/۲۱	۰/۰۸	۰/۰۵	۰/۲۲	۰/۱۸	۳
۳/۲۷	۳/۰۴	۰/۵۸	-۰/۱۵	۰/۰۶	۰/۰۴	۰/۱۹	۰/۱۷	۴
۳/۴۱	۳/۱۰	۰/۴۸	-۰/۱۵	۰/۰۵	۰/۰۴	۰/۱۷	۰/۱۵	۵
۳/۹۴	۳/۰۹	۰/۷۰	-۰/۲۶	۰/۰۴	۰/۰۳	۰/۱۵	۰/۱۴	۶
۳/۴۲	۲/۷۹	۰/۵۶	-۰/۰۵	۰/۰۴	۰/۰۳	۰/۱۵	۰/۱۴	۷
۳/۶۰	۳/۳۴	۰/۶۰	-۰/۰۲	۰/۰۴	۰/۰۳	۰/۱۴	۰/۱۴	۸
۴/۱۶	۳/۱۴	۰/۶۲	-۰/۱۹	۰/۰۳	۰/۰۳	۰/۱۲	۰/۱۲	۹
۳/۶۶	۳/۳۵	۰/۶۴	-۰/۳۰	۰/۰۳	۰/۰۲	۰/۱۲	۰/۱۲	۱۰
۳/۳۷	۳/۱۱	۰/۴۴	-۰/۱۷	۰/۰۳	۰/۰۲	۰/۱۲	۰/۱۲	۱۱
۳/۴۱	۳/۰۲	۰/۵۴	۰/۰۱	۰/۰۲	۰/۰۲	۰/۱۲	۰/۱۲	۱۲
۳/۵۳	۲/۹۶	۰/۴۴	-۰/۱۵	۰/۰۲	۰/۰۲	۰/۱۰	۰/۱۲	۱۳
۳/۲۱	۳/۲۷	۰/۳۴	-۰/۲۰	۰/۰۲	۰/۰۲	۰/۱۰	۰/۱۰	۱۴
۳/۵۶	۳/۱۴	۰/۵۲	-۰/۰۳	۰/۰۲	۰/۰۲	۰/۱۰	۰/۱۰	۱۵
۳/۳۱	۲/۹۱	۰/۴۱	-۰/۰۷	۲/۰۰۲	۰/۰۲	۰/۱۰	۰/۱۰	۱۶
۳/۲۵	۲/۷۳	۰/۳۲	-۰/۰۹	۰/۰۲	۰/۰۲	۰/۱۰	۰/۱۰	۱۷

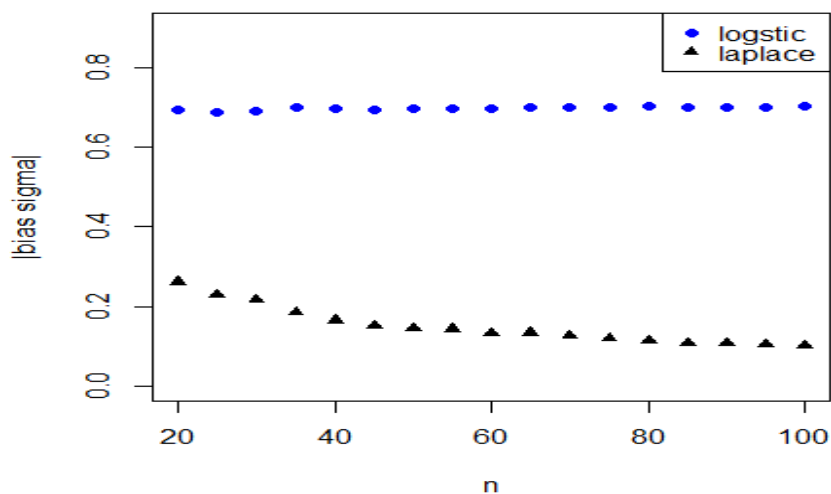
جدول ۴.۳: نتایج شبیه سازی در مدل لوجستیک

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۴/۲۰	۳/۱۶	۰/۷۰	۰/۰۶	۰/۵۰	۰/۲۰	۰/۶۹	۰/۳۶	۱
۳/۲۰	۳/۴۶	۰/۴۴	-۰/۰۱	۰/۴۹	۰/۱۶	۰/۶۹	۰/۳۱	۲
۳/۱۵	۳/۳۴	۰/۴۰	-۰/۱۵	۰/۴۹	۰/۱۴	۰/۶۹	۰/۲۹	۳
۳/۰۱	۳/۵۰	۰/۳۱	۰/۰۷	۰/۵۰	۰/۱۱	۰/۷۰	۰/۲۵	۴
۳/۲۵	۳/۳۰	۰/۴۳	۰/۱۲	۰/۴۹	۰/۰۹	۰/۷۰	۰/۲۴	۵
۳/۳۱	۳/۰۶	۰/۳۴	۰/۰۲	۰/۴۹	۰/۰۹	۰/۶۹	۰/۲۴	۶
۲/۹۱	۳/۰۴	۰/۲۳	-۰/۱۰	۰/۵۰	۰/۰۷	۰/۷۰	۰/۲۱	۷
۲/۷۲	۲/۸۰	۰/۲۳	۰/۰۳	۰/۴۹	۰/۰۷	۰/۷۰	۰/۲۲	۸
۲/۹۱	۲/۸۹	۰/۲۷	-۰/۰۴	۰/۴۹	۰/۰۶	۰/۷۰	۰/۲۰	۹
۳/۰۷	۳/۰۷	۰/۲۳	-۰/۰۵	۰/۵۰	۰/۰۶	۰/۷۰	۰/۱۹	۱۰
۲/۹۵	۳/۲۹	۰/۱۹	۰/۰۳	۰/۵۰	۰/۰۶	۰/۷۰	۰/۱۹	۱۱
۲/۹۹	۳/۰۹	۰/۱۹	-۰/۰۴	۰/۴۹	۰/۰۵	۰/۷۰	۰/۱۸	۱۲
۲/۹۷	۲/۹۴	۰/۲۰	۰/۱۲	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۳
۲/۶۳	۳/۰۵	۰/۱۶	۰/۰۱	۰/۵۰	۰/۰۴	۰/۷۰	۰/۱۷	۱۴
۳/۰۱	۳/۰۰	۰/۱۶	-۰/۰۵	۰/۴۹	۰/۰۴	۰/۷۰	۰/۱۷	۱۵
۲/۹۸	۳/۰۳	۰/۲۹	۰/۰۸	۰/۴۹	۰/۰۴	۰/۷۰	۰/۱۶	۱۶
۳/۳۸	۳/۲۵	۰/۱۸	-۰/۱۶	۰/۵۰	۰/۰۴	۰/۷۰	۰/۱۵	۱۷



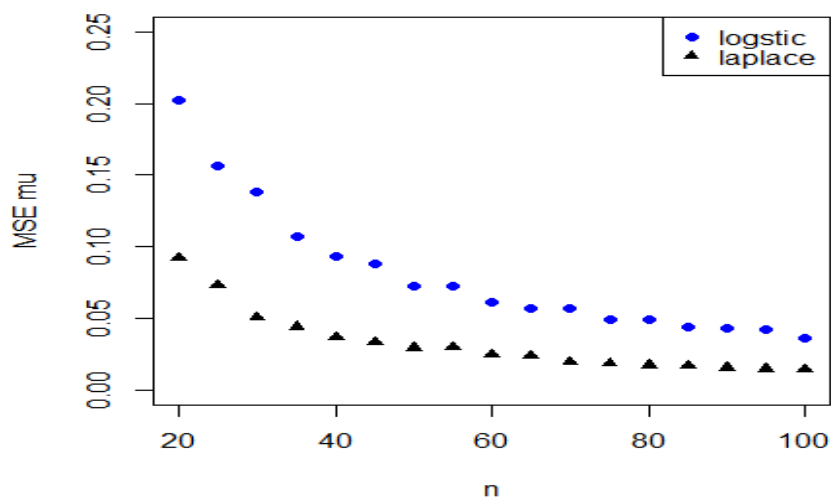
شکل ۹.۳: نمودار قدر مطلق آریبی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۹.۳ نشان می دهد که مقدار قدر مطلق آریبی میانگین μ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



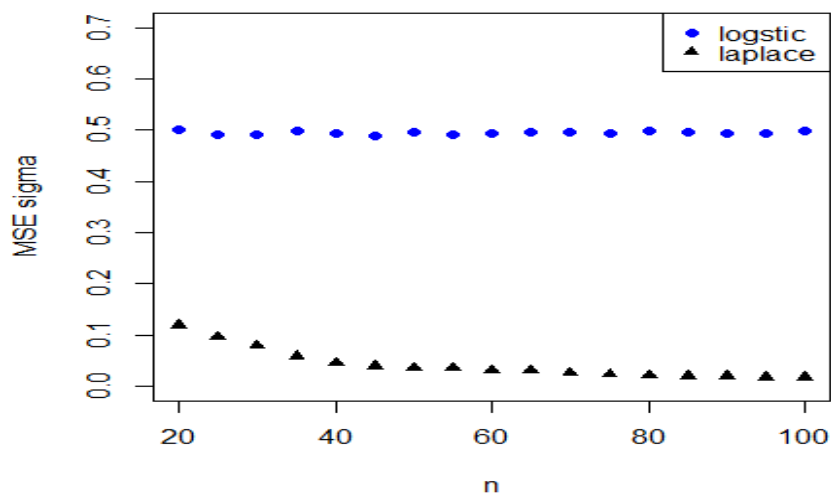
شکل ۱۰.۳: نمودار قدر مطلق آریبی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۱۰.۳ نشان می دهد که مقدار قدر مطلق آریبی σ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



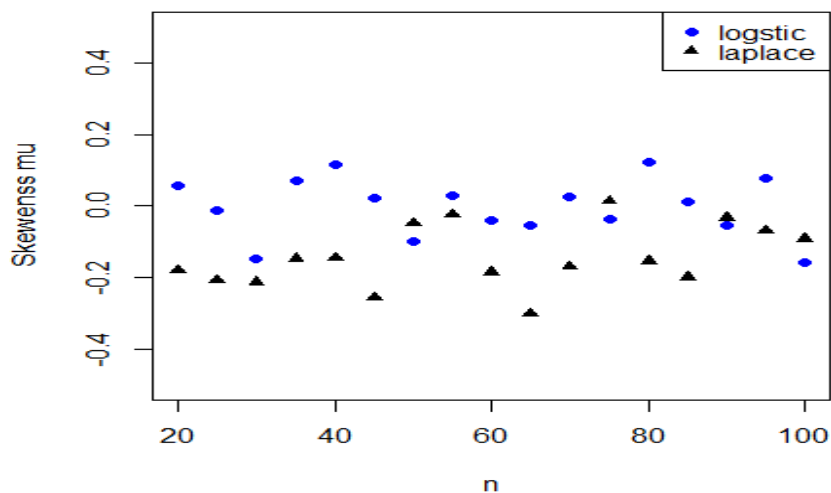
شکل ۱۱.۳: نمودار MSE پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۱۱.۳ MSE پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم خطای توزیع لاپلاس کمتر است.



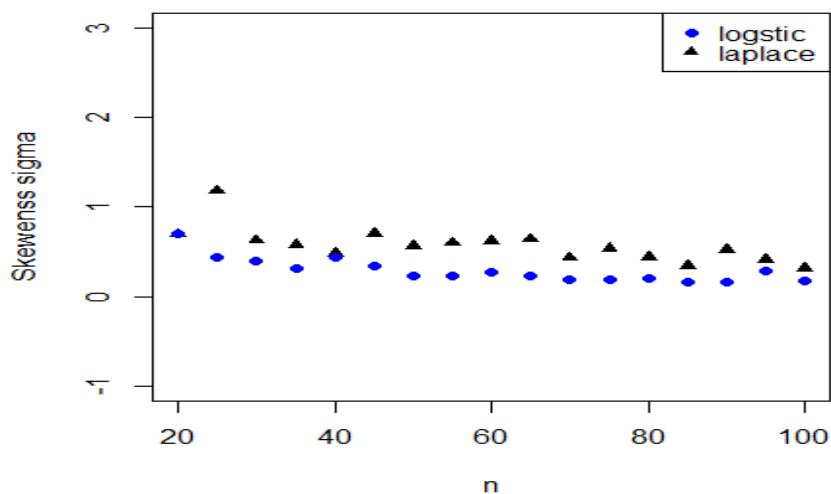
شکل ۱۲.۳: نمودار MSE پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۱۲.۳ MSE پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم خطای توزیع لاپلاس کمتر است.



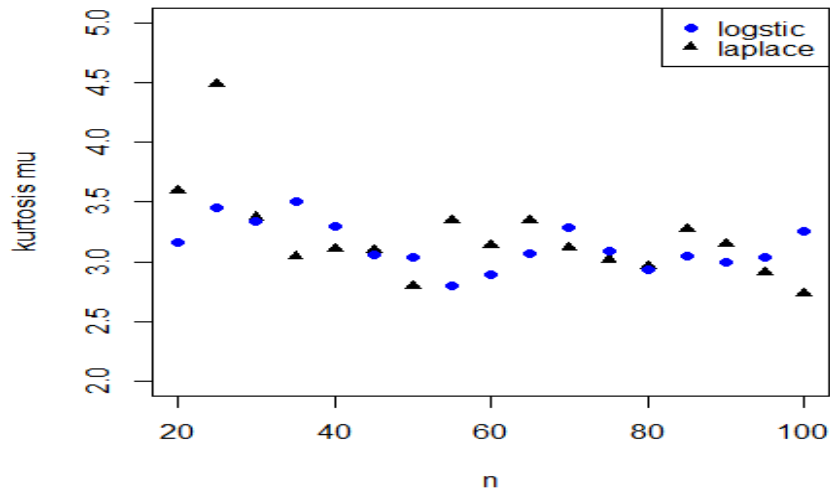
شکل ۱۳.۳: نمودار چولگی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۱۳.۳ چولگی پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی ذو توزیع مشابه می باشد.



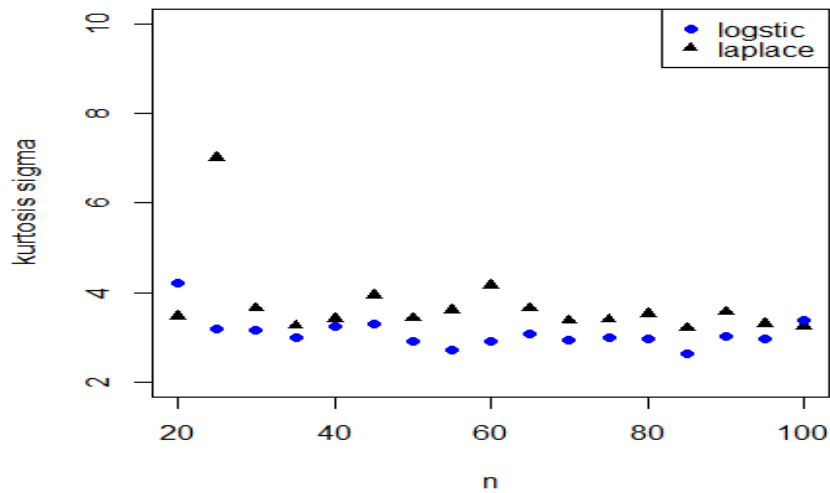
شکل ۱۴.۳: نمودار چولگی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۱۴.۳ چولگی پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی برای توزیع لوجستیک کمتر است.



شکل ۱۵.۳: نمودار کشیدگی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۱۵.۳ کشیدگی پارامتر μ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم کشیدگی برای توزیع لوجستیک کمتر است.



شکل ۱۶.۳: نمودار کشیدگی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۱۶.۳ کشیدگی پارامتر σ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم کشیدگی برای توزیع لوجستیک کمتر است.

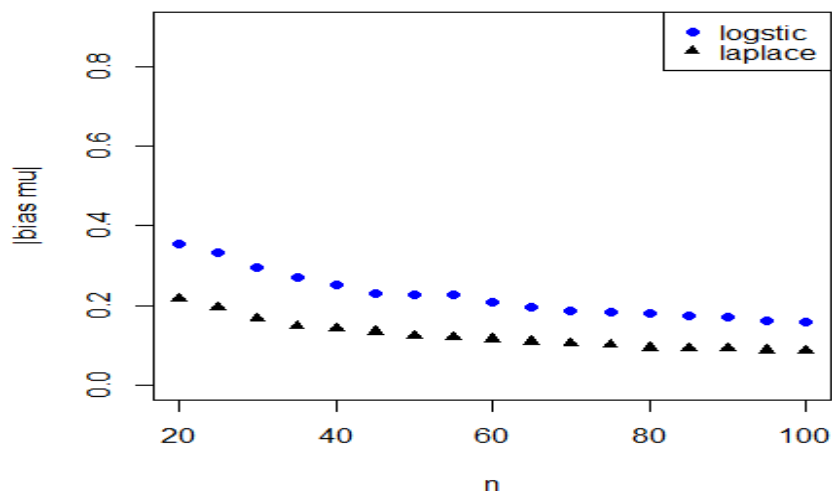
مثال ۳.۰.۳. فرض کنید که مقدار $\tau = (-\infty, -0.25, 0, 1, 2, 2.5, \infty)$ باشد. معیارهای قدرمطلق اریبی، میانگین مربعات خطا، چولگی و کشیدگی برای دو مدل لاپلاس و لوجستیک به صورت زیر می باشند.

جدول ۵.۳: نتایج شبیه سازی در مدل لاپلاس

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۶/۶۹	۳/۵۰	۱/۴۱	۰/۰۵	۰/۱۷	۰/۰۸	۰/۲۹	۰/۲۲	۱
۸/۹۱	۳/۹۲	۱/۷۳	-۰/۰۲	۰/۱۴	۰/۰۷	۰/۲۶	۰/۲۰	۲
۴/۸۷	۳/۵۲	۱/۰۰	۰/۰۵	۰/۰۹	۰/۰۵	۰/۲۲	۰/۱۷	۳
۴/۷۴	۳/۵۳	۰/۸۹	۰/۱۸	۰/۰۸	۰/۰۴	۰/۲۱	۰/۱۵	۴
۴/۶۷	۵/۷۶	۰/۸۸	۰/۵۱	۰/۰۶	۰/۰۴	۰/۱۹	۰/۱۴	۵
۴/۸۹	۳/۶۷	۰/۸۵	۰/۰۰	۰/۰۵	۰/۰۳	۰/۱۷	۰/۱۳	۶
۴/۷۲	۲/۹۶	۰/۷۸	۰/۰۴	۰/۰۴	۰/۰۲	۰/۱۶	۰/۱۲	۷
۳/۳۰	۳/۱۸	۰/۴۸	۰/۰۱	۰/۰۴	۰/۰۲	۰/۱۶	۰/۱۲	۸
۳/۶۰	۳/۳۱	۰/۵۹	-۰/۰۸	۰/۰۴	۰/۰۲	۰/۱۴	۰/۱۲	۹
۴/۸۱	۳/۹۷	۰/۷۹	-۰/۰۱	۰/۰۴	۰/۰۲	۰/۱۴	۰/۱۱	۱۰
۳/۸۸	۳/۶۴	۰/۳۹	۰/۱۸	۰/۰۳	۰/۰۲	۰/۱۴	۰/۱۰	۱۱
۳/۲۰	۳/۶۸	۰/۵۱	-۰/۰۲	۰/۰۳	۰/۰۲	۰/۱۳	۰/۱۰	۱۲
۳/۵۳	۳/۶۳	۰/۵۳	-۰/۰۵	۰/۰۳	۰/۰۱	۰/۱۳	۰/۰۹	۱۳
۳/۶۲	۳/۵۱	۰/۵۳	۰/۱۸	۰/۰۲	۰/۰۱	۰/۱۲	۰/۰۹	۱۴
۳/۷۴	۳/۰۸	۰/۷۰	-۰/۰۱	۰/۰۳	۰/۰۱	۰/۱۲	۰/۰۹	۱۵
۳/۳۳	۳/۴۳	۰/۵۳	۰/۰۳	۰/۰۲	۰/۰۱	۰/۱۲	۰/۰۹	۱۶
۳/۹۷	۳/۷۹	۰/۵۷	۰/۱۲	۰/۰۲	۰/۰۱	۰/۱۱	۰/۰۹	۱۷

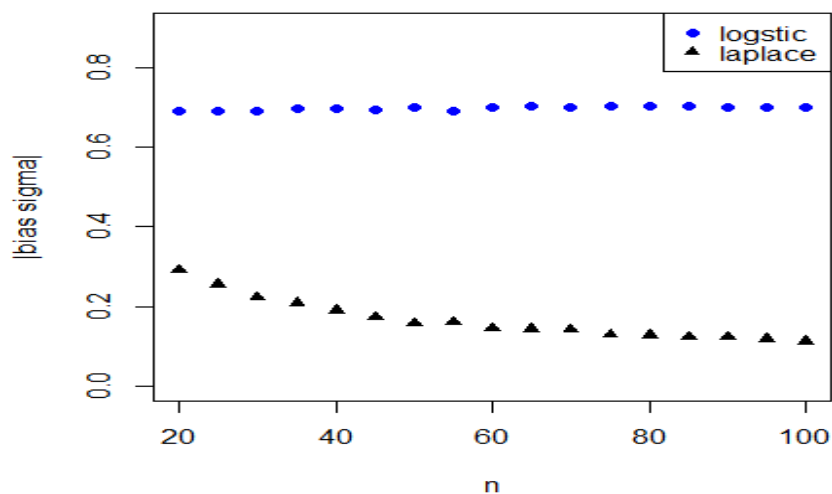
جدول ۶.۳: نتایج شبیه سازی در مدل لوجستیک

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۵/۱۴	۳/۷۹	۰/۷۲	-۰/۲۳	۰/۵۰	۰/۲۱	۰/۶۹	۰/۳۵	۱
۳/۲۴	۳/۰۱	۰/۵۴	-۰/۱۱	۰/۴۹	۰/۱۸	۰/۶۹	۰/۳۳	۲
۳/۴۹	۳/۳۶	۰/۴۳	۰/۰۷	۰/۴۹	۰/۱۴	۰/۶۹	۰/۳۰	۳
۳/۱۸	۳/۴۸	۰/۲۴	۰/۰۳	۰/۵۰	۰/۱۲	۰/۷۰	۰/۲۷	۴
۲/۶۸	۲/۸۶	۰/۱۷	-۰/۰۸	۰/۵۰	۰/۱۰	۰/۷۰	۰/۲۵	۵
۳/۶۷	۳/۱۲	۰/۴۶	-۰/۰۳	۰/۴۹	۰/۰۸	۰/۶۹	۰/۲۳	۶
۳/۳۰	۳/۱۲	۰/۳۴	۰/۱۰	۰/۵۰	۰/۰۸	۰/۷۰	۰/۲۳	۷
۳/۱۴	۳/۰۲	۰/۳۲	۰/۰۵	۰/۴۹	۰/۰۸	۰/۶۹	۰/۲۳	۸
۲/۷۷	۲/۸۲	۰/۰۶	-۰/۰۷	۰/۵۰	۰/۰۸	۰/۷۰	۰/۲۱	۹
۳/۱۰	۳/۳۶	۰/۲۶	-۰/۰۲	۰/۵۰	۰/۰۸	۰/۷۰	۰/۲۰	۱۰
۳/۰۰	۲/۹۶	۰/۲۲	-۰/۱۱	۰/۴۹	۰/۰۵	۰/۷۰	۰/۱۹	۱۱
۳/۱۰	۲/۷۸	۰/۲۶	۰/۰۷	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۲
۳/۲۱	۲/۸۶	۰/۲۹	۰/۰۰	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۳
۲/۸۹	۳/۱۱	۰/۲۵	۰/۰۸	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۷	۱۴
۲/۹۵	۳/۱۲	۰/۲۶	۰/۰۳	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۷	۱۵
۳/۰۲	۳/۰۹	۰/۲۷	-۰/۰۱	۰/۴۹	۰/۰۵	۰/۷۰	۰/۱۶	۱۶
۳/۱۲	۳/۰۲	۰/۲۴	-۰/۱۶	۰/۴۹	۰/۰۵	۰/۷۰	۰/۱۶	۱۷



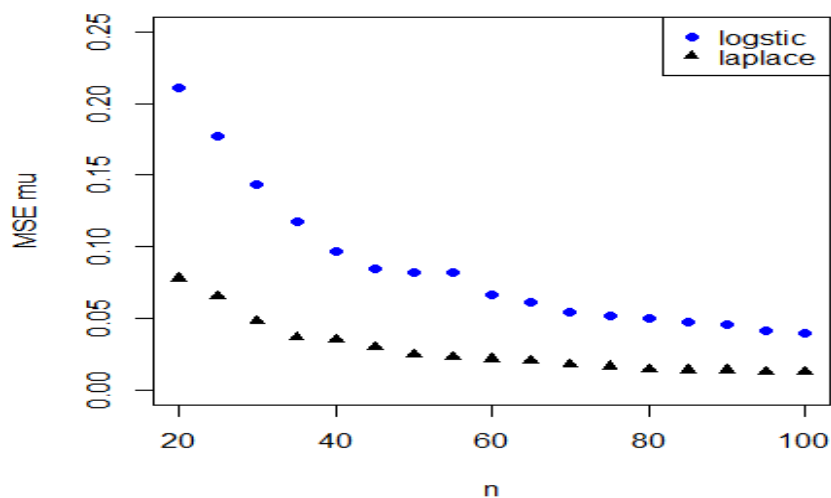
شکل ۱۷.۳: نمودار قدر مطلق آریبی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۱۷.۳ نشان می‌دهد که مقدار قدر مطلق آریبی μ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



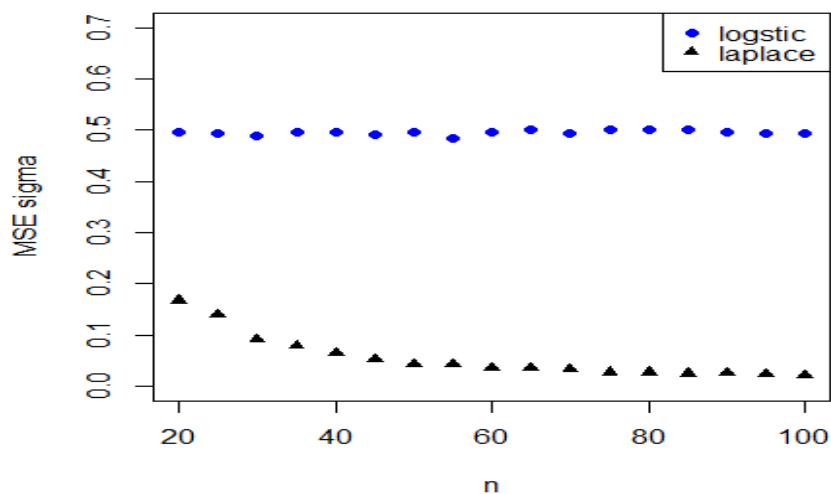
شکل ۱۸.۳: نمودار قدر مطلق آریبی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۱۸.۳ نشان می‌دهد که مقدار قدر مطلق آریبی σ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



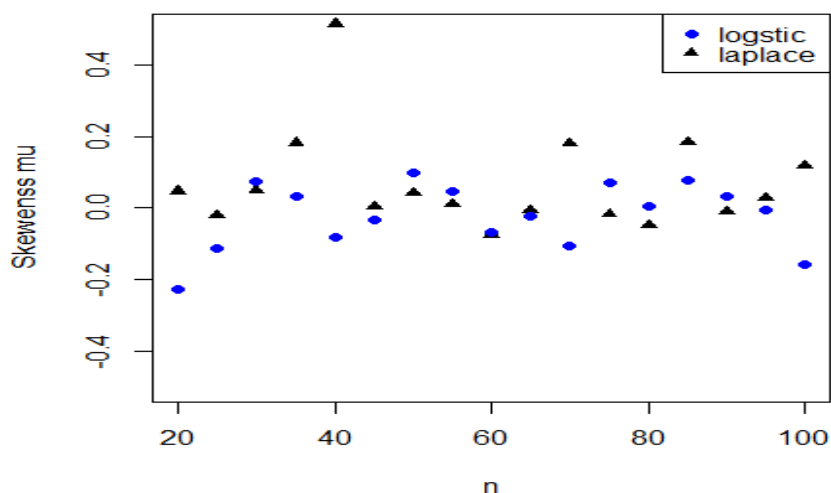
شکل ۱۹.۳: نمودار MSE پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۱۹.۳ MSE پارامتر μ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم خطای توزیع لاپلاس کمتر است.



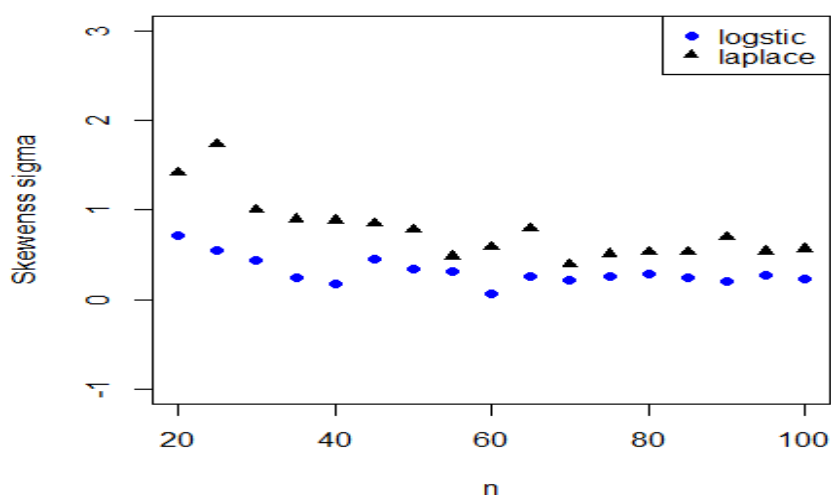
شکل ۲۰.۳: نمودار MSE پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۲۰.۳ MSE پارامتر σ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم خطای توزیع لاپلاس کمتر است.



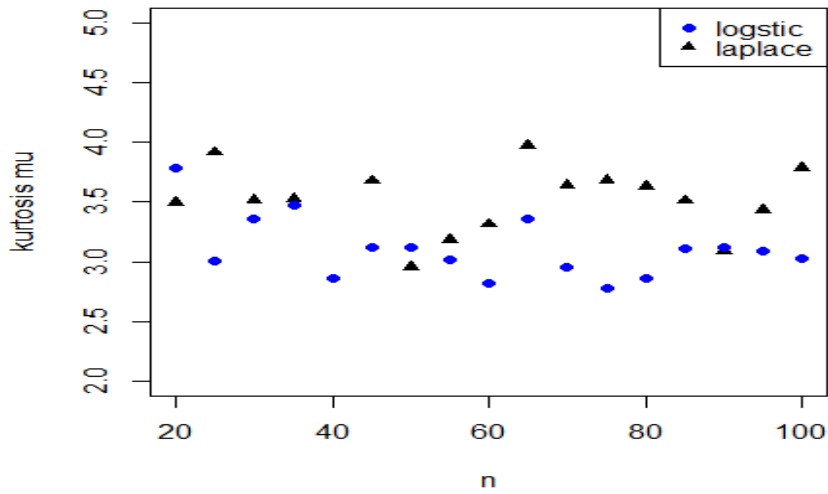
شکل ۲۱.۳: نمودار چولگی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۲۱.۳ چولگی پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی ذو توزیع مشابه می باشد.



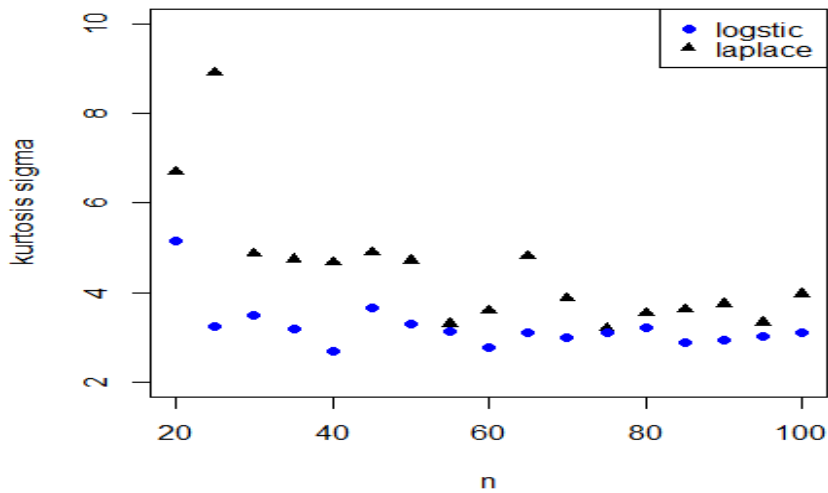
شکل ۲۲.۳: نمودار چولگی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۲۲.۳ چولگی پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی توزیع لوجستیک کمتر است.



شکل ۲۳.۳: نمودار کشیدگی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۲۳.۳ کشیدگی پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم کشیدگی برای توزیع لوجستیک کمتر است.



شکل ۲۴.۳: نمودار کشیدگی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۲۴.۳ کشیدگی پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم کشیدگی برای توزیع لوجستیک کمتر است.

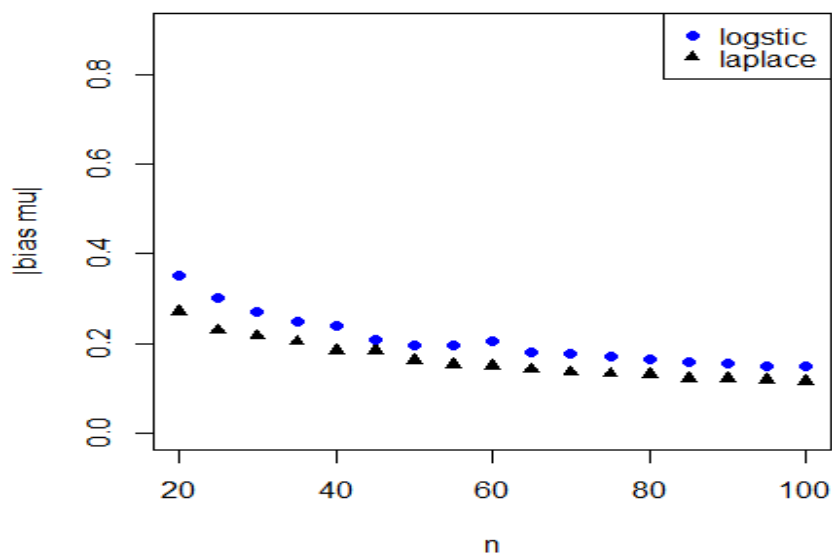
مثال ۴.۰.۳. فرض کنید که مقدار $\tau = (-\infty, -2, -1, 0.5, 1, 3, \infty)$ باشد. معیارهای قدرمطلق اریبی، میانگین مربعات خطا، چولگی و کشیدگی برای دو مدل لاپلاس و لوجستیک به صورت زیر می باشند.

جدول ۷.۳: نتایج شبیه سازی در مدل لاپلاس

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۷/۱۶	۳/۳۶	۱/۳۷	-۰/۱۷	۰/۱۱	۰/۱۲	۰/۲۴	۰/۲۷	۱
۴/۶۳	۴/۱۶	۱/۰۳	-۰/۵۱	۰/۰۷	۰/۰۹	۰/۲۰	۰/۲۳	۲
۴/۶۸	۳/۶۶	۰/۹۴	-۰/۱۳	۰/۰۶	۰/۰۸	۰/۱۸	۰/۲۲	۳
۳/۸۲	۲/۸۴	۰/۶۰	-۰/۱۳	۰/۰۴	۰/۰۷	۰/۱۶	۰/۲۰	۴
۴/۲۴	۳/۰۲	۰/۷۳	-۰/۱۷	۰/۰۴	۰/۰۵	۰/۱۵	۰/۱۸	۵
۳/۵۶	۳/۰۵	۰/۵۱	-۰/۲۳	۰/۰۳	۰/۰۵	۰/۱۵	۰/۱۸	۶
۳/۹۲	۳/۳۰	۰/۶۴	-۰/۱۲	۰/۰۳	۰/۰۴	۰/۱۳	۰/۱۶	۷
۳/۳۴	۳/۱۱	۰/۴۸	-۰/۰۶	۰/۰۳	۰/۰۴	۰/۱۳	۰/۱۵	۸
۳/۸۳	۲/۹۱	۰/۷۰	-۰/۰۶	۰/۰۲	۰/۰۳	۰/۱۲	۰/۱۵	۹
۳/۶۹	۳/۳۲	۰/۵۳	-۰/۱۴	۰/۰۲	۰/۰۳	۰/۱۲	۰/۱۴	۱۰
۴/۲۴	۳/۰۲	۰/۶۹	-۰/۱۲	۰/۰۲	۰/۰۳	۰/۱۲	۰/۱۴	۱۱
۳/۸۸	۳/۶۵	۰/۷۱	-۰/۲۴	۰/۰۲	۰/۰۳	۰/۱۲	۰/۱۳	۱۲
۳/۹۵	۲/۸۵	۰/۵۸	-۰/۱۵	۰/۰۲	۰/۰۳	۰/۱۰	۰/۱۳	۱۳
۳/۴۸	۳/۱۷	۰/۵۵	-۰/۱۶	۰/۰۲	۰/۰۲	۰/۱۰	۰/۱۲	۱۴
۳/۲۷	۳/۱۸	۰/۴۹	-۰/۱۶	۰/۰۲	۰/۰۲	۰/۱۰	۰/۱۲	۱۵
۳/۳۲	۲/۷۰	۰/۵۲	-۰/۰۷	۰/۰۱	۰/۰۲	۰/۰۹	۰/۱۲	۱۶
۳/۵۱	۲/۸۸	۰/۳۷	-۰/۰۹	۰/۰۱	۰/۰۲	۰/۰۹	۰/۱۲	۱۷

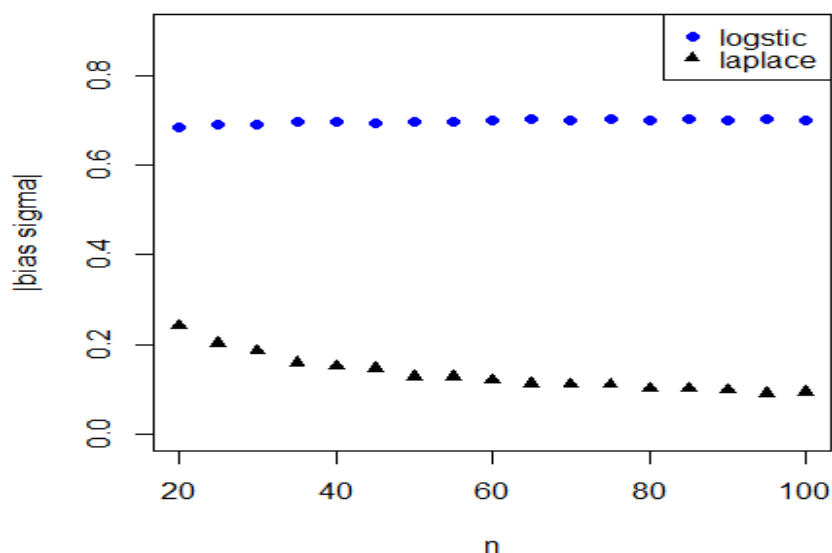
جدول ۸.۳: نتایج شبیه سازی در مدل لوجستیک

$kurtosis(\hat{\sigma})$	$kurtosis(\hat{\mu})$	$skewens(\hat{\sigma})$	$skewens(\hat{\mu})$	$mse(\hat{\sigma})$	$mse(\hat{\mu})$	$bias(\hat{\sigma})$	$bias(\hat{\mu})$	
۳/۵۴	۳/۲۲	۰/۰۱	۰/۰۱	۰/۴۸	۰/۲۰	۰/۶۹	۰/۳۵	۱
۳/۲۵	۳/۳۹	۰/۳۶	۰/۰۰	۰/۴۹	۰/۱۵	۰/۶۹	۰/۳۰	۲
۳/۶۲	۳/۴۱	۰/۴۲	-۰/۱۳	۰/۴۹	۰/۱۲	۰/۶۹	۰/۲۷	۳
۲/۹۰	۳/۴۱	۰/۱۸	-۰/۰۸	۰/۴۹	۰/۱۰	۰/۷۰	۰/۲۵	۴
۲/۸۹	۳/۲۴	۰/۲۴	-۰/۰۳	۰/۴۹	۰/۰۹	۰/۷۰	۰/۲۴	۵
۲/۹۵	۳/۲۰	۰/۲۳	۰/۰۵	۰/۴۹	۰/۰۷	۰/۶۹	۰/۲۱	۶
۳/۱۵	۲/۹۱	۰/۳۶	۰/۰۱	۰/۴۹	۰/۰۶	۰/۷۰	۰/۱۹	۷
۳/۳۹	۳/۰۴	۰/۲۷	۰/۰۳	۰/۴۹	۰/۰۶	۰/۷۰	۰/۲۰	۸
۳/۰۱	۲/۷۱	۰/۲۱	۰/۰۳	۰/۴۹	۰/۰۶	۰/۷۰	۰/۲۱	۹
۳/۱۴	۳/۰۲	۰/۲۷	-۰/۰۱	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۸	۱۰
۲/۸۹	۳/۵۶	۰/۱۹	۰/۰۷	۰/۴۹	۰/۰۵	۰/۷۰	۰/۱۸	۱۱
۲/۷۷	۳/۴۲	۰/۱۵	۰/۱۴	۰/۵۰	۰/۰۵	۰/۷۰	۰/۱۷	۱۲
۲/۹۱	۲/۹۶	۰/۲۴	۰/۰۵	۰/۴۹	۰/۰۴	۰/۷۰	۰/۱۷	۱۳
۳/۲۷	۳/۱۱	۰/۲۸	۰/۰۰	۰/۵۰	۰/۰۴	۰/۷۰	۰/۱۶	۱۴
۳/۲۷	۳/۳۹	۰/۱۷	-۰/۰۸	۰/۴۹	۰/۰۴	۰/۷۰	۰/۱۶	۱۵
۳/۱۱	۲/۹۰	۰/۲۲	-۰/۰۵	۰/۴۹	۰/۰۳	۰/۷۰	۰/۱۵	۱۶
۳/۰۰	۳/۰۵	۰/۱۸	-۰/۱۱	۰/۴۹	۰/۰۴	۰/۷۰	۰/۱۵	۱۷



شکل ۲۵.۳: نمودار قدر مطلق آریبی پارامتر μ برای دو مدل لاپلاس و لوجستیک

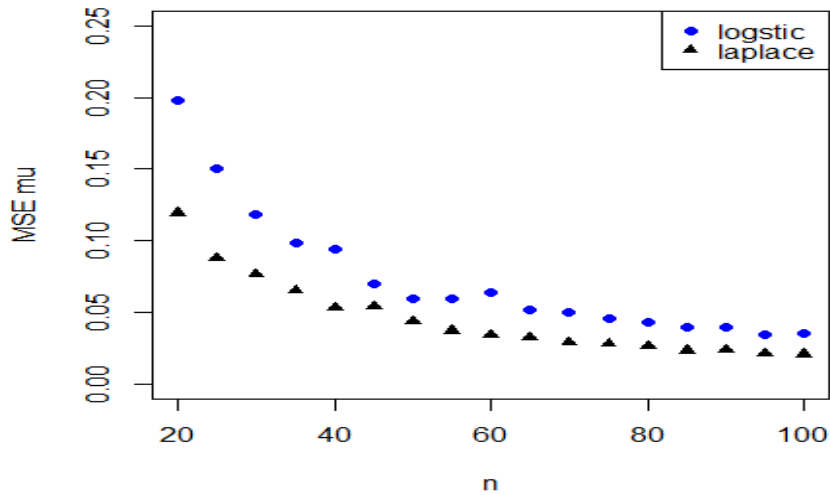
شکل ۲۵.۳ نشان می‌دهد که مقدار قدر مطلق آریبی μ در توزیع لاپلاس کمتر از توزیع لوجستیک است.



شکل ۲۶.۳: نمودار قدر مطلق آریبی پارامتر σ برای دو مدل لاپلاس و لوجستیک

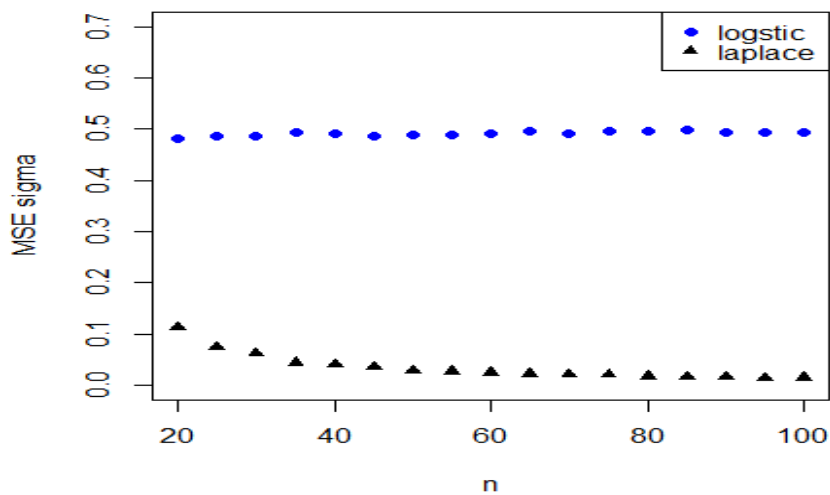
شکل ۲۶.۳ نشان می‌دهد که مقدار قدر مطلق آریبی σ در توزیع لاپلاس کمتر از توزیع

لوجستیک است.



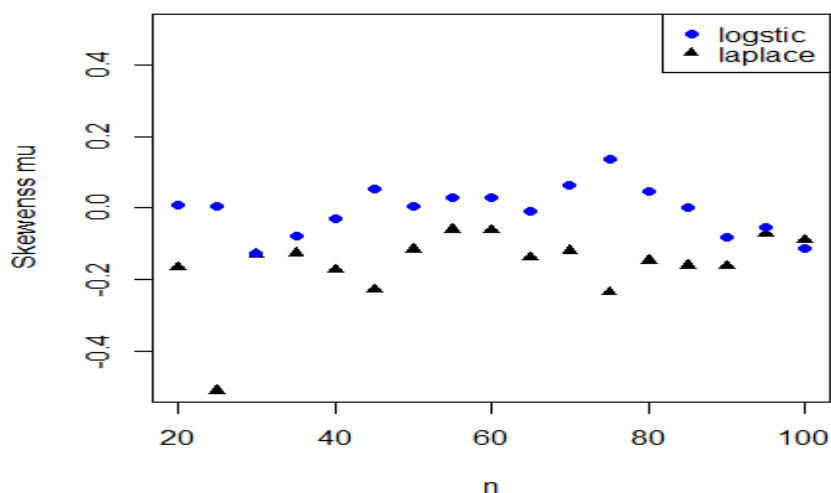
شکل ۲۷.۳: نمودار MSE پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۲۷.۳ MSE پارامتر μ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم خطای توزیع لاپلاس کمتر است.



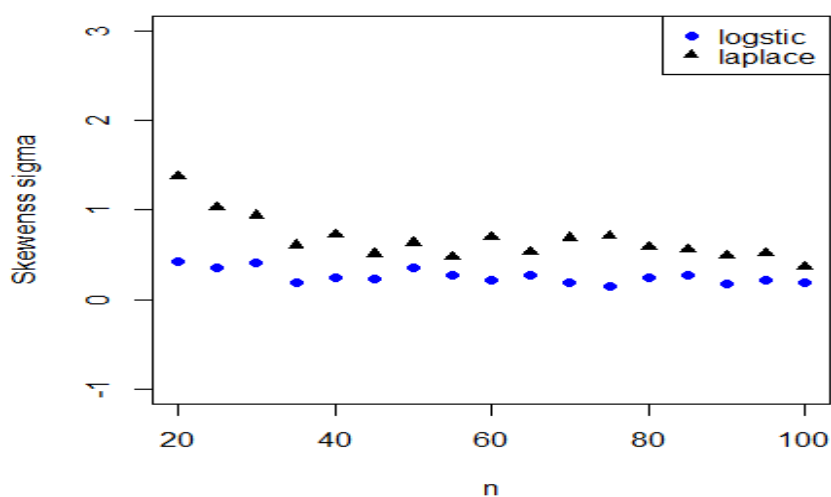
شکل ۲۸.۳: نمودار MSE پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۲۸.۳ MSE پارامتر σ را برای دو توزیع نشان می‌دهد، مشاهده می‌کنیم خطای توزیع لاپلاس کمتر است.



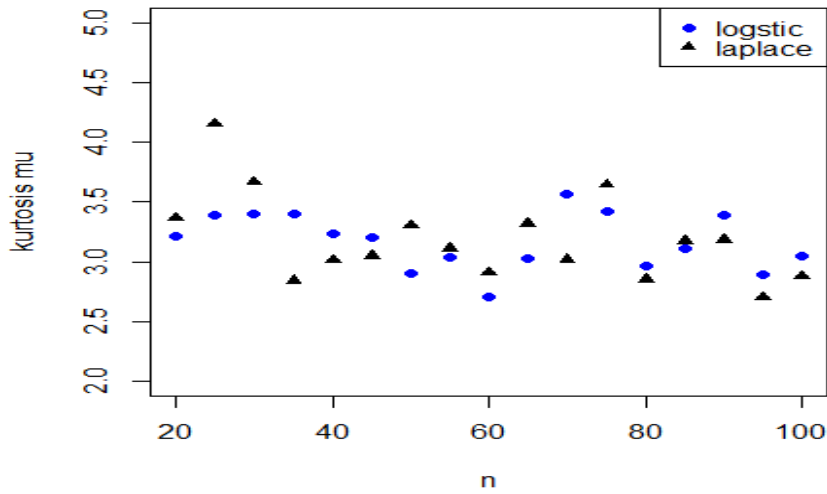
شکل ۲۹.۳: نمودار چولگی پارامتر μ برای دو مدل لاپلاس و لوجستیک

شکل ۲۹.۳ چولگی پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی توزیع لاپلاس کمتر است



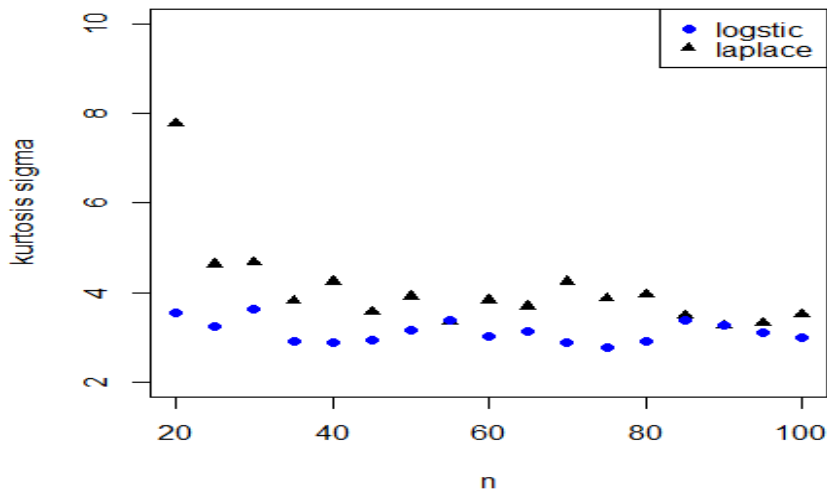
شکل ۳۰.۳: نمودار چولگی پارامتر σ برای دو مدل لاپلاس و لوجستیک

شکل ۳۰.۳ چولگی پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم چولگی توزیع لوجستیک کمتر است.



شکل ۳۱.۳: نمودار کشیدگی پارامتر μ برای دومدل لاپلاس و لوجستیک

شکل ۳۱.۳ کشیدگی پارامتر μ را برای دو توزیع نشان می دهد، مشاهده می کنیم کشیدگی برای توزیع لوجستیک کمتر است.



شکل ۳۲.۳: نمودار کشیدگی پارامتر σ برای دومدل لاپلاس و لوجستیک

شکل ۳۲.۳ کشیدگی پارامتر σ را برای دو توزیع نشان می دهد، مشاهده می کنیم کشیدگی برای توزیع لوجستیک کمتر است.

۲.۰.۳ نتیجه گیری

- اریبی و میانگین مربعات خطا $\hat{\mu}$ و $\hat{\sigma}$ در مدل لاپلاس به طور کلی کوچکتر است.
- تغییرات اریبی و میانگین مربعات خطا $\hat{\mu}$ و $\hat{\sigma}$ با توجه به مقادیر مختلف τ در مدل لاپلاس به طور کلی کوچکتر است.
- مقادیر اریبی $\hat{\mu}$ و $\hat{\sigma}$ برای هر دو مدل می توانند منفی باشند.
- قدرمطلق اریبی به طور کلی با افزایش n کاهش می یابد.
- میانگین مربعات خطا به طور کلی با افزایش n کاهش می یابد.
- مقادیر چولگی برای $\hat{\mu}$ به طور کلی نزدیک صفر است و برای $\hat{\sigma}$ داری مقادیر مثبت است.
- مقادیر کشیدگی برای هر دو برآوردگر $\hat{\mu}$ و $\hat{\sigma}$ به طور کلی ۳ است و با افزایش n کاهش می یابد.
- این مشاهدات نشان می دهد که توزیع معرفی شده به توزیع واقعی نزدیکتر است.

۱.۳ کاربرد

در این بخش، مجموعه داده‌ی کنترل لوکوموتیو که ژو و همکاران [۳۰] تهیه کردند را در نظر می گیریم. این مجموعه داده مطالعه‌ی آزمایش کنترل‌های لوکوموتیو می باشد. ژو و همکاران داده‌ها از مطالعه‌ی ”پیمودن مسافتی که در آن ۹۶ بار کنترل کننده لوکوموتیوهای مختلف شکست خوردند. و آزمایش پس از ۱۳۵۰۰۰ مایل خاتمه یافت، که در آن ۳۷ خرابی اتفاق افتاد. علاوه بر این، ۵۹ بار سانسور شد که برابر با ۱۳۵۰۰۰ مایل است.” ژو و همکاران اطلاعات را به دوازده گروه تقسیم کرد و تعداد آنها را ثبت نمودند فواصل تقسیم شده و فرکانس های مشاهده شده در جدول ۹.۳ داده شده است.

جدول ۹.۳: برآورد پارامترها

تعداد	فاصله	لگاریتم فاصله	فراوانی مشاهده شده	فراوانی مورد انتظار در مدل لاپلاس	فراوانی مورد انتظار در مدل لوجستیک
۱	< ۳۵	$(-\infty, ۳/۵۵۵]$	۱	۲/۲۵۴۵۳	۱/۱۳۷
۲	۳۵ - ۴۵	$(۳/۵۵۵, ۳/۸۰۷]$	۱	۱/۵۴۶۳۴۵	۱/۵۵۵
۳	۴۵ - ۵۵	$(۳/۸۰۷, ۴/۰۰۷]$	۵	۱/۹۵۲۲۷۳	۲/۱۳۳۷
۴	۵۵ - ۶۵	$(۴/۰۰۷, ۴/۱۷۴]$	۱	۲/۳۷۹۳۴۸	۲/۲۲۵
۵	۶۵ - ۷۵	$(۴/۱۷۴, ۴/۳۱۷]$	۳	۲/۸۰۵۵۴۵	۳/۲۶۸۴
۶	۷۵ - ۸۵	$(۴/۳۱۷, ۴/۴۴۳]$	۸	۳/۲۶۴۰۸۵	۳/۱۶۷۹
۷	۸۵ - ۹۵	$(۴/۴۴۳, ۴/۵۵۴]$	۲	۳/۶۷۳۶۷۶	۴/۱۰۶
۸	۹۵ - ۱۰۵	$(۴/۵۵۴, ۴/۶۵۴]$	۱	۴/۱۱۶۷۹۴	۴/۳۶۷۱
۹	۱۰۵ - ۱۱۵	$(۴/۶۵۴, ۴/۷۴۵]$	۴	۴/۵۶۴۸۶	۴/۵۱۴۶
۱۰	۱۱۵ - ۱۲۵	$(۴/۷۴۵, ۴/۸۲۸]$	۷	۴/۹۸۵۰۰۵	۴/۵۲۷۲
۱۱	۱۲۵ - ۱۳۵	$(۴/۸۲۸, ۴/۹۰۵]$	۴	۵/۴۵۷۷۳۴	۴/۴۹۲۶
۱۲	> ۱۳۵	$(۴/۹۰۵, \infty)$	۵۹	۵۸/۹۹۹۸	۵۸/۸۲۷

جدول ۱۰.۳: برآورد پارامترها

$f(x)$	$\hat{\mu}$	$Var(\hat{\mu})$	$\hat{\sigma}$	$Var(\hat{\sigma})$	$Cov(\hat{\mu}, \hat{\sigma})$	لگاریتم تابع درستنمایی
مدل لاپلاس	۵/۰۳۱	۰/۰۰۰۴	۰/۴۸۲	۰/۰۰۰۷	۰/۰۰۰۲	۱۵۱/۵
مدل لوجستیک	۵/۰۸۰	۰/۰۰۰۸	۰/۳۸۰	۰/۰۰۰۳	۰/۰۰۰۳	۱۵۰/۹

با توجه به جدول ۱۰.۳ مشاهده می‌کنیم مدل لاپلاس بهتر است.

فصل ۴

برنامه‌های کامپیوتری با نرم افزار R

در این پیوست تمام دستورهایی مورد نیاز برای رسم شکل‌ها و جداول با نرم‌افزار R-۳.۳.۲ آمده است.

```
rm(list=ls(all=T))
set.seed(1371)
library(rmutil)
library(moments)
#tu=c(-Inf,0,0.5,1,1.5,2,Inf)
#tu=c(-Inf,-0.5,0.25,1,1.75,2.5,Inf)
#tu=c(-Inf,-0.25,0,1,2,2.5,Inf)
tu=c(-Inf,-2,-1,0.5,1,3,Inf)
n=c(20,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100)
Output<-matrix(0,ncol=8,nrow=length(n))
for(l in 1:length(n)){
m1<-matrix(0,nr=1000,nc=2)
m2<-matrix(0,nr=1000,nc=2)
star=c(0,1.2)
for( j in 1:1000){
```

```
mu=0
sigma=1
x=rlaplace(n[1], m=0, s=1)
n1=sum(x<tu[2])
n2=sum(x>=tu[2]&x<tu[3])
n3=sum(x>=tu[3]&x<tu[4])
n4=sum(x>=tu[4]&x<tu[5])
n5=sum(x>=tu[5]&x<tu[6])
n6=sum(x>=tu[6])
t0=tu[1]
t1=tu[2]
t2=tu[3]
t3=tu[4]
t4=tu[5]
t5=tu[6]
t6=tu[7]
ni=c(n1,n2,n3,n4,n5,n6)
ti=c(t0,t1,t2,t3,t4,t5,t6)
k=length(ti)
loglike=function(par,...){
  ni<- as.vector(ni)
  ti<- as.vector(ti)
  theta1=par[1];theta2=par[2]
  i=2
  sum=0
  while(i<=k){
    if((theta2*ti[i]-theta1)>=0){
      F1=1-1/2*(exp(-(theta2*ti[i]-theta1)))}else{
      F1=1/2*exp(theta2*ti[i]-theta1)
    }
    if((theta2*ti[i-1]-theta1)>=0){
      F2=1-1/2*(exp(-(theta2*ti[i-1]-theta1)))}else{
```

```

        F2=1/2*exp(theta2*ti[i-1]-theta1)
    }
    sum=sum+ni[i-1]*log(F1-F2)
    i=i+1
}
#return(sum)
loglike=-sum
}

    m1[j,]<-optim(star,loglike,method="Nelder-Mead")$par
}
rb<-c(0,1)
rb=as.matrix(rb)
rb=t(rb)
Output[1,1]<-sum(abs(m1[,1]-rb[,1]))/1000
Output[1,2]<-sum(abs(m1[,2]-rb[,2]))/1000
Output[1,3]<-(sum((m1[,1]-rb[,1])^2)/1000)
Output[1,4]<-(sum((m1[,2]-rb[,2])^2)/1000)
Output[1,5]=skewness((m1[,1]))
Output[1,6]=skewness((m1[,2]))
Output[1,7]=kurtosis((m1[,1]))
Output[1,8]=kurtosis((m1[,2]))
colnames(Output)=c("biasmu","biassigma","msemu","msesigma","skewens mu"
,"skewens sigma","kurtosis mu","kurtosis sigma")
}
#rm(list=ls(all=T))
library(rmutil)
library(moments)
#tu=c(-Inf,0,0.5,1,1.5,2,Inf)
#tu=c(-Inf,-0.5,0.25,1,1.75,2.5,Inf)
#tu=c(-Inf,-0.25,0,1,2,2.5,Inf)
tu=c(-Inf,-2,-1,0.5,1,3,Inf)
n=c(20,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100)
Output1<-matrix(0,ncol=8,nrow=length(n))

```

```

for(l in 1:length(n)){
  m11<-matrix(0,nr=1000,nc=2)
  m21<-matrix(0,nr=1000,nc=2)
  star=c(0,1.2)
  for( j in 1:1000){

    mu1=0
    sigma1=6/sqrt(pi)
    x=rlogis(n[1], location = mu1, scale = sigma1)
    n1=sum(x<tu[2])
    n2=sum(x>=tu[2]&x<tu[3])
    n3=sum(x>=tu[3]&x<tu[4])
    n4=sum(x>=tu[4]&x<tu[5])
    n5=sum(x>=tu[5]&x<tu[6])
    n6=sum(x>=tu[6])
    t0=tu[1]
    t1=tu[2]
    t2=tu[3]
    t3=tu[4]
    t4=tu[5]
    t5=tu[6]
    t6=tu[7]
    ni=c(n1,n2,n3,n4,n5,n6)
    ti=c(t0,t1,t2,t3,t4,t5,t6)
    k=length(ti)
    loglike1=function(par,...){
      ni<- as.vector(ni)
      ti<- as.vector(ti)
      theta1=par[1];theta2=par[2]
      i=2
      sum=0
      while(i<=k){

```

```

F11=1/(1+(exp(-(theta2*ti[i]-theta1))))

F21=1/(1+(exp(-(theta2*ti[i-1]-theta1))))

sum=sum+ni[i-1]*log(F11-F21)
i=i+1
}
#return(sum)
loglike1=-sum
}

m11[j,]<-optim(star,loglike1,method="Nelder-Mead")$par
}
rb1<-c(0,1)
rb1=as.matrix(rb1)
rb1=t(rb1)
Output1[1,1]<-sum(abs(m11[,1]-rb1[,1]))/1000
Output1[1,2]<-sum(abs(m11[,2]-rb1[,2]))/1000
Output1[1,3]<-(sum((m11[,1]-rb1[,1])^2)/1000)
Output1[1,4]<-(sum((m11[,2]-rb1[,2])^2)/1000)
Output1[1,5]=skewness((m11[,1]))
Output1[1,6]=skewness((m11[,2]))
Output1[1,7]=kurtosis((m11[,1]))
Output1[1,8]=kurtosis((m11[,2]))
colnames(Output1)=c("biasmu","biassigma","msemu","msesigma","skewens
mu","skewens sigma","kurtosis mu","kurtosis sigma")
}

plot(n,Output[,1],ylab="|bias mu|",pch=17,ylim=c(0,0.9))
lines(n,Output1[,1],type="p",pch=16,col="blue")
legend("topright",
      c("logstic",

```

```

        "laplace"
    ),
    pch=c(16,17),
    col=c("blue", "black")
)
plot(n,Output[,2],ylab="|bias sigma|",pch=17,ylim=c(0,0.9))
lines(n,Output1[,2],type="p",pch=16,col="blue")
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
)
plot(n,Output[,3],pch=17,ylab="MSE mu",ylim=c(0,0.25))
lines(n,Output1[,3],type="p",pch=16,col="blue")
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
)
plot(n,Output[,4],pch=17,ylab="MSE sigma",ylim = c(0,0.7))
lines(n,Output1[,4],type="p",pch=16,col="blue")
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
)
plot(n,Output[,5],pch=17,ylab="Skewenss mu",ylim=c(-0.5,0.5))
lines(n,Output1[,5],type="p",pch=16,col="blue")

```

```
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
    )
plot(n,Output[,6],pch=17,ylab="Skewenss sigma",ylim=c(-1,3))
lines(n,Output1[,6],type="p",pch=16,col="blue")
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
    )
plot(n,Output[,7],pch=17,ylab="kurtosis mu",ylim=c(2,5))
lines(n,Output1[,7],type="p",pch=16,col="blue")
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
    )
plot(n,Output[,8],pch=17,ylab="kurtosis sigma",ylim=c(2,10))
lines(n,Output1[,8],type="p",pch=16,col="blue")
legend( "topright",
        c( "logstic",
            "laplace"
        ),
        pch=c(16,17),
        col=c("blue", "black")
    )
)
```

Output

Output1

مراجع

- [1] Chen, D. G. and Lio, Y. L. (2010). "Parameter estimations for generalized exponential distribution under progressive type-I interval censoring". **Computational Statistics and Data Analysis**, 54, 1581-1591.
- [2] Tomazella. V. L. D. and Nadarajah. S. (2010). "Estimation of parameters in Laplace distributions with interval censored data". **Computational Statistics and Data Analysis**
- [3] Cleveland, W. S. (1979). "Robust locally weighted regression and smoothing scatter plots". **Journal of the American Statistical Association**, 74, 829-836.
- [4] Cleveland, W. S. (1981). "LOWESS: A program for smoothing scatterplots by robu locally weighted regression". **The American Statisticians**, 35, 54-67.
- [5] Corradini, M. G., Normand, M. D., Nussinovitch, A., Horowitz, J. and Peleg, M. (2001). "Estimating the frequency of high microbial counts in commercial food products using various distribution functions". **Journal of Food Protection**, 64, 674-681.
- [6] Ding, C., Yang, C. and Tse, S. -K. (2010). "Accelerated life test sampling plans for the Weibull distribution under type I progressive interval censoring with random removals". **Journal of Statistical Computation and Simulation**, 80, 903-914.
- [7] Fay, M. P. and Shaw, P. A. (2010). "Exact and asymptotic weighted logrank tests for interval censored data: The interval R package". **Journal of Statistical Software**, 36, 1-34.
- [8] Ferguson, T. S. (1996). "A Course in Large Sample Theory". **Chapman and Hall, London**.
- [9] Hashimoto, E. M., Ortega, E. M. M., Cancho, V. G. and Cordeiro, G. M. (2010). "The log-exponential Weibull regression model for interval censored data". **Computational Statistics and Data Analysis**, 54, 1017-1035.
- [10] Hurlimann, W. (2001). "Financial data analysis with two symmetric distributions". **Astin Bulletin**, 31, 187-211.

-
- [11] Kozubowski, T. J. and Nadarajah, S. (2010). "Multitude of Laplace distributions". **Statistical Papers**, 51, 127-148.
- [12] Lawless, F. J. (2003). "Statistical Models and Methods for Lifetime Data". **John Wiley and Sons, New York**.
- [13] Lin, C. -T., Wu, S. J. S. and Balakrishnan, N. (2009). "Planning life tests with progressively type-I interval censored data from the lognormal distribution". **Journal of Statistical Planning and Inference**, 139, 54-61.
- [14] Lindsey, J. K. (1998). "A study of interval censoring in parametric regression models". **Life-time Data Analysis**, 4, 329-354.
- [15] Lindsey, J. K., Byrom, W. D., Wang, J., Jarvis, P. and Jones, B. (2000). "Generalized non-linear models for pharmacokinetic data". **Biometrics**, 56, 81-88.
- [16] Lio, Y. L., Chen, D. -G. and Tsai, T. -R. (2011). "Parameter estimations for generalized Rayleigh distribution under progressively type-I interval censored data". **Open Journal of Statistics**, 1, 46-57.
- [17] Lu, W. and Tsai, T. -R. (2009). "Interval censored sampling plans for the gamma lifetime model". **European Journal of Operational Research**, 192, 116-124.
- [18] Ng, H. K. T. and Wang, Z. (2009). "Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample". **Journal of Statistical Computation and Simulation**, 79, 145-159.
- [19] Patel, M. N. and Gajjar, K. A. (2010). "Inference under progressive interval censoring for geometric distribution". **International Journal of Statistics and Economics**, 5, 21- 36.
- [20] Peleg, M. (2002). "Interpretation of the irregularly fluctuating microbial counts in commercial dairy products". **International Dairy Journal**, 12, 255-262.
- [21] R Development Core Team (2014). "R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing". **Vienna, Austria**.
- [22] Scallan, A. J. (1999). "Regression modelling of interval censored failure time data using the Weibull distribution". **Journal of Applied Statistics**, 26, 613-618.
- [23] Sparling, Y. H., Younes, N., Lachin, J. M. and Bautista, O. M. (2006). "Parametric survival models for interval censored data with time-dependent covariates". **Biostatistics**, 7, 599-614.

- [24] Sun, J. (2006). "The Statistical Analysis of Interval Censored Failure Time Data". **Springer Verlag, New York**.
- [25] Teimouri, M. and Gupta, A. K. (2012). "Estimation methods for the GompertzMakeham distribution under progressively type-I interval censoring scheme". **National Academy Science Letters—India**, 35, 227-235.
- [26] Tse, S. -K., Ding, C. and Yang, C. (2008). "Optimal accelerated life tests under interval censoring with random removals: The case of Weibull failure distribution". **Statistics**, 42, 435-451.
- [27] Tse, S. -K., Yuen, H. -K. and Yang, C. (2002). "Statistical analysis of exponential lifetimes under an integrated type-II interval censoring scheme". **Journal of Statistical Computation and Simulation**, 72, 461-471.
- [28] Tomazella.V.L.D. and Nadarajah.S.(2010),"Estimation of Parametres in Laplace distributions with interval censored data". **Computational Statistics and Data Analysis**
- [29] Uppuluri, V. R. R. (1981). "Some properties of the log-Laplace distribution". **In: Statistical Distributions in Scientific Work**, volume 4, pp. 105-110, NATO Advanced Study Institute Series C: Mathematical and Physical Sciences, 79, Reidel, Dordrecht.
- [30] Zhou, Y. Y., Mi, J. and Guo, S. (2007). "Estimation of parameters in logistic and log-logistic distribution with grouped data". **Lifetime Data Analysis**, 13, 421-429.

واژه‌نامه فارسی به انگلیسی

Probabilistic	احتمالی
Valuation	ارزیابی
Bias	اریبی
Trace	اثر
Regression to the mean	بازگشت به میانگین
Restricted estimator	برآوردگر محدود شده
Function Space	فضای تابع
Optimality	بهینگی
parameter	پارامتر
Rank	رتبه
Multiple linear regression	رگرسیون چند متغیره خطی
Response variable	متغیر پاسخ
Regressor variable	متغیر رگرسیونی
Symetric	متقارن
Singular linear model	مدل خطی منفرد
Inverse	معکوس
Generalized inverse	معکوس تعمیم یافته
Positive definite	معین مثبت
Constrast	مقابله
Linear dependence	وابسته خطی
Correlation	همبستگی
Multicollinearity	همخطی چندگانه
Asymptotic Distributional	توزیع مجانبی
Asymptotic Normal	نرمال مجانبی
Determinant	دترمینان
Eigen Value	مقدار ویژه

Full Rank	رتبه کامل
Generalized Least Squares	کمترین توان های دوم تعمیم یافته
Idempotent	خودتوان
Ill Conditioned	بدشرطی
Independent Variable	متغیر مستقل
Indicator Function	تابع نشانگر
Lagrangian	لاگرانژ
Linear Non-Stochastic Constrains	محدودیت های خطی غیر تصادفی
Regularity Conditions	شرایط نظم
Residual	مانده
Unbiased	نااریب
Variance Inflation Factor	عامل تورم واریانس
Weighted Least Square Estimator	برآوردگر کمترین توان های دوم وزنی

واژه‌نامه انگلیسی به فارسی

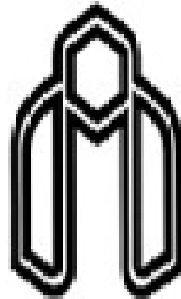
Probabilistic	احتمالی
Valuation	ارزیابی
Bias	اریبی
Trace	اثر
Regression to the mean	بازگشت به میانگین
Estimator	برآوردگر
Function Space	فضای تابع
Optimality	بهینگی
Parameter	پارامتر
Rank	رتبه
Multiple linear regression	رگرسیون چند متغیره خطی
Response variable	متغیر پاسخ
Regressor variable	متغیر رگرسیونی
Symetric	متقارن
Singular linear model	مدل خطی منفرد
Inverse	معکوس
Generalized inverse	معکوس تعمیم یافته
Positive definite	معین مثبت
Constrast	مقابله
Linear dependence	وابسته خطی
Hat matrix	هت ماتریس - ماتریس برازش
Correlation	همبستگی
Multicollinearity	همخطی چندگانه
Asymptotic Distributional	توزیع مجانبی
Asymptotic Normal	نرمال مجانبی
Coefficient of Determination	ضریب تعیین

Determinant	دترمینان
Eigen Value	مقدار ویژه
Exact Collinearity	همخطی کامل
Full Rank	رتبه کامل
Generalized Least Squares	کمترین توان های دوم تعمیم یافته
Idempotent	خودتوان
Ill Conditioned	بدشرطی
Independent Variable	متغیر مستقل
Indicator Function	تابع نشانگر
Lagrangian	لاگرانژ
Regularity Conditions	شرایط نظم
Residual	مانده
Regression	رگرسیون
Unbiased	نااریب
Weighted Least Square Estimator	برآوردگر کمترین توان های دوم وزنی

Abstract

We derive maximum likelihood estimators for the parameters of the Laplace distribution for interval censored data. Existence and uniqueness of the estimators are proved. Simulations and real data applications show that the Laplace distribution can be a better model for interval censored data than competing models in spite of being simpler.

Keywords: Laplace distribution, Logistic distribution, Maximum likelihood estimators, Censored, Interval censored, Reliability, Survival function.



Shahrood University of Technology

Faculty Of Mathematical Sciences

M.Sc. Thesis in Mathematical Statistics

**Estimation of parameters in Laplace
distributions with interval censored data**

By: Zahra Nadri

Supervisor

Ahmad Nezakati Rezazadeh

December 2017