دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده ریاضی و علوم کامپیوتر

پایان نامه کارشناسی ارشد

آمار ریاضی

موضوع:

توزیع طول عمر جدید

(آماره‌های ترتیبی توزیع نمایی — پواسن)

نگارش :
فاطمه حسن تبار

اساتید راهنما:
دکتر صادق رضایی
دکتر عادل محمدپور

استاد مشاور:
دکتر اسماعیل خرم

آبان ۱۳۸۷

بسمه تعالی

تاریخ:۱۳۸۷/۱۰/۱۱
شماره:

معاونت پژوهشی
فرم پروژه تحصیلات تکمیلی ۷

فرم اطلاعات پایان‌نامه
کارشناسی- ارشد و دکترا

دانشگاه صنعتی امیرکبیر
(پلی‌تکنیک تهران)

مشخصات دانشجو:

نام و نام خانوادگی: فاطمه حسن تبار درزی   **دانشجوی آزاد** ●   **بورسیه** ○   **معادل** ○

شماره دانشجوئی: ۸۵۱۱۳۰۰۱   **دانشکده:** علوم ریاضی و کامپیوتر   **رشته تحصیلی:** آمار ریاضی   **گروه:** آمار

مشخصات استاد راهنما:

نام و نام خانوادگی: دکتر صادق رضایی   **درجه و رتبه:** دانشیار

نام و نام خانوادگی: دکتر عادل محمدپور   **درجه و رتبه:** استادیار

مشخصات استاد مشاور:

نام و نام خانوادگی: دکتر اسماعیل خرم   **درجه و رتبه:** دانشیار

نام و نام خانوادگی: **درجه و رتبه:**

عنوان پایان‌نامه به فارسی: توزیع طول عمر جدید (آماره های ترتیبی توزیع نمایی-پواسن)

عنوان پایان‌نامه به انگلیسی:

A New lifetime distribution (order statistics for exponential Poisson distribution)

نوع پروژه: کارشناسی ○   ارشد ●   دکترا ○   سال تحصیلی: ۱۳۸۷-۱۳۸۵

کاربردی ○   بنیادی ○   توسعه‌ای ○   نظری ●

تاریخ شروع: ۸۶/۱۱/۲۰   تاریخ خاتمه: ۸۷/۸/۱۵   تعداد واحد: ۶   سازمان تأمین کننده اعتبار:

واژه‌های کلیدی به فارسی: توزیع نمایی، نرخ شکست نزولی(DFR)، نرخ شکست صعودی (IFR) ، توزیع پواسن بریده شده، الگوریتم EM ، شبیه سازی ، برآورد درستنمایی ماکسیمم.

واژه‌های کلیدی به انگلیسی :

exponential distribution, decreasing failure rate(DFR), increasing failure rate(IFR), truncate Poisson distribution, EM algorithm, simulation, and maximum likelihood estimation.

مشخصات ظاهری	تعداد صفحات			تعداد مراجع	تعداد صفحات ضمائم
	۸۳	تصویر ○ جدول ● نمودار ● نقشه ○ واژه‌نامه ●		۳۰	۱۳
زبان متن	فارسی ●	انگلیسی ○	چکیده	فارسی ●	انگلیسی ●
یادداشت					

نظرها و پیشنهادها به منظور بهبود فعالیت‌های پژوهشی دانشگاه

استاد:

دانشجو:

امضاء استاد راهنما:

تاریخ:

سپاسگزاری

سپاس مهربانی که، نیازمندی‌هایت را در اختیارت قرار می‌دهد. تشکر و سپاس نسبت به تمام عزیزانی که مرا در انجام این پروژه یاری رساندند. بخصوص اساتید ارجمندم: جناب آقای دکتر صادق رضائی که با راهنمائی‌های ارزنده، بنده را در به ثمر رسیدن این پروژه یاری دادند و جناب آقای دکتر عادل محمدپور که دانش و تخصص خود را وقف آموزش و ارتقاء علم نموده‌اند نهایت سپاس و احترام را دارم. از جناب آقای دکتر اسماعیل خرم نیز که مشاوره پایان‌نامه را به عهده داشتند، کمال تشکر و سپاس را ابراز می‌کنم.

از آقای رسول طهماسبی که در راهنمایی و گره‌گشایی این پایان‌نامه از هیچ تلاشی دریغ نکردند تشکر و سپاس دارم.

در نهایت از دوست گرامی‌ام جناب آقای مهرداد اسلامی که با بردباری‌هایشان بنده را در به ثمر رسیدن این پایان نامه یاری رساندن کمال تشکر را دارم و امیدوارم در ادامه پاسخگوی مهربانی‌هایشان باشم.

تقدیم به‌پدر و مادر عزیزم

آنانکه

شاهدانی شایسته و کوشا بر پویایی‌ام بودند.
شکوه انسان را در علم و تقوا می‌دانستند.
و حضورشان الهامی پیوسته و جاری بر دانش‌اندوزیم بود.

چکیده

در این پایان‌نامه، توزیع OEP معرفی می‌شود. این توزیع، توزیع kامین آماره ترتیبی است که از ترکیب توزیع نمایی با پواسن بریده شده بدست آمده است. نرخ شکست این توزیع به ازاء مقادیر مختلف پارامترها نزولی (DFR)، صعودی (IFR) و صعودی—نزولی $(IDFR)$ است. ویژگی‌های توزیع پیشنهاد شده مورد بحث و بررسی قرار گرفته و پارامترهای آن با استفاده از الگوریتم EM برآورد شده است و سپس واریانس و کواریانس مجانبی این پارامترها بدست آورده شده است. به منظور بررسی دقت تقریب واریانس و کواریانس برآوردگرهای ماکسیمم درستنمایی، شبیه‌سازی انجام شده و نتایج تجربی با داده‌های واقعی بیان شده است.

لغات کلیدی: توزیع نمایی، نرخ شکست نزولی (DFR)، نرخ شکست صعودی (IFR)، توزیع پواسن بریده شده، الگوریتم EM، شبیه‌سازی، برآورد ماکسیمم درستنمایی.

فهرست مندرجات

فصل ۱

مفاهیم اساسی و تعاریف پایه‌ای

۱.۱ مقدمه

مبحث قابلیت اعتماد یکی از مفاهیمی است که سالیان متمادی مورد توجه بشریت بوده است. از دهه ۵۰ علم قابلیت اعتماد به عنوان شاخه‌ای از علوم مطرح گردید و هر روزه شاهد ظهور روشهای جدید قابلیت اعتماد و کاربرد گسترده‌تر این علم در صنایع مختلف می‌باشیم. در دهه‌های اخیر صنایع هسته‌ای گامهای بسزایی در تکامل این علم برداشته‌اند.

موقعیت‌های که علاقمند به دانستن زمان رخداد یک پیشامد خاص هستید را طول عمر می‌نامند که این مفهوم با نرخ شکست رابطه مستقیم دارد. نرخ شکست احتمال از کار افتادن بلافاصله بعد از زمان معین مثلاً t است با این فرض که واحد تحت بررسی تا زمان t کار کرده باشد.

توزیع‌های طول عمری که دارای نرخ شکست نزولی می‌باشد توسط نویسندگان و محققان متعددی مورد مطالعه و بررسی قرار گرفته است. لوماکس[۱] در سال ۱۹۵۴ مثالی در مورد فناپذیری یا رکود اقتصادی ارائه داد و پس از آن نیز نقص در سیستم هوایی هواپیماهای بوئینگ ۷۲۰ و طول عمر ماجول‌های مدار مجتمع، توسط پروشان[۲] در سال ۱۹۶۳ وساندرز و

[۱] Lomax

[۲] Proschan

مایر[۳] در سال ۱۹۸۳ گزارش شده است. کاکس[۴][۱۶] با ترکیب توزیع نمایی و پواسن به توزیع طول عمری دست یافت که دارای نرخ شکست نزولی بود و پس از آن آدامیس[۵] [۱] نیز به توزیع‌های با چنین خصوصیتی دست یافت که آن‌را از ترکیب توزیع نمایی و هندسی بدست آورد. طهماسبی و رضایی[۶] [۲۸] نیز با ترکیب توزیع نمایی و لگاریتمی، توزیع‌های طول عمر را مورد بررسی قرار دادند. در این پایان‌نامه به بررسی این ترکیب پرداخته و نشان می‌دهیم که این توزیع جانشین مناسبی برای توزیع‌های قبلی است.

اما همیشه نرخ شکست نزولی نمی‌باشد بلکه به‌صورت‌های صعودی و U شکل نیز می‌باشد. تانگ[۷] در سال ۲۰۰۲ به بررسی مدل وایبل تعمیم‌یافته برای داده‌های بقا با نرخ مخاطره U شکل پرداخته است. لی[۸] همچنین توزیع‌های طول عمر با نرخ شکست U شکل را در سال ۲۰۰۱ گزارش داده و قبل از او توسط چن[۹] در سال ۲۰۰۰ یک توزیع طول عمر جدید با نرخ شکست U شکل یا صعودی مورد مطالعه قرار گرفت. توزیع‌های که دارای نرخ شکست صعودی (IFR) می‌باشند در سال ۲۰۰۴ توسط بزونک و گائو[۱۰] [۸] و در سال ۲۰۰۱ توسط ریچگ[۱۱] [۲۶] گزارش شده است.

امروزه توزیع آماره‌های ترتیبی در بسیاری از شاخه‌های علم از جمله زمین لرزه و ژنتیک و... مورد توجه قرار گرفته است. توزیع آماره‌های ترتیبی از توزیع نرمال چوله دو متغیره و توزیع چوله دوتائی توسط جمال‌زاده و بلاکریشنا[۱۲] [۱۵] و ترکیبات خطی از متغیرهای تصادفی مستقل، ارلان و جنتون[۱۳] [۳] را می‌توان در این زمینه بیان کرد. در نظریه قابلیت اعتماد برای بررسی توزیع‌های طول عمر، اولین آماره ترتیبی توزیع را مورد بررسی قرار داده‌اند. یعنی اینکه در مسائل قابلیت، مینیمم فاصله زمانی شکست که سیستم دچار اختلال می‌شود مورد توجه قرار

[۳] Saunders and Myhre
[۴] Kus
[۵] Adamidis
[۶] Tahmasbi and Rezaei
[۷] Tang
[۸] Lie
[۹] chen
[۱۰] Belzunce and Gao
[۱۱] Rychlik
[۱۲] Jamalizadeh and Balakrishnan
[۱۳] Arellano-Valle and Genton

گرفته است. حال سوالی که مطرح می‌شود این است که آیا همیشه مینیمم مخاطره سیستم مهم می‌باشد؟ مخاطره‌های دیگری که برای سیستم بوجود می‌آیند و آن‌را تحت الشعاع قرار می‌دهند نباید مورد توجه قرار گیرند. این سوال مبنا و اساس این پایان‌نامه را شکل داده است. لذا در این پایان‌نامه مطالعه آمارهای ترتیبی توزیع‌های طول عمر نمایی—پواسن، نمایی—هندسی و بررسی خصوصیات و رفتار آن مدنظر می‌باشد.

در بخش بعدی این فصل ابتدا مفاهیم و تعاریف مورد نیاز یادآوری شده است و روشی برای تولید اعداد تصادفی معرفی گردیده است. روش نیوتون—رافسون و الگوریتم EM که روشهای عددی برای برآورد پارامترهای مجهول می‌باشند نیز معرفی شده‌اند. بطوریکه در این فصل بیشتر از مراجع [۱۲] و [۲۱] استفاده شده است. در فصل دوم یک توزیع دوپارامتری با نرخ شکست نزولی که ترکیبی از توزیع نمایی—لگاریتمی است معرفی شده و ویژگی‌های مختلف، از جمله: تابع چگالی، تابع توزیع، گشتاور، تابع بقا و ریسک و میانگین باقیمانده طول مورد مطالعه قرار گرفته است. در بخش۴ پارامترهای مجهول این توزیع با استفاده از برآورد درستنمایی ماکسیمم برآورد شده است. در بخش ۵ واریانس و کواریانس مجانبی برآوردگرها محاسبه و در نهایت دقت برآوردگرها با شبیه‌سازی نشان داده شده است. مرجع[۲۸] ، مرجع اصلی استفاده شده در این فصل می‌باشد.

در فصل سوم و چهارم این پایان‌نامه کاری است که اولین بار خودمان آن‌را مورد مطالعه و بررسی قرار داده‌ایم و ایده اصلی آن از مقاله کاکس و آدامیس گرفته شده است. به این ترتیب که در مقاله آنها با ترکیب توزیع نمایی با پواسن و هندسی ، اولین شکست یا اولین مخاطره سیستم، مورد مطالعه گرفته بود و ما در این پایان‌نامه توزیع Kامین آماره ترتیبی، ترکیبی از توزیع‌های نمایی با پواسن و نمایی با هندسی را مورد مطالعه قرار می‌دهیم و ویژگی‌های مختلف آن‌را بررسی می‌کنیم. نکته حائز اهمیت در این مطالعه، شکل تابع نرخ شکست آن بود که به‌صورت صعودی—نزولی($DIFR$) بدست آمد. سپس با استفاده از الگوریتم EM پارامترهای مجهول، برآورد شده و با استفاده از شبیه‌سازی، دقت برآوردگرهای مجانبی واریانس—کواریانس مورد بررسی قرار گرفته‌اند. نتایج با استفاده از داده‌های واقعی برای هر توزیع در انتهای هر فصل نشان داده شده است.

۲.۱ تعریف‌های اولیه

برای تعاریف بیان شده در این فصل بیشتر از کتاب آمار ریاضی مود[۳۰]، سایت wikipedia و مراجع [۱۲] و [۲۱] استفاده شده است.

تعریف ۱.۲.۱ اگر متغیر تصادفی X دارای چگالی به صورت

$$f_X(x;\lambda) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$$

باشد که در آن $\beta > 0$ است، در آن صورت گوییم X دارای توزیع نمایی (منفی) است و با نماد $X \sim E(\lambda)$ نشان می‌دهیم.

تعریف ۲.۲.۱ اگر متغیر تصادفی X دارای تابع احتمال زیر باشد:

$$f_\theta(x) = \alpha \frac{\theta^x}{x} I_{\{1,2,\dots\}}(x) \qquad 0 < \theta < 1 \ , \ \ \alpha = -[\ln(1-\theta)]^{-1}$$

گوییم X دارای توزیع سریهای لگاریتمی با پارامتر θ است و آن را با نماد $X \sim LS(\theta)$ نشان می‌دهیم.

تعریف ۳.۲.۱ گوییم متغیر تصادفی X دارای توزیع پواسن است اگر چگالی X به صورت زیر باشد.

$$f_X(x;\lambda) = \lambda \frac{e^{-\lambda x}}{x!} I_{\{0,1,\dots\}}(x)$$

که در آن، پارامتر λ در $\lambda > 0$ صدق می‌کند. این متغیر را با نماد $X \sim P(\lambda)$ نمایش می‌دهیم.

تعریف ۴.۲.۱ گوییم متغیر تصادفی X دارای توزیع هندسی است، اگر چگالی X به صورت زیر باشد.

$$f_X(x;p) = p(1-p)^{x-1} I_{\{1,2,\dots\}}(x) \quad , \quad 0 < p \leq 1$$

تعریف ۵.۲.۱ توزیع پواسن بریده شده در صفر تا $(k-1)$: همان توزیع پواسن است با این تفاوت که دامنه تابع چگالی به‌صورت $\{k, k+1, \ldots\}$ می‌باشد. تابع چگالی آن به‌صورت

$$f_Z(z;\lambda) = \frac{\lambda^z e^{-\lambda}}{\Gamma(z+1)(1-\sum_{i=0}^{k-1}\frac{e^{-\lambda}\lambda^i}{i!})} I_{\{k,k+1,\ldots\}}(z)$$

است.

تعریف ۶.۲.۱ توزیع هندسی بریده شده در صفر تا $(k-1)$: همان توزیع هندسی است با این تفاوت که دامنه تابع چگالی به‌صورت $\{k, k+1, \ldots\}$ می‌باشد. تابع چگالی آن به‌صورت $f_Z(z;p) = (1-p)p^{z-k} I_{\{k,k+1,\ldots\}}$ است.

تعریف ۷.۲.۱ تابع ریسک : تابع ریسک (بقا) یکی از مفاهیم مهم در مبحث قابلیت اطمینان است و با گذشت زمان به صورت نزولی می‌باشد زیرا با گذشت زمان قابلیت اعتماد سیستمی با طول عمر t افزایش پیدا نمی‌کند و به صورت زیر نمایش داده می‌شود.

$$\overline{F(t)} = 1 - F(t) = P(T > t)$$

تعریف ۸.۲.۱ تابع نرخ شکست : این تابع اغلب به عنوان تابع نرخ مخاطره شناخته می‌شود.

$$h(t) = \frac{f(t)}{\overline{F(t)}}$$

از طرفی نرخ شکست (از کار افتادگی) را می‌توان به ۳ دسته زیر تقسیم‌بندی کرد:

۱) نرخ شکست یکنواخت، که تابع مخاطره یا صعودی و یا نزولی است.

۲) نرخ شکست U شکل که تابع مخاطره در این حالت به شکل وان حمامی است که ابتدا کاهشی و سپس در یک بازه زمانی ثابت است و پس از آن روند افزایشی دارد.

۳) در حالت $(IDFR)$ تابع مخاطره ابتدا افزایشی و سپس کاهشی می‌باشد که در این پایان‌نامه به طور کامل در مورد آن بحث شده است.

تعریف ۹.۲.۱ میانگین باقیمانده طول عمر(MRL): میانگین باقیمانده عمر یک سیستم در صورتی که بدانیم که سیستم تا زمان t عمر کرده است را گویند و به‌صورت زیر تعریف می‌شود.

$$\mu(t) = E(T-t|T>t) = \frac{[\int_t^{\infty} \overline{F(t)}]}{\overline{F(t)}}$$

۳.۱ تولید اعداد تصادفی

متغیر تصادفی پیوسته‌ای را با تابع توزیع F در نظر بگیرید. یک روش کلی برای تولید این قبیل متغیرهای تصادفی که روش معکوس نامیده می‌شود مبتنی بر قضیه زیر است.

قضیه: فرض کنید U یک متغیر تصادفی یکنواخت در بازه $(۰, ۱)$ است. برای هر تابع توزیع پیوسته F متغیر تصادفی X که با $X = F^{-۱}(U)$ تعریف می‌شود دارای توزیع F است. قضیه فوق نشان می‌دهد که با تولید کردن متغیر تصادفی U از توزیع یکنواخت در بازه $(۰, ۱)$ و سپس قرار دادن $X = F^{-۱}(U)$ می‌توانیم متغیر تصادفی X را از تابع توزیع پیوسته F تولید کنیم.

۴.۱ روش نیوتون—رافسون

یکی از مسائلی که‌غالباً در کارهای عملی پیش می‌آید پیدا کردن ریشه‌های معادلاتی به‌صورت $f(x) = ۰$ یعنی، صفرهای تابع f است. یافتن ریشه‌های دقیق معادله(صفرهای تابع) در خیلی

از موارد سخت و پیچیده است که در چنین شرایطی، با اتکا بر تکنیک‌های عددی سعی می‌شود تا جواب‌های تقریبی معادله بدست آیند. روش نیوتون–رافسون یک روش عددی مشهور و بسیار مناسبی برای پیدا کردن ریشه‌های صفر می‌باشد. این روش یک روش عددی با تکرار است که از مشتق خطا در مراحل متوالی استفاده می‌کند. نکته قابل توجه در این روش انتخاب مقادیر اولیه است. زیرا در سرعت همگرائی بسیار مؤثر می‌باشد. با انتخاب مقدار اولیه مناسب و جایگذاری آن در رابطه

$$x_{n+۱} = x_n - \frac{f(x)}{f'(x)}$$

مقداری بدست می‌آید که به ریشه معادله نزدیکتر است و با تکرار این الگوریتم به تقریب مناسبی از ریشه واقعی دست خواهید یافت.

۵.۱ الگوریتم EM

در مسائل برآورد کلی با داده های گمشده، قسمت مشاهده شده را با ζ و قسمت گمشده را با X نشان می‌دهیم، (X,ζ) با هم داده‌های کامل فرضی را می‌سازند. برای قابل استفاده بودن از لحاظ روش شناختی، نیاز به دانستن چگونگی ارتباط توابع درستنمایی داده‌های کامل و درستنمایی داده‌های مشاهده شده، داریم. به‌طور دقیق‌تر نیاز به مشخص کردن توزیع داده‌های ناقص به‌شرط داده‌های مشاهده شده، داریم.

فرض کنید $f_\theta(y)$ و $g_\theta(x|y)$ به ترتیب تابع چگالی داده‌های مشاهده شده ζ و تابع چگالی شرطی برای داده‌های گمشده X به شرط $y = \zeta$ باشد. سپس تابع چگالی توأم X و ζ، $g_\theta(x|y)f_\theta(y)$ درستنمایی داده‌های کامل می باشد. در نتیجه $f_\theta(y)$ ، بعنوان تابعی از θ بررسی می شود، که درستنمایی داده‌های مشاهده شده است. با گرفتن لگاریتم داریم:

$$\log[g_\theta(x|y)f_\theta(y)] = \log[g_\theta(x|y)] + \log[f_\theta(y)]$$

یا

$$l(\theta; x, y) = l^*(\theta; y) + \log[g_\theta(x|y)]$$

که l، لگاریتم تابع درستنمایی داده‌های کامل و l^*، لگاریتم درستنمایی داده‌های مشاهده شده می‌باشد. این رابطه چرایی بحرانی بودن تعیین $g_\theta(x|y)$ را مشخص می‌کند. الگوریتم EM، یک روش ماکزیمم کردن لگاریتم درستنمایی داده‌های مشاهده شده، به صورت تکراری با استفاده از ماکزیمم کردن لگاریتم درستنمایی داده‌های کامل با قسمت گمشده X می‌باشد. الگوریتم با یک مقدار آغازین $\theta^{(۰)}$ در فضای پارامتری Θ شروع می‌شود و امید ریاضی $l(\theta; X, y)$، به شرط داده‌های مشاهده شده ζ و با استفاده از $\theta^{(۰)}$ به عنوان پارامتر توزیع شرطی X محاسبه می‌شود. مرحله E الگوریتم EM نامیده می‌شود.یعنی مرحله E بصورت

$$l_۱(\theta; X, \zeta) = E[l(\theta)|\zeta, \theta^{(۰)}]$$

می‌باشد. سپس، $l_۱(\theta)$ را نسبت به θ در فضای پارامتر Θ ماکزیمم می‌کنیم و مقدار پارامتر ماکزیمم را بدست می‌آوریم،$\theta^{(۱)}$ می‌نامیم. که مرحله M الگوریتم نامیده می‌شود. سپس مرحله E و M دیگری را اجرا می‌کنیم با در نظر گرفتن

$$l_۲(\theta; X, \zeta) = E[l(\theta)|\zeta, \theta^{(۱)}]$$

$\theta^{(۲)}$ را به عنوان مقداری از θ که $l_۲$ را ماکزیمم می‌کند بدست می‌آوریم. تکرار مابین این دو ادامه پیدا می‌کند تا اینکه مقدارهای $\theta^{(t)}$ و $\theta^{(t+۱)}$ در دو تکرار متوالی یکی شوند. در این نقطه الگوریتم متوقف می‌شود، زیرا که در تکرار بیشتر ما همین مقدار را پارامتری را بدست خواهیم آورد. در عمل ما تکرار کردن‌ها را وقتی که مقدارهای متوالی $\theta^{(t)}$ و $\theta^{(t+۱)}$ فقط در یک مقدار خیلی کوچک فرق داشته باشند، متوقف می‌کنیم.

فصل ۲

توزیع نمایی – لگاریتمی

۱.۲ مقدمه

مطالعه طول‌عمر موجودات زنده، ابزارها، مواد و غیره در علوم مهندسی و زیستی بسیار حائز اهمیت می‌باشند. به‌طور کلی انتظار می‌رود طول‌عمر ابزارها با گذشت زمان دارای نرخ شکست نزولی باشند. توزیع‌هایی که در آن تابع نرخ شکست با گذشت زمان کاهش می‌یابد توسط نویسندگان متعددی گزارش شده است. مثال‌های روشن آن شامل: فناپذیری یا رکود اقتصادی [1] توسط لوماکس[2] [۱۸]، نقص در سیستم تهویه هوایی هواپیماهای بوئینگ ۷۲۰ و یا نقص صنایع هوایی در نیم رساناهای که حاوی اجزای الکترونیکی بسیار هستند توسط پروشان[3] [۲۳] و طول عمر ماجول‌های مدار مجتمع، ساندرز و مایر[4] [۲۷] می‌باشد. محققان دیگری از جمله بارلو[5] [۶]، بارلو و مارشال[6] [۷]، مارشال و پروشان[7] [۲۰] داهیا و گرلند[8] [۱۱]،

[1] Business mortality
[2] Lomax
[3] Proschan
[4] Saunders and Myhre
[5] Barlowetal
[6] Barlow Marashal
[7] Marashal Proschan
[8] Dahiya Gurland

نولتی[۹] [۲۲]، گلسر[۱۰] [۱۳]، گارلند و ستارمن[۱۱] [۱۴] توزیع‌هایی با نرخ شکست نزولی را مورد مطالعه و بررسی قرار داده‌اند. کاکس[۱٦] در سال ۲۰۰۷ و طهماسبی و رضایی[۲۸] در سال ۲۰۰۸ به ترتیب با ترکیب توزیع‌های نمایی با پواسن و نمایی با لگاریتمی به توزیعی از این دسته دست یافته و آن‌را مورد مطالعه قرار داده‌اند.

در بخش ۲ این فصل، یک توزیع دو پارامتری با نرخ شکست نزولی که ترکیبی از توزیع نمایی و توزیع لگاریتمی است معرفی می‌شود. ویژگی‌های گوناگون این توزیع در بخش ۳، مورد بحث و بررسی قرار می‌گیرد. در بخش ۴، برآورد پارامترها با استفاده از روش ماکسیمم درستنمایی مورد مطالعه قرار گرفته و همچنین پارامترها با استفاده از الگوریتم EM برآورد شده‌اند. ماتریس اطلاع مشاهده شده و مورد انتظار و ماتریس واریانس—کوواریانس مجانبی برآوردگرهای ماکسیمم درستنمایی بدست آورده شده است. صحت و دقت واریانس و کواریانس مجانبی برآوردگرها با استفاده از شبیه‌سازی در بخش ۵، مورد مطالعه قرار گرفته و در بخش ٦ مطالب بیان شده با داده‌های واقعی نشان داده شده است. داده‌های واقعی که در این فصل استفاده شده است شامل ۳ سری داده می‌باشد که سری اول شامل ۲۱۳ مشاهده از شکست‌های متوالی سیستم تهویه هوای هواپیمای بوئینگ ۷۲۰ می‌باشد و دو سری دیگر مربوط به ۲۴ مشاهده از فواصل زمانی زمین لرزه‌های متوالی در قرن اخیر آنتالیای شمالی و زمین لرزه ویرانگر بم در ایران می‌باشد.

۲.۲ توزیع نمایی – لگاریتمی

فرض کنید $Y_1, Y_2, ..., Y_z$ متغیرهای تصادفی مستقل و هم‌توزیع و دارای توزیع نمایی با تابع چگالی زیر باشند.

$$f(y;\beta) = \beta e^{-\beta y} \qquad \beta > 0 \tag{۱.۲.۲}$$

[۹] McNolty

[۱۰] Gleser

[۱۱] Gurland Sethuraman

که در آن Z متغیر تصادفی و دارای توزیع لگاریتمی با تابع احتمال $P_Z(z;p) = \frac{(1-p)^z}{-z\ln p}$ به ازای $z \in N$ و $p \in (0,1)$ است. حال فرض کنید $X = \min(\{Y_i\}_{i=1}^{z})$ در این صورت تابع چگالی شرطی و تابع چگالی احتمال X به ترتیب عبارتند از:

$$f_{X|Z}(x|z;\beta) = \beta z e^{-\beta z x} \qquad (2.2.2)$$

$$f_X(x;\theta) = -\frac{\beta}{\ln p}\frac{(1-p)e^{-\beta x}}{1-(1-p)e^{-\beta x}} I_{(0,\infty)} \qquad (3.2.2)$$

زمانیکه متغیر تصادفی X دارای تابع چگالی احتمال $f_X(x;\theta)$ باشد آنگاه گفته می‌شود X دارای توزیع نمایی–لگاریتمی (EL) است. که برای محاسبه آن ابتدا توزیع توأم دو متغیر x و z که بصورت $f_{X,Z}(x,z) = f_{X|Z}(x|z)P_Z(z,p)$ می‌باشد را بدست می آوریم و سپس با جمع بستن روی مقادیر Z، توزیع حاشیه‌ای X به‌صورت زیر محاسبه می‌گردد.

$$\begin{aligned} f_X(x;\theta) &= \sum_{i=1}^{\infty} f(x;z,\beta) \\ &= -\frac{\beta}{\ln p}\sum_{z=1}^{\infty}((1-p)e^{-\beta x})^k \\ &= -\frac{\beta}{\ln p}\frac{(1-p)e^{-\beta x}}{1-(1-p)e^{-\beta x}} \end{aligned}$$

تابع چگالی EL به ازای مقادیر مختلف پارامترها در شکل ۱.۲ رسم شده است، همانطور که در شکل مشاهده می‌شود توزیع EL اکیداً نزولی است. در زیر با استفاده از مشتق تابع نیز نشان داده شده است که تابع چگالی EL نزولی می‌باشد.

$$\begin{aligned} \frac{\partial f(x;\beta)}{\partial x} &= -\frac{\beta}{\ln p}\frac{-\beta(1-p)e^{-\beta x}}{(1-(1-p)e^{-\beta x})^2} \\ &= \frac{1}{\ln p}\frac{\beta^2(1-p)e^{-\beta x}}{(1-(1-p)e^{-\beta x})^2} \end{aligned}$$

با توجه به اینکه p عددی بین صفر و یک است، لذا لگاریتم آن $(\ln p)$ منفی می‌باشد در نتیجه مشتق تابع $(\frac{\partial f(x;\beta)}{\partial x} < 0)$ منفی و تابع نزولی اکید می‌باشد.