

توزیع کسینوس هیپربولیک لگ لوژستیک و برآورد پارامترهای آن با استفاده از روش‌های ماکسیمم درست‌نمایی، بیزی و خودگردان

امید خوارزمی^۱، علی سعادت‌نیک^۲

تاریخ دریافت: ۱۳۹۶/۶/۲

تاریخ پذیرش: ۱۳۹۶/۹/۱۷

چکیده:

در این مقاله یک توزیع احتمال جدید براساس خانواده توزیع‌های کسینوس هیپربولیک $(HCF) - F$ معرفی و ویژگی‌های مختلف آماری و قابلیت اعتماد آن بررسی می‌شود. رده جدید توزیع‌های HCF از طریق ترکیب یک توزیع انباشته F با تابع کسینوس هیپربولیک به دست می‌آید. براساس توزیع پایه لگ لوژستیک، یک توزیع جدید تحت عنوان کسینوس هیپربولیک لگ لوژستیک ($HCLL$) معرفی و ویژگی‌های مختلف توزیع مانند گشتاورها، چندک‌ها، تابع مولد گشتاور، تابع نرخ شکست، میانگین باقیمانده عمر، آماره‌های ترتیبی و پارامتر تنش-مقاومت را ارائه شده‌است. برآورد پارامترهای توزیع $HCLL$ برای یک مجموعه داده واقعی از طریق سه روش ماکسیمم درست‌نمایی، بیزی و خودگردان (پارامتری و ناپارامتری) بررسی می‌شود. کارایی روش برآورد ماکسیمم درست‌نمایی را از طریق شبیه‌سازی مونت کارلو مورد ارزیابی قرار گرفته است. همچنین، در بخش کاربرد با استفاده از یک مجموعه داده واقعی برتری مدل $HCLL$ نسبت به توزیع‌های نمایی تعمیم‌یافته، وایبل، کسینوس هیپربولیک نمایی، گاما و نمایی وزنی از طریق معیارهای مختلف انتخاب مدل نشان داده شده است.

واژه‌های کلیدی: تابع کسینوس هیپربولیک، توزیع لگ لوژستیک، میانگین باقیمانده عمر، برآورد ماکسیمم درست‌نمایی، روش خودگردان.

۱ مقدمه

واقعی لازم است تا توزیع‌های جدیدتری با قابلیت انعطاف بیشتر در زمینه داده‌های علوم و مهندسی ارائه شود. برخی از مهمترین روش‌های ساخت توزیع عبارت‌اند از خانواده توزیع‌های چوله آزالینی[؟]، توزیع‌های طول عمر متناسب بخت‌ها مارشال و اولکین[؟]، خانواده توزیع‌های بتا-نرمال ایوجن و همکاران[؟] و توزیع‌های ارائه شده توسط الزاتره و همکاران[؟]. این موضوع توجه بسیاری از پژوهشگران را در سال‌های اخیر به خود جلب کرده‌اند. هدف این پژوهش معرفی یک توزیع جدید مرتبط با خانواده توزیع‌های پیوسته HCF (خوارزمی و سعادت‌نیک[؟]) و بررسی بعضی از ویژگی‌های آماری این مدل جدید است. علاوه براین، پارامترهای توزیع

در سالیان اخیر، توزیع‌های آماری به‌طور گسترده مورد توجه پژوهشگران آماری قرار گرفته‌اند. به دلیل اهمیت توزیع‌های آماری در توصیف پدیده‌های تصادفی، نظریه این شاخه از آمار روز به روز گسترش یافته و مدل‌های جدیدتری براساس توزیع‌های کلاسیک پایه ساخته می‌شوند. در دهه‌های اخیر استفاده از توزیع‌های آماری برای مدل‌سازی داده‌های واقعی در حوزه مهندسی، آمار، محیط زیست، پزشکی، جمعیت‌شناسی، اقتصاد، مالی و بیمه به‌طور روز افزونی مورد توجه قرار گرفته است. با این وجود، به دلیل تنوع داده‌های

^۱ عضو هیئت علمی گروه آمار، دانشگاه ولی عصر (عج) رفسنجان، ایران

^۲ دانش‌آموخته کارشناسی ارشد آمار، دانشگاه مازندران، ایران

خواهد بود. که در آن $x < 0$ و $a > 0$. تابع توزیع انباشته (۲) به صورت

$$G(x, a) = \frac{2e^a}{e^{2a}-1} \sinh(aF(x)). \quad (3)$$

به دست می‌آید. همچنین تابع قابلیت اطمینان $\bar{G}(x)$ و تابع نرخ شکست $h(x)$ برای توزیع HCF به ترتیب به صورت

$$\bar{G}(x) = 1 - \frac{2e^a}{e^{2a}-1} \sinh(aF(x)) \quad (4)$$

$$\bar{G}(x) = \frac{\frac{2ae^a}{e^{2a}-1} f(x) \cosh(aF(x))}{1 - \frac{2e^a}{e^{2a}-1} \sinh(aF(x))}. \quad (5)$$

تعیین می‌شوند. چندک مرتبه p ، x_p مربوط به توزیع به صورت

$$x_p = F^{-1}\left(\frac{1}{a} \operatorname{arcsinh}\left(\frac{2ae^a}{e^{2a}-1}\right)\right). \quad (6)$$

است.

۳ توزیع HCLL و برخی از ویژگی‌های آن

در این بخش با استفاده از روش ساخت توزیع HCF ، یک توزیع جدید سه پارامتری برپایه توزیع لگ لوژستیک معرفی و سپس برخی از ویژگی‌های آماری و قابلیت اعتماد آن مطالعه می‌شود.

تعریف ۱.۳. متغیر تصادفی X دارای توزیع کسینوس هیپربولیک لگ لوژستیک $HCLL$ است، هرگاه تابع چگالی احتمال آن به صورت

$$g(x; a, \beta, \lambda) = \frac{2ae^a}{e^{2a}-1} \frac{\beta \lambda^\beta x^{\beta-1}}{(\lambda^\beta + x^\beta)^2} \cosh\left(\frac{ax^\beta}{\lambda^\beta + x^\beta}\right) \quad (7)$$

است. که در آن $a > 0$ ، $\beta > 0$ ، $\lambda > 0$ و $x > 0$ و با نماد $HCLL(a, \beta, \lambda)$ نشان داده می‌شود.

شکل ۱ حالت‌هایی از توزیع $HCLL(a, \beta, \lambda)$ را برای مقادیر مختلف a ، β و λ نشان می‌دهد.

^۳ Bayesian

^۴ Bootstrap

ارائه شده از طریق سه روش ماکسیم درست‌نمایی،^۳ بیزی و بوت‌استرپ^۴ خودگردان مطالعه می‌شود. در پایان کاربرد توزیع جدید و برتری آن نسبت به برخی از مدل‌های آماری از طریق تحلیل داده‌های یک مثال واقعی نشان داده می‌شود. در ادامه این مقاله، در بخش ۲ مدل کلی HCF و ویژگی‌های مهم این خانواده از توزیع‌ها مرور می‌شود. در بخش ۳، به معرفی توزیع جدید $HCLL$ و ویژگی‌های آن می‌پردازیم. سه روش مختلف برآورد پارامترهای توزیع پیشنهادی در بخش ۴ بررسی می‌شود. شبیه‌سازی، کاربرد و مقایسه مدل پیشنهادی با توزیع‌های معمول آماری بر اساس یک مثال از داده‌های واقعی در بخش ۵ ارائه شده است. در بخش ۶ مقاله نتیجه‌گیری می‌شود.

۲ خانواده توزیع‌های کسینوس هیپربولیک F-(HCF)

در این بخش، ابتدا تعریف خانواده توزیع‌های کسینوس هیپربولیک $F-(HCF)$ بیان و سپس بعضی از ویژگی‌های مهم این خانواده مرور می‌شود. تابع کسینوس هیپربولیک برحسب تابع نمایی به صورت

$$\cosh(x) = \frac{e^x + e^{-x}}{2},$$

تعریف می‌شود. تابع $\cosh(x)$ زوج است و دارای عبارت سری تیلور با نماهای زوج برای x به صورت

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, \quad (1)$$

است.

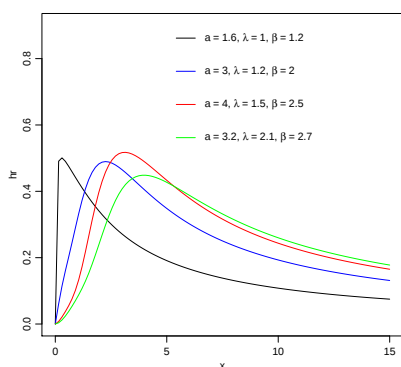
تعریف ۱.۲. فرض کنید X یک متغیر تصادفی پیوسته با تابع توزیع تجمعی $F(x)$ (CDF) باشد، آنگاه توزیع کسینوس هیپربولیک $F-(HCF)$ دارای تابع چگالی احتمال (PDF) به صورت

$$g(x, a) = \frac{2ae^a}{e^{2a}-1} f(x) \cosh(aF(x)) \quad (2)$$

محاسبه می‌شوند. تابع بقای شرطی نقش مهمی در طبقه‌بندی توزیع‌های طول عمر ایفا می‌کند. با استفاده از رابطه‌های (۴۴) و (۴۵) می‌توان تابع نرخ شکست را به صورت

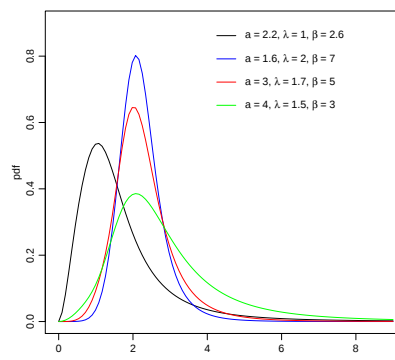
$$h(x; a, \beta, \lambda) = \frac{\frac{\gamma a e^a}{e^{\gamma a} - 1} \frac{\beta \lambda^\beta x^{\beta-1}}{(\lambda^\beta + x^\beta)^\gamma} \cosh\left(\frac{ax^\beta}{\lambda^\beta + x^\beta}\right)}{1 - \frac{\gamma a e^a}{e^{\gamma a} - 1} \sinh\left(\frac{ax^\beta}{\lambda^\beta + x^\beta}\right)} \quad (12)$$

تعیین کرد. نرخ شکست یک مفهوم کلیدی در تحلیل قابلیت اطمینان و اندازه‌گیری روند سالخوردگی است. دانستن شکل و رفتار نرخ شکست در نظریه قابلیت اطمینان، تحلیل ریسک و موارد مشابه دیگر دارای اهمیت بسیار است. در نظریه قابلیت اطمینان، مفاهیم مربوط به رفتار نرخ شکست از قبیل صعودی، نزولی، وان حمامی شکل (ابتدا نزولی و سپس صعودی) و وارون وان حمامی شکل (ابتدا صعودی و سپس نزولی) برای تحلیل داده‌های واقعی یک ابزار کلیدی محسوب می‌شوند. بر اساس همین مفاهیم رده‌های طول عمر تحت عنوان‌های IFR، BUT، DFR، و UBT رده‌بندی می‌شوند. نرخ شکست توزیع را به دلیل شکل پیچیده نمی‌توان به صورت تحلیلی بررسی کرد. با بررسی نموداری تابع نرخ شکست توزیع مشاهده می‌شود که این تابع می‌تواند رفتار صعودی، نزولی و وارون وانی شکل داشته باشد. این حالت‌ها را می‌توان در شکل‌های ۲، ۳ و ۴ مشاهده کرد.



شکل ۲. شکل نرخ شکست توزیع $HCLL(a, \beta, \lambda)$ برای برخی از مقادیر انتخاب شده پارامترها (نرخ شکست وارون وانی شکل).

۵ Stress-Strength



شکل ۱. تابع چگالی $HCLL(a, \beta, \lambda)$ برای مقادیر مختلف پارامترهای a ، β و λ .

۱.۳ خواص آماری و قابلیت اطمینان توزیع HCLL

در این بخش، برخی از خواص مهم آماری و قابلیت اعتماد توزیع HCLL مانند تابع توزیع انباشته (CDF)، تابع بقا (SF)، تابع بقای شرطی (CSF)، تابع نرخ شکست (خطر) (FR)، تابع مولد گشتاور (MGF)، میانگین باقیمانده عمر (MRL)، j -مین گشتاور، توزیع آماره‌های ترتیبی و پارامتر تنش-مقاومت بررسی می‌شود. تابع توزیع انباشته برای HCLL به صورت

$$G(x; a, \beta, \lambda) = \frac{\gamma a e^a}{e^{\gamma a} - 1} \sinh\left(\frac{ax^\beta}{\lambda^\beta + x^\beta}\right), \quad (8)$$

تعیین می‌شود. همچنین تابع بقا، چندک و تابع بقای شرطی به ترتیب به صورت

$$\bar{G}(x; a, \beta, \lambda) = 1 - \frac{\gamma a e^a}{e^{\gamma a} - 1} \sinh\left(\frac{ax^\beta}{\lambda^\beta + x^\beta}\right) \quad (9)$$

$$x_p = \lambda \left(\frac{\operatorname{arcsinh}\left(\frac{\gamma a e^a}{e^{\gamma a} - 1}\right)}{a - \operatorname{arcsinh}\left(\frac{\gamma a e^a}{e^{\gamma a} - 1}\right)} \right)^{\frac{1}{\beta}}, \quad (10)$$

$$\begin{aligned} \bar{G}(x; a, \beta, \lambda | t) &= \frac{\bar{G}(x+t; a, \beta, \lambda)}{\bar{G}(x; a, \beta, \lambda)} \\ &= \frac{1 - \frac{\gamma a e^a}{e^{\gamma a} - 1} \sinh\left(\frac{a(x+t)^\beta}{\lambda^\beta + x^\beta}\right)}{1 - \frac{\gamma a e^a}{e^{\gamma a} - 1} \sinh\left(\frac{ax^\beta}{\lambda^\beta + x^\beta}\right)}, \end{aligned} \quad (11)$$

$$M_X(t) = \left(\frac{2ae^a}{e^{2a} - 1} \right)^2 \times \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \frac{a^{2n} \lambda^j}{(2n)! j!} B\left(1 - \frac{j}{\beta}, 1 + \frac{j}{\beta} + 2n\right),$$

به دست آورد. j -مین گشتاور مرکزی توزیع HCLL را می‌توان از طریق رابطه

$$\mu_j = E(X^j) = \left(\frac{2ae^a}{e^{2a} - 1} \right)^2 \times \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \frac{a^{2n} \lambda^j}{(2n)! j!} B\left(1 - \frac{j}{\beta}, 1 + \frac{j}{\beta} + 2n\right). \quad (13)$$

محاسبه کرد. از طریق رابطه (۱۳)، می‌توان میانگین و واریانس توزیع پیشنهادی را به ترتیب به صورت

$$E(X) = \frac{2ae^a}{e^{2a} - 1} \sum_{n=0}^{\infty} \frac{a^{2n} \lambda}{(2n)!} B\left(1 - \frac{1}{\beta}, 1 + \frac{1}{\beta} + 2n\right)$$

و

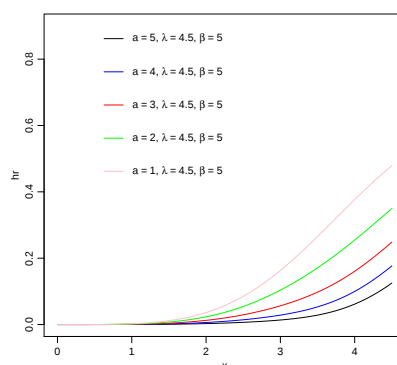
$$Var(X) = \mu_2 - E^2(X)$$

تعیین کرد. در رابطه‌های بالا $B(\cdot, \cdot)$ تابع بتا است. یکی از مهمترین ویژگی‌های هر توزیع طول عمر، میانگین باقیمانده عمر آن توزیع است، در حالت توزیع HCLL، میانگین باقیمانده عمر به صورت

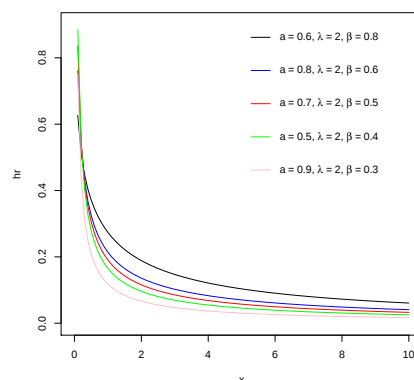
$$\begin{aligned} m(t) &= E(X - t | X > t) \\ &= \frac{2ae^a \sum_{n=0}^{\infty} \frac{\lambda a^{2n}}{(2n)!}}{a^{2a} - 1 - 2e^a \sinh\left(\frac{at^\beta}{\lambda^\beta + t^\beta}\right)} \\ &\quad \times \left(\int_0^{\frac{\lambda^\beta}{\lambda^\beta + t^\beta}} (1-z)^{\frac{1}{\beta} + 2n} z^{\frac{1}{\beta} - 1} dz \right. \\ &\quad \left. - \frac{t}{2n+1} \left(1 - \left(\frac{\lambda^\beta}{\lambda^\beta + t^\beta} \right)^{2n+1} \right) \right). \end{aligned}$$

۳.۳ آماره‌های ترتیبی و پارامتر تنش-مقاومت

محاسبه می‌شود. فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع $HCLL(a, \beta, \lambda)$ باشد. i -امین آماره ترتیبی را با $X_{i:n}$ نشان می‌دهیم. تابع چگالی احتمال آن



شکل ۳. شکل نرخ شکست توزیع $HCLL(a, \beta, \lambda)$ برای برخی از مقادیر انتخاب شده پارامترها (نرخ شکست صعودی).



شکل ۴. شکل نرخ شکست توزیع $HCLL(a, \beta, \lambda)$ برای برخی از مقادیر انتخاب شده پارامترها (نرخ شکست نزولی).

۲.۳ تابع مولد گشتاور و میانگین باقیمانده طول عمر

بعضی از ویژگی‌های مهم و خصوصیات یک توزیع را می‌توان از طریق گشتاورهای آن مانند توابع مولد گشتاور، j -امین گشتاور و میانگین باقیمانده طول عمر بررسی کرد. با استفاده از رابطه (؟؟) تابع مولد گشتاور مرتبط با رابطه (؟؟) را می‌توان به صورت

به صورت

به صورت

$$\begin{aligned} \ell(a, \beta, \lambda, x) &= \log L(a, \beta, \lambda, x) \\ &= n \log \left(\frac{e^{a_1 + a_2}}{(e^{a_1} - 1)(e^{a_2} - 1)} \right) + n \log \beta \\ &\quad + n\beta \log \lambda + (\beta - 1) \sum_{i=1}^n \log x_i \\ &\quad - 2 \sum_{i=1}^n \log(\lambda^\beta + x_i^\beta) \\ &\quad + \sum_{i=1}^n \log \left(\cosh \left(\frac{ax_i^\beta}{\lambda^\beta + x_i^\beta} \right) \right). \end{aligned}$$

نوشته می شود. برآوردگرهای ماکسیم درست نمایی برای پارامترهای a ، β و λ به ترتیب با $\hat{a}$ ، $\hat{\beta}$ و $\hat{\lambda}$ نشان داده و از طریق حل همزمان معادلات نرمال زیر حاصل می شود. به علت غیرخطی بودن این معادلات، برآوردگرهای ماکسیم درست نمایی پارامترها به صورت عددی محاسبه می شود. در این مقاله از نرم افزار آماری R و روش عددی نیوتن-رافسون برای حل معادلات نرمال استفاده شده است.

$$\begin{aligned} f_{X_{i:n}}(x) &= \frac{n!}{(i-1)!(n-i)!} \frac{e^{a_1 + a_2}}{(e^{a_1} - 1)(e^{a_2} - 1)} \frac{\beta \lambda^\beta x^{\beta-1}}{(\lambda^\beta + x^\beta)^\beta} \\ &\quad \times \cosh \left(\frac{ax^\beta}{\lambda^\beta + x^\beta} \right) \left(\frac{e^{a_1}}{e^{a_1} - 1} \sinh \left(\frac{ax^\beta}{\lambda^\beta + x^\beta} \right) \right)^{i-1} \\ &\quad \times \left(1 - \frac{e^{a_2}}{e^{a_2} - 1} \sinh \left(\frac{ax^\beta}{\lambda^\beta + x^\beta} \right) \right)^{n-i}. \end{aligned}$$

به دست می آید. فرض کنید $X_i \sim HCLL(a_i, \beta_i, \lambda_i)$ ، $i = 1, 2$ و مستقل باشند، آنگاه پارامتر تنش-مقاومت به صورت

$$\begin{aligned} P(X_1 < X_2) &= \frac{e^{a_1 + a_2}}{(e^{a_1} - 1)(e^{a_2} - 1)} \\ &\quad \times \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{a_1^{n+m+1} a_2^{n+m+1}}{(2n+1)!(2m+1)!} \\ &\quad \times \int_0^1 \frac{(1-z)^{2n+\frac{\beta_1}{\beta_2}(2m+1)}}{\left(\left(\frac{\lambda_1}{\lambda_2} \right)^{\beta_1} z^{\frac{\beta_1}{\beta_2}} + (1-z)^{\frac{\beta_1}{\beta_2}} \right)^{2m+1}} dz \end{aligned}$$

است.

۴ برآورد پارامترهای توزیع HCLL

در این بخش، سه روش برآورد برای پارامترهای a ، β و λ توزیع HCLL بیان می کنیم. این سه روش عبارت اند از روش ماکسیم درست نمایی، روش خودگردان و روش بیزی. در هر سه روش فرض می کنیم که همه پارامترها نامعلوم هستند.

۱.۴ برآورد ماکسیم درست نمایی

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از چگالی $f(x, \theta)$ باشد، آنگاه تابع درست نمایی براساس مقادیر مشاهده شده $x_1, x_2, \dots, x_n$ به صورت

$$L(\theta, x) = \prod_{i=1}^n f(x_i, \theta). \quad (14)$$

نوشته می شود. برآوردگرهای حاصل از ماکسیم سازی رابطه (۱۴)، برآورد ماکسیم درست نمایی (MLE) بردار θ نامیده می شود. در حالت توزیع HCLL لگاریتم تابع درست نمایی

۲.۴ برآوردکننده و بازه اطمینان خودگردان

عدم قطعیت در پارامترهای توزیع برازش داده را می توان توسط نمونه برداری خودگردان پارامتری (نمونه گیری از توزیع برازش داده شده) و ناپارامتری (نمونه گیری با جایگزینی از داده های اصلی) برآورد کرد (افرون و تییشیرانی [۲۰]). در بخش ۵ پارامترهای توزیع HCLL و بازه های اطمینان متناظر با استفاده از تکنیک خودگردان در دو حالت پارامتری و ناپارامتری برای یک مثال از داده های واقعی ارائه شده است.

۳.۴ برآورد بیزی و بازه اطمینان باورپذیری

در مسئله برآوردیابی، با توجه به این که از مقدار واقعی پارامتر اطلاع نداریم ممکن است با انتخاب یک برآوردگر دچار زیان شویم. این زیان را می توان با تابعی از پارامتر مورد بررسی و برآوردگر آن بیان کرد. برخی از توابع زیان شناخته شده و برآوردگرهای بیزی متناظر به صورت زیر می باشند. کالابریا

و پولیسنی [؟] را ملاحظه کنید.

فرض کنید $X_1, X_2, \dots, X_n$ یک نمونه تصادفی از توزیع $HCLL(a, \beta, \lambda)$ باشد. بنا بر این، تابع چگالی پیشین توام به صورت

۱. تابع زیان درجه دوم خطا

$$\pi(a, \beta, \lambda) = \frac{g^h c^b e^d}{\Gamma(h)\Gamma(b)\Gamma(d)} \lambda^{h-1} a^{b-1} \beta^{d-1} e^{-(g\lambda+ca+e\beta)}. \quad L(\gamma(\theta), d(\underline{x})) = (d(\underline{x}) - \gamma(\theta))^2,$$

برآورد گر بیز:

نوشته می‌شود. همچنین با توجه به رابطه توزیع پسین

$$d_B(\underline{x}) = E(\gamma(\theta)|\underline{x}).$$

$$\pi^*(a, \beta, \lambda|\underline{x}) \propto \pi(a, \beta, \lambda)f(\underline{x}, a, \beta, \lambda) \quad (15)$$

۲. تابع زیان قدر مطلق خطا

$$L(\gamma(\theta), d(\underline{x})) = |d(\underline{x}) - \gamma(\theta)|,$$

خواهیم داشت

برآورد گر بیز:

$$\begin{aligned} \pi^*(a, \beta, \lambda|\underline{x}) = & K \frac{g^h c^b e^d}{\Gamma(h)\Gamma(b)\Gamma(d)} \\ & \times \lambda^{h-1} a^{b-1} \beta^{d-1} e^{-(g\lambda+ca+e\beta)} \\ & \times \prod_{i=1}^n \frac{\gamma a e^a}{e^{\gamma a} - 1} \frac{\beta \lambda^\beta x_i^{\beta-1}}{(\lambda^\beta + x_i^\beta)^2} \cosh\left(\frac{a x_i^\beta}{\lambda^\beta + x_i^\beta}\right), \end{aligned}$$

۳. تابع زیان لاینکس

$$L(\gamma(\theta), d(\underline{x})) = (e^{c(d(\underline{x})-\gamma(\theta))} - c(d(\underline{x}) - \gamma(\theta)) - 1),$$

برآورد گر بیز:

که در آن K به صورت

$$d_B(\underline{x}) = -\frac{1}{c} \ln(E(e^{-c\gamma(\theta)}|\underline{x})).$$

$$\begin{aligned} K^{-1} = & \int_0^\infty \int_0^\infty \int_0^\infty \frac{g^h c^b e^d}{\Gamma(h)\Gamma(b)\Gamma(d)} \\ & \times \lambda^{h-1} a^{b-1} \beta^{d-1} e^{-(g\lambda+ca+e\beta)} \\ & \times \prod_{i=1}^n \frac{\gamma a e^a}{e^{\gamma a} - 1} \frac{\beta \lambda^\beta x_i^{\beta-1}}{(\lambda^\beta + x_i^\beta)^2} \cosh\left(\frac{a x_i^\beta}{\lambda^\beta + x_i^\beta}\right) da d\beta d\lambda. \end{aligned}$$

۴. تابع زیان آنتروپی تعمیم یافته

$$L(\gamma(\theta), d(\underline{x})) = \left[\left(\frac{d(\underline{x})}{\gamma(\theta)} \right)^c - c \log\left(\frac{d(\underline{x})}{\gamma(\theta)} \right) - 1 \right]$$

برآورد گر بیز:

$$d_B(\underline{x}) = (E[\gamma^{-c}(\theta)|\underline{x}])^{-\frac{1}{c}}.$$

است. با توجه به این که برآوردگرهای بیزی برای هر کدام از پارامترهای توزیع HCLL تحت توابع زیان بیان شده در بالا دارای شکل صریح نمی‌باشند، شبیه‌سازی از توزیع پسین مربوط با استفاده از الگوریتم MCMC و روش نمونه‌گیری گیس انجام و برآوردگرهای بیزی متناظر و بازه باورپذیری بیزی محاسبه می‌شود. برآوردکننده‌های بیزی پارامترهای a ، β و λ را به صورت $\hat{a}_B$ ، $\hat{\beta}_B$ و $\hat{\lambda}_B$ نشان می‌دهیم.

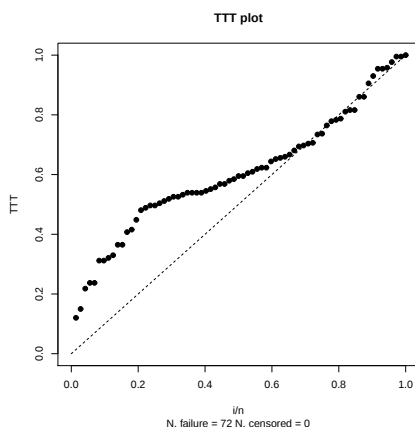
با توجه به این که پارامترهای توزیع HCLL دارای مقادیر نامنفی هستند، توزیع‌های پیشین مستقل برای هر کدام را به صورت

$$a \sim \text{Gamma}(b, c), \quad \lambda \sim \text{Gamma}(h, g),$$

$$\beta \sim \text{Gamma}(d, e).$$

در نظر می‌گیریم. همه پارامترهای h, g, d, c, b مثبت هستند.

۶۸ ۷۰ ۷۰ ۷۲ ۷۳ ۷۵ ۷۶ ۷۶ ۸۱ ۸۳ ۸۴ ۸۵ ۸۷ ۹۱ ۹۵ ۹۶ ۹۸
 ۹۹ ۱۰۹ ۱۱۰ ۱۲۱ ۱۲۷ ۱۲۹ ۱۳۱ ۱۴۳ ۱۴۶ ۱۴۶ ۱۷۵ ۱۷۵
 ۲۱۱ ۲۳۳ ۲۵۸ ۲۵۸ ۲۶۳ ۲۹۷ ۳۴۱ ۳۴۱ ۳۷۶.



شکل ۵. نمودار TTT مربوط به مجموعه داده واقعی.

قبل از تحلیل داده‌های مذکور از نمودار TTT برای تعیین اعتبار مدل خود استفاده می‌کنیم. نمودار TTT یک ابزار گرافیکی برای تشخیص شکل تابع خطر است (آرست [۱]) را ملاحظه کنید). شکل ۵ نشان می‌دهد نمودار TTT مقعر و بنا بر این تابع نرخ شکست این مجموعه داده صعودی است. در ادامه این بخش پارامترهای توزیع به سه روش ماکسیم درست‌نمایی، بیزی و خودگردان برآورد می‌شود. جدول ۳ محاسبات مربوط به برآورد پارامترها و فواصل اطمینان تحت سه روش مذکور را نشان می‌دهد. در روش ماکسیم درست‌نمایی علاوه بر برآورد پارامترها بازه اطمینان ۹۵ درصد برای هر کدام از پارامترها محاسبه شده است. در حالت برآورد بیزی، چهار تابع زیان مختلف در نظر گرفته شده است. همچنین بازه اطمینان باورپذیری برای هر کدام از پارامترها به صورت جداگانه ارائه شده است. در حالت خودگردان، برآورد پارامترها در دو حالت خودگردان پارامتری و ناپارامتری محاسبه شده است. در این حالت نیز بازه اطمینان خودگردان برای هر کدام از پارامترها آورده شده است. با توجه به نتایج جدول ۲، مشاهده می‌شود که برآورد پارامترها و بازه‌های اطمینان متناظر تحت هر سه روش به کار برده شده منجر به نتایج یکسان شده است. مقدار لگاریتم تابع درست‌نمایی در همه حالت‌ها به جز برآورد بیز

۵ کاربرد و شبیه‌سازی توزیع HCLL

در این بخش ابتدا کارایی روش ماکسیم درست‌نمایی به کمک شبیه‌سازی مونته‌کارلو با ۲۰۰۰ بار تکرار بررسی می‌شود. در این شبیه‌سازی نمونه‌هایی با اندازه n از توزیع HCLL تحت انتخاب‌های مختلف از مقادیر پارامترها تولید کرده و سپس برآوردهای ماکسیم درست‌نمایی برای پارامترهای α ، β و λ محاسبه می‌شود. از جدول ۱ مشاهده می‌شود که با افزایش اندازه نمونه همان‌طور که انتظار می‌رود میانگین توان دوم خطا و اریبی روند کاهشی دارد. بنا بر این می‌توان پذیرفت که برآورد گره‌های ماکسیم درست‌نمایی سازگار هستند.

در ادامه این بخش با استفاده از داده‌های واقعی اهمیت و کاربرد توزیع HCLL را نشان می‌دهیم. ابتدا، پارامترهای توزیع HCLL برای یک مجموعه داده واقعی از طریق سه روش ماکسیم درست‌نمایی، بیزی و خودگردان (پارامتری و ناپارامتری) برآورد و سپس برتری توزیع HCLL را نسبت به توزیع‌های گاما، وایبل، نمایی تعمیم‌یافته (GE)، نمایی کسینوس هیپربولیک HCE و نمایی وزنی (WE) نشان می‌دهیم.

$$f_{\text{Gamma}}(x) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x}, \quad x \geq 0,$$

$$f_{\text{Weibull}}(x) = \frac{\beta}{\lambda^\beta} x^{\beta-1} e^{-(\frac{x}{\lambda})^\beta}, \quad x \geq 0,$$

$$f_{\text{GE}}(x) = \alpha \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha-1}, \quad x \geq 0,$$

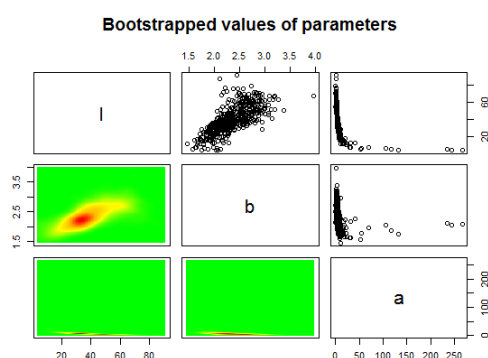
$$f_{\text{HCE}}(x) = \frac{2a e^a}{e^{2a} - 1} \lambda e^{-\lambda x} \cosh(a(1 - e^{-\lambda x})), \quad x \geq 0,$$

$$f_{\text{WE}}(x) = \frac{\alpha + 1}{\alpha} \lambda e^{-\lambda x} (1 - e^{-\lambda x}), \quad x \geq 0.$$

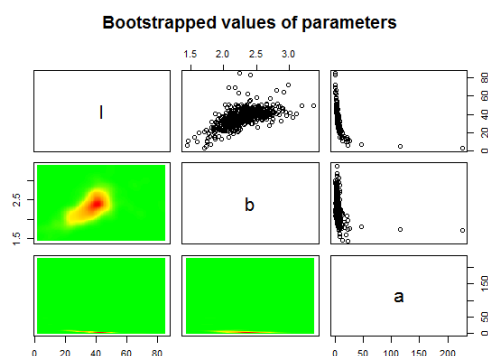
مثال ۱.۵. داده‌های کاربردی این بخش مربوط به مطالعه زمان بقای خوکچه‌های هندی در مقابل مقادیر مختلف تزریق باسیل سل است که گوپتا و کوندو [۲] برای نشان دادن برتری توزیع نمایی وزنی بر توزیع‌های گاما، وایبل و نمایی تعمیم‌یافته مورد استفاده قرار داده‌اند. این مجموعه داده‌ها عبارت‌اند از:

۱۲ ۱۵ ۲۲ ۲۴ ۲۴ ۳۲ ۳۲ ۳۳ ۳۴ ۳۸ ۳۸ ۴۳ ۴۴ ۴۸ ۵۲ ۵۳ ۵۴
 ۵۴ ۵۵ ۵۶ ۵۷ ۵۸ ۵۸ ۵۹ ۶۰ ۶۰ ۶۰ ۶۰ ۶۱ ۶۲ ۶۳ ۶۵ ۶۵ ۶۷

مدل جدول ۲، مشاهده می‌شود که کمترین مقدار AIC^* ، W^* و $HCLL$ مربوط به توزیع $HCLL$ است. واضح است که این معیارها برای مدل $HCLL$ به‌طور چشمگیری کاهش یافته‌اند. با توجه به نتایج این جدول، کفایت مدل مورد بررسی نسبت به مدل‌های رایج دیگر تأیید می‌شود. بنا براین توزیع پیشنهادی، بهترین مدل برای توصیف و تحلیل داده‌های مذکور است. بافت‌نگاشت نسبی مربوط به داده‌ها و توابع چگالی برازش داده شده $HCLL$ ، گاما، وایبل، کسینوس هیپربولیک نمایی، نمایی تعمیم‌یافته و نمایی وزنی، نمودار تابع انباشته تجربی و توابع انباشته برازش داده شده، نمودارهای $P - P$ و $Q - Q$ برای $HCLL$ و توزیع‌های مذکور در شکل ۹ آورده شده است. این نمودارها نتایج جدول ۳ را تأیید می‌کنند.



شکل ۶. نمودار برآورد پارامترهای توزیع $HCLL$ در حالت خودگردان پارامتری.



شکل ۷. نمودار برآورد ناپارامترهای توزیع $HCLL$ در حالت خودگردان ناپارامتری.

تحت تابع زیان لاینکس با مقدار $c = 3$ تقریباً یکسان هستند. تفاوت مقدار لگاریتم تابع درست‌نمایی تحت تابع زیان لاینکس به علت مقدار پارامتر c است و باعث بیش‌برآوردی و کم‌برآوردی می‌شود، بنا براین چنین نتیجه‌ای منطقی به نظر می‌رسد. شکل‌های ۶ و ۷ به‌ترتیب نمودارهای مربوط به برآورد خودگردان پارامتری و ناپارامتری پارامترهای سه‌گانه توزیع $HCLL$ را نشان می‌دهند. همچنین شکل ۸ نمودار چگالی پسین مربوط به پارامترهای سه‌گانه توزیع $HCLL$ را نشان می‌دهد.

در پایان این بخش توزیع جدید را به داده‌های مثال ۱.۵ برازش داده و آن را با چگالی‌های وایبل، نمایی تعمیم‌یافته، گاما، کسینوس هیپربولیک نمایی HCE و نمایی وزنی مقایسه می‌کنیم. یک مسئله مهم برای برازش داده‌ها توسط یک توزیع، آزمون نیکویی برازش است. برای این منظور، آزمون‌هایی بر مبنای اختلاف بین تابع توزیع برازش داده شده و تابع توزیع تجربی ارائه شده است. معیار آکائیکه $AIC = -2 \ln(L) + 2K$ که در آن L مقدار تابع درست‌نمایی و K تعداد پارامترهای مدل است، تعریف می‌شود. همچنین آماره‌های کرامر-ون میزز (W^*) و اندرسون-دارلینگ (A^*) به ترتیب به‌صورت

$$w^* = \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 f(x) dx$$

و

$$A^* = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\log(F(x_i)) + \log(1-F(x_{n-i+1}))].$$

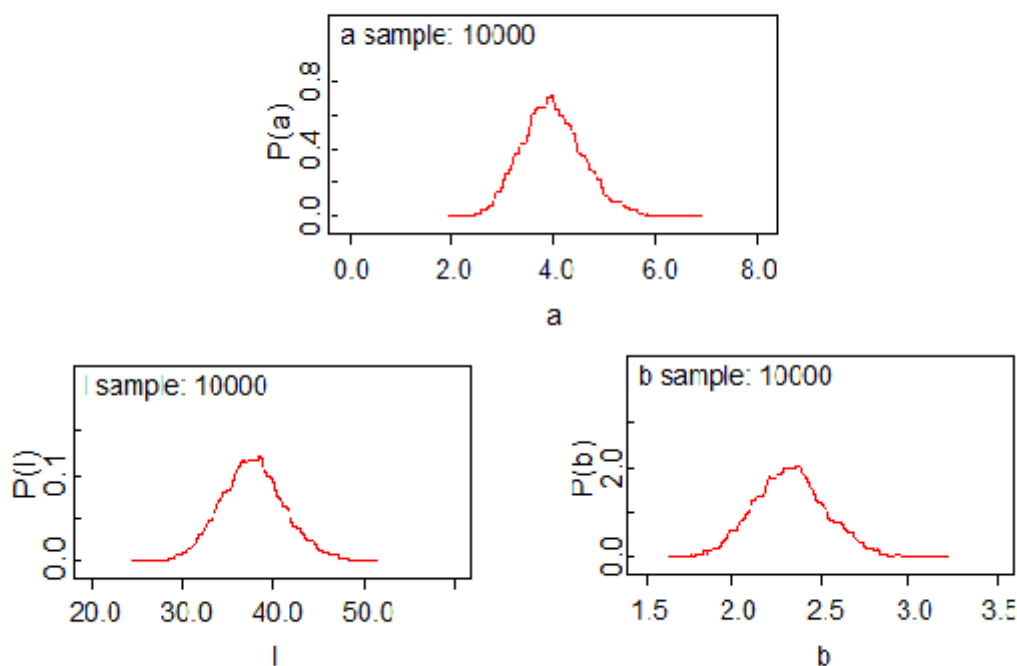
تعریف می‌شوند. F نشان دهنده تابع توزیع برازش داده شده است. در مسائل کاربردی هرچه مقدار این معیارها کوچک‌تر باشند، برازش بهتری از توزیع مورد بررسی فراهم آورده می‌شود. جدول ۲ برآورد ماکسیمم درست‌نمایی پارامترها، معیار لگاریتم تابع درست‌نمایی، (LL) معیار اطلاع آکائیکه (AIC)، آماره آزمون‌های کرامر-ون میزز (W^*) و اندرسون-دارلینگ (A^*) را نشان می‌دهد. با توجه به معیارهای انتخاب

جدول ۱. میانگین توان دوم خطا و (اریبی) برای براورد پارامترهای توزیع HCLL

$\lambda = 0.5$	$\beta = 2$	$a = 2$		
$(-0.0035) / 0.266$	$(-0.0104) / 0.0716$	$(0.07903) / 2.4127$	۵۰	n
$(-0.0006) / 0.224$	$(-0.0169) / 0.0550$	$(0.02083) / 2.3968$	۷۵	
$(0.0017) / 0.198$	$(-0.0238) / 0.372$	$(0.1610) / 1.7665$	۱۰۰	
$(0.0054) / 0.174$	$(-0.0363) / 0.225$	$(0.060) / 1.2251$	۲۰۰	
$\lambda = 2$	$\beta = 4$	$a = 1$		
$(-0.0661) / 0.509$	$(0.0880) / 2.696$	$(0.1716) / 0.8693$	۵۰	n
$(-0.0695) / 0.457$	$(0.0276) / 1.599$	$(0.1656) / 0.8261$	۷۵	
$(-0.0681) / 0.443$	$(-0.0058) / 1.176$	$(0.1669) / 0.8113$	۱۰۰	
$(-0.0639) / 0.380$	$(-0.0224) / 0.634$	$(0.1323) / 0.8143$	۲۰۰	
$\lambda = 0.5$	$\beta = 1$	$a = 0.5$		
$(-0.0473) / 0.249$	$(0.0337) / 0.169$	$(0.2700) / 0.6345$	۵۰	n
$(-0.0518) / 0.220$	$(0.0240) / 0.1056$	$(0.2749) / 0.6339$	۷۵	
$(-0.0537) / 0.203$	$(0.0163) / 0.075$	$(0.2744) / 0.6485$	۱۰۰	
$(-0.0545) / 0.1632$	$(0.0107) / 0.040$	$(0.2347) / 0.6227$	۲۰۰	
$\lambda = 0.2$	$\beta = 4$	$a = 0.1$		
$(-0.0096) / 0.0003$	$(0.1508) / 2.787$	$(0.6250) / 0.9154$	۵۰	n
$(-0.0095) / 0.0003$	$(0.1154) / 1.833$	$(0.6105) / 0.9103$	۷۵	
$(-0.0087) / 0.00029$	$(0.0967) / 1.304$	$(0.5919) / 0.89303$	۱۰۰	
$(-0.0083) / 0.0002$	$(0.0629) / 0.649$	$(0.5440) / 0.7905$	۲۰۰	
$\lambda = 0.1$	$\beta = 0.2$	$alpha = 1$		
$(0.1596) / 0.5164$	$(0.0038) / 0.0006$	$(0.1379) / 0.8312$	۵۰	n
$(0.1147) / 0.1455$	$(0.0019) / 0.0003$	$(0.0852) / 0.7717$	۷۵	
$(0.1077) / 0.1218$	$(0.0012) / 0.0002$	$(0.0414) / 0.7404$	۱۰۰	
$(0.0658) / 0.467$	$(-0.0008) / 0.0001$	$(0.0580) / 0.7506$	۲۰۰	

جدول ۲. برآورد گرهای ماکسیمم درست‌نمایی پارامترها برای داده‌های واقعی.

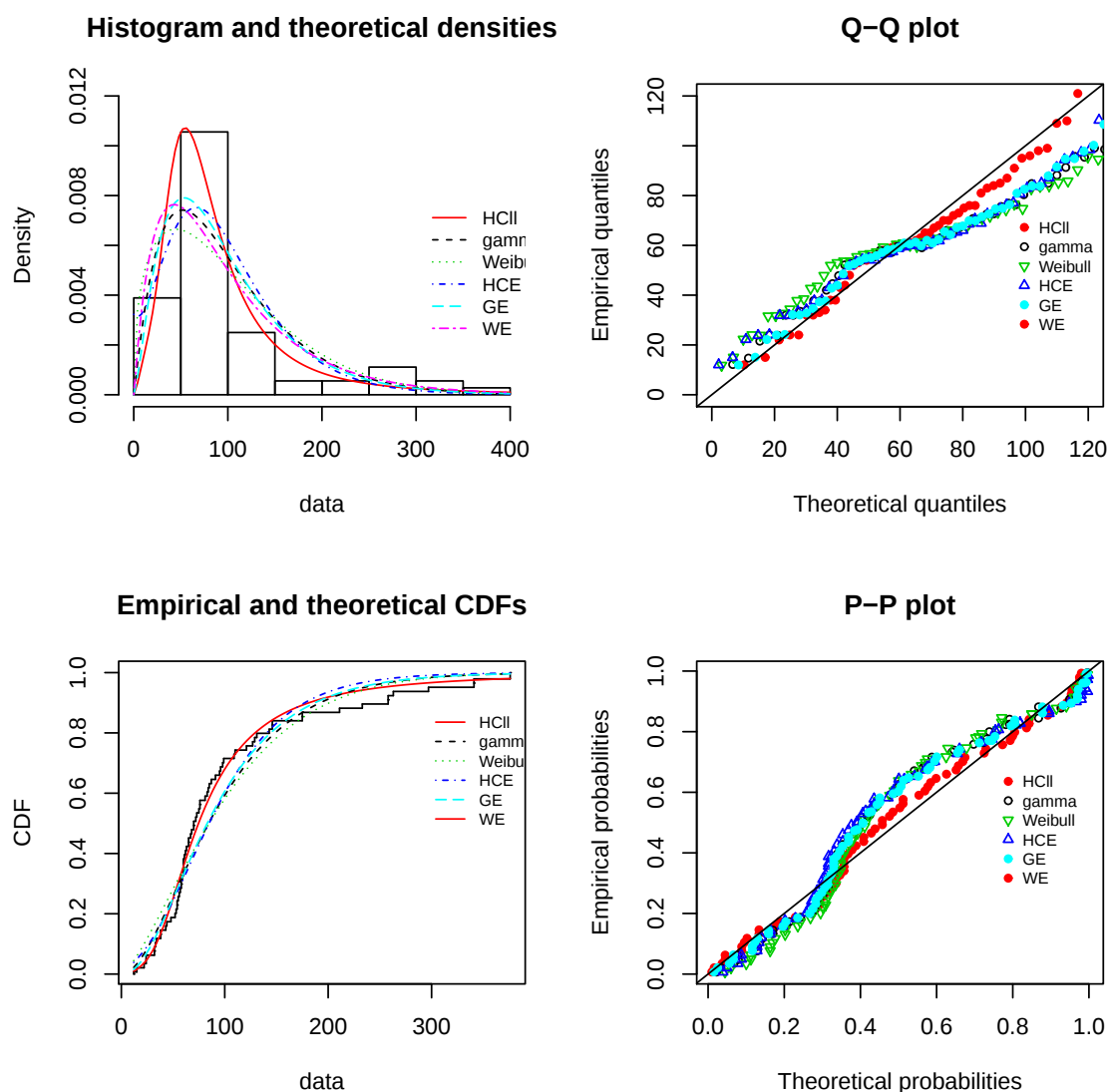
مدل	برآورد پارامترها	LL	AIC	A^*	W^*
$HCLL$	$(\hat{a}, \hat{\beta}, \hat{\lambda}) = (4/0.3, 2/3, 37/68)$	-389/3935	784/786	0/505	0/08
HCE	$(\hat{a}, \hat{\lambda}) = (3/894, 0/02)$	-397/4949	798/989	2/35	0/37
$Gamma$	$(\hat{\alpha}, \hat{\lambda}) = (2/081, 0/02)$	-394/2476	792/495	1/89	0/35
$Weibull$	$(\hat{\beta}, \hat{\lambda}) = (1/392, 110/5)$	-397/147	798/295	2/37	0/43
GE	$(\hat{\alpha}, \hat{\lambda}) = (2/476, 0/016)$	-393/110	790/22	1/64	0/30
WE	$(\hat{\alpha}, \hat{\lambda}) = (1/63, 0/014)$	-393/568	791/138	1/62	0/29



شکل ۸. نمودارهای مربوط به توزیع پسین پارامترهای توزیع $HCLL$.

جدول ۳. برآورد پارامترهای توزیع HCLL با استفاده از سه روش ماکسیمم درست‌نمایی، بیزی و خودگردان.

$\hat{\lambda}$	$\hat{\beta}$	$\hat{a}$	روش برآورد	
۳۷/۶۸ (۵۵/۶۶۸ و ۱۹/۶۹۱)	۲/۳۰۲ (۲/۸۳۷ و ۱/۷۶۷)	۴/۰۳۰ (۶/۶۲ و ۱/۴۴)	نقطه‌ای بازه اطمینان	ماکسیمم درست‌نمایی
		----- -۳۸۹/۳۸۳		مقدار لگاریتم تابع درست‌نمایی
$\hat{\lambda}_B$	$\hat{\beta}_B$	$\hat{a}_B$	روش برآورد بیزی	
۳۷/۸۲ (۴۵/۰۸ و ۳۱/۱۳)	۲/۳۲۷ (۲/۷۵ و ۱/۹۴۱)	۳/۹۷۵ (۵/۲۶۹ و ۲/۸۹۶)	نقطه‌ای بازه اطمینان باورپذیر	تابع زیان درجه دوم خطا
		----- -۳۸۹/۴۰۶		مقدار لگاریتم تابع درست‌نمایی
۳۷/۷۷ (۴۵/۰۸ و ۳۱/۱۳)	۲/۳۲۲ (۲/۷۵ و ۱/۹۴۱)	۳/۹۶۸ (۵/۲۶۹ و ۲/۸۹۶)	نقطه‌ای بازه اطمینان باورپذیر	تابع زیان قدر مطلق خطاها
		----- -۳۸۹/۴۰۶		مقدار لگاریتم تابع درست‌نمایی
۲۷/۸۷۶ (۳۱/۱۱ و ۲/۷۴۹)	۲/۴۱۰ (۲/۷۴۹ و ۲/۳۲۲)	۳/۲۱۵ (۵/۲۶۹ و ۲/۸۹۶)	نقطه‌ای بازه اطمینان باورپذیر	تابع زیان لاینکس ($c = 3$)
		----- -۴۰۵/۸۲		مقدار لگاریتم تابع درست‌نمایی
۳۷/۵۹۱ (۴۵/۰۴ و ۳۱/۱۰)	۲/۳۱۲ (۱/۹۴۱ و ۲/۷۵)	۳/۸۷ (۵/۲۶۹ و ۲/۸۹۶)	نقطه‌ای بازه اطمینان باورپذیر	تابع زیان آنتروپی تعمیم‌یافته ($c = 0.5$)
		----- -۳۸۹/۴۶۳		مقدار لگاریتم تابع درست‌نمایی
$\hat{\lambda}_P$	$\hat{\beta}_P$	$\hat{a}_P$	روش برآورد خودگردان پارامتری	
۳۷/۳۰ (۷۰.۴۳ و ۱۰/۸۸۹)	۲/۲۹۸ (۳/۰۶۶ و ۱/۷۱۱۵)	۴/۰۵۶ (۲۷/۳۹۱ و ۱/۰۷۳)	نقطه‌ای بازه اطمینان چندکی	خودگردان پارامتری
		----- -۳۸۹/۳۹۶		مقدار لگاریتم تابع درست‌نمایی
$\hat{\lambda}_{NP}$	$\hat{\beta}_{NP}$	$\hat{a}_{NP}$	روش برآورد خودگردان ناپارامتری	
۳۸/۲۰ (۵۲/۷۹۱ و ۱۴/۷۰۳)	۲/۲۹۹ (۲/۸۷۳ و ۱/۷۷۱)	۴/۰۴۵ (۱۳/۷۷۱ و ۲/۱۲۸)	نقطه‌ای بازه اطمینان چندکی	خودگردان ناپارامتری
		----- -۳۸۹/۴۱۷		مقدار لگاریتم تابع درست‌نمایی



شکل ۹. نمودار برآورد پارامترهای توزیع $HCLL$ در حالت خود گردان پارامتری .

۶ نتیجه گیری

بسیار قوی برای دیگر توزیع‌های شناخته شده به کار برود. برآورد پارامترها به سه روش روش ماکسیمم درست‌نمایی، خود گردان (پارامتری و ناپارامتری) و بیزی بررسی شده است. علاوه بر این با تحلیل یک مثال واقعی از داده‌های طول عمر مشاهده شده است که توزیع $HCLL$ برازش بهتری را نسبت به توزیع‌های وایبل، گاما، نمایی تعمیم یافته، کسینوس هیپربولیک نمایی و نمایی وزنی فراهم می‌آورد.

در این مقاله یک توزیع احتمال جدید تحت عنوان $HCLL$ پیشنهاد شده است. مدل $HCLL$ از ترکیب توزیع مبنای لگ لوژستیک با تابع کسینوس هیپربولیک ساخته می‌شود. انتظار می‌رود که این توزیع پیشنهادی به‌طور گسترده در زمینه شاخه‌های مختلف علوم و مهندسی مورد استفاده قرار گیرد. توزیع $HCLL$ می‌تواند به عنوان یک مدل رقیب

مراجع

- [1] Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, **36**(1), 106-108.
- [2] Alzaatreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, **71**(1), 63-79.
- [3] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian journal of statistics*, 171-178.
- [4] Calabria, R., and Pulcini, G. (1996). Point estimation under asymmetric loss functions for left-truncated exponential samples. *Communications in Statistics-Theory and Methods*, **25**(3), 585-600.
- [5] Efron, B., and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- [6] Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, **31**(4), 497-512.
- [7] Gupta, R. D., and Kundu, D. (2009). A new class of weighted exponential distributions. *Statistics*, **43**(6), 621-634.
- [8] Marshall, A. W., and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, **84**(3), 641-652.
- [9] Kharazmi, O., and Saadatinik, A. (2016). Hyperbolic Cosine-F Family of Distributions with an Application to Exponential Distribution. *Gazi University Journal of Science*, **29**(4), 811-829.
- [10] Team, R. C. (2014). R: A language and environment for statistical computing. *Computing*.