

مقایسه دو توزیع آماری کوماراسوآمی و کوماراسوآمی نمایی شده و ارائه کاربردهای آن‌ها

آنیتا عبدالهی نانوایشه
گروه آمار، دانشگاه علامه طباطبایی

چکیده

روش‌های ریاضی و توزیع‌های آماری، نتایجی دقیق در مسائل کاربردی ارائه می‌دهند که در مقاله حاضر به دو نمونه از این توزیع‌های آماری و کاربردهای عملی آن‌ها پرداخته شده است. در این مقاله با فرض این که X از مدل کوماراسوآمی و کوماراسوآمی نمایی شده پیروی می‌کند، پس از توضیح مختصری در مورد ویژگی‌های این دو توزیع، اعتبار این مدل‌ها برای مجموعه‌ای از داده‌های واقعی بررسی شد. نتایج نیز مناسب بودن مدل‌های ارائه شده را برای مجموعه داده‌ها تأیید می‌کنند.

واژه‌های کلیدی: توزیع کوماراسوآمی، توزیع کوماراسوآمی نمایی شده، گشتاورها و برآورد درست‌نمایی ماکسیمم.

۱ مقدمه

نمایی شده و ارائه کاربردهایی از آن‌ها می‌باشد. ابتدا برخی از ویژگی‌های این توزیع‌ها را محاسبه خواهیم کرد و در نهایت با استفاده از دو مثال کاربردی، اعتبار این مدل‌ها را برای مجموعه داده‌های واقعی بررسی می‌کنیم.

توزیع کوماراسوآمی یک توزیع تک متغیره، دو پارامتری پیوسته با تکیه‌گاه $(0, 1)$ است که توسط کوماراسوآمی در سال ۱۹۸۰ ارائه شد [۲]. این توزیع برای اولین بار در موضوعات آب و هوا مورد بررسی قرار گرفت و سپس برخی از ویژگی‌های آن مانند امید ریاضی، واریانس، تابع مولد گشتاور، تابع نرخ مخاطره و ... ارائه شدند. این توزیع در بسیاری از پدیده‌های طبیعی که شامل حدود بالا و پائین، مانند قد افراد، نمرات امتحانی، درجه حرارت هوا، داده‌های آبی و ... هستند کاربرد دارد. هدف از این مقاله، مرور ویژگی‌های مدل کوماراسوآمی و کوماراسوآمی

۲ توزیع کوماراسوآمی و ویژگی‌های آن

اگر X دارای توزیع کوماراسوآمی با پارامترهای α و β باشد در این صورت تابع توزیع آن به صورت زیر ارائه

می‌شود:

$$(۱)$$

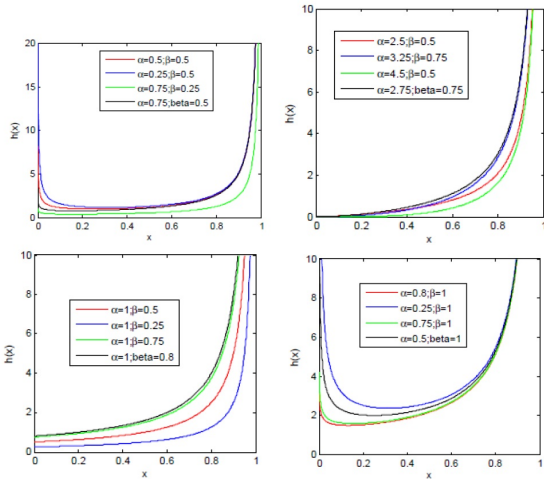
$$F(X) = 1 - (1 - x^\alpha)^\beta, \quad 0 < x < 1, \alpha, \beta > 0$$

تابع چگالی احتمال این توزیع نیز عبارت است از:

$$(۲)$$

$$f(x) = \alpha\beta x^{\alpha-1} (1 - x^\alpha)^{\beta-1}, \quad 0 < x < 1, \alpha, \beta > 0$$

تابع چندک و تابع نرخ شکست آن نیز به صورت زیر است:



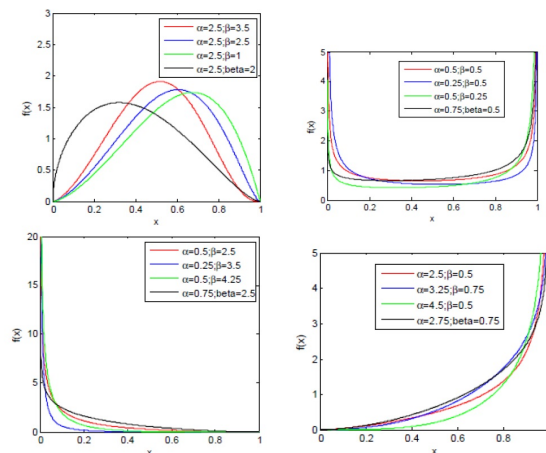
شکل ۲: تابع نرخ شکست توزیع کوماراسوآمی به ازای مقادیر مختلف α و β

$$Q(u) = F^{-1}(u) = \left\{ 1 - \left(1 - u^{\frac{1}{\beta}} \right) \right\}^{1/\alpha},$$

$$h(x) = \frac{\alpha\beta x^{\alpha-1} (1 - x^\alpha)^{\beta-1}}{(1 - x^\alpha)^\beta}, \quad x > 0$$

در نمودارهای ۱ و ۲، شکل تابع چگالی احتمال و تابع نرخ شکست این توزیع برای مقادیر مختلف α و β در محیط نرم‌افزار متلب^۱ نشان داده شده است.

همانطور که در شکل‌ها نشان داده شده است، این توزیع دارای نرخ شکست افزایشی و u شکل می‌باشد.



حال امید ریاضی و واریانس و سایر ویژگی‌های این توزیع را ارائه می‌کنیم [۲، ۳، ۴، ۵، ۶، ۷]. اگر X دارای توزیع کوماراسوآمی باشد در این صورت گشتاورهای X از رابطه زیر به دست می‌آیند:

$$E[X^r] = \beta B\left(1 + \frac{r}{\alpha}, \beta\right).$$

شکل ۱: تابع چگالی احتمال توزیع کوماراسوآمی به ازای مقادیر مختلف α و β

به همین ترتیب امید ریاضی، واریانس و ضریب تغییرات

^۱Matlab

آن نیز به صورت زیر است:

زیر به دست می‌آید:

$$L(\alpha, \beta) = n \log \alpha \beta + (\alpha - 1) \sum_{i=1}^n \log(x_i) + (\beta - 1) \sum_{i=1}^n \log(1 - x_i^\alpha). \quad (3)$$

$$E[X] = \int_0^1 x \alpha \beta x^{\alpha-1} (1 - x^\alpha)^{\beta-1} dx = \beta B\left(1 + \frac{1}{\alpha}, \beta\right).$$

که در این صورت برآورد درست نمایی ماکسیمم پارامتر β به صورت زیر خواهد بود

$$Var[X] = E[X^2] - E^2[X] = \beta B\left(1 + \frac{2}{\alpha}, \beta\right) - \left\{ \beta B\left(1 + \frac{1}{\alpha}, \beta\right) \right\}^2.$$

$$\hat{\beta} = \frac{-n}{\sum_{i=1}^n \log(1 - x_i^\alpha)}. \quad (4)$$

$$Cv[X] = Cv[Y] = \frac{\sigma}{\mu} = \frac{\sqrt{\beta B\left(1 + \frac{2}{\alpha}, \beta\right) - \left\{ \beta B\left(1 + \frac{1}{\alpha}, \beta\right) \right\}^2}}{\beta B\left(1 + \frac{1}{\alpha}, \beta\right)}.$$

و برآورد درست‌نمایی ماکسیمم α از حل معادلات زیر به دست می‌آید [۱]:

که $B(\alpha, \beta)$ نشان‌دهنده توزیع بتا است.

$$S(\alpha) = \frac{n}{\alpha} \left[1 + T_1(\alpha) + \frac{T_2(\alpha)}{T_3(\alpha)} \right] = 0, \quad (5)$$

ضریب چولگی و ضریب برجستگی X به ترتیب $\gamma_1[X]$ و $\gamma_2[X]$ به صورت زیر است:

$$T_1(\alpha) = n^{-1} \sum_{i=1}^n \frac{\log x_i^\alpha}{1 - x_i^\alpha},$$

$$T_2(\alpha) = n^{-1} \sum_{i=1}^n \frac{x_i^\alpha \log x_i^\alpha}{1 - x_i^\alpha},$$

$$T_3(\alpha) = n^{-1} \sum_{i=1}^n \log(1 - x_i^\alpha).$$

$$\begin{aligned} \gamma_1[X] &= \frac{E(X - E(X))^3}{(E(X - E(X))^2)^{3/2}} \\ &= \frac{B\left(1 + \frac{1}{\alpha}, \beta\right) - 6B\left(1 + \frac{1}{\alpha}, 2\beta\right)}{B\left(1 + \frac{1}{\alpha}, \beta\right) - 2B\left(1 + \frac{1}{\alpha}, 2\beta\right)} \\ &\quad + \frac{6B\left(1 + \frac{1}{\alpha}, 3\beta\right)}{B\left(1 + \frac{1}{\alpha}, \beta\right) - 2B\left(1 + \frac{1}{\alpha}, 2\beta\right)}. \end{aligned}$$

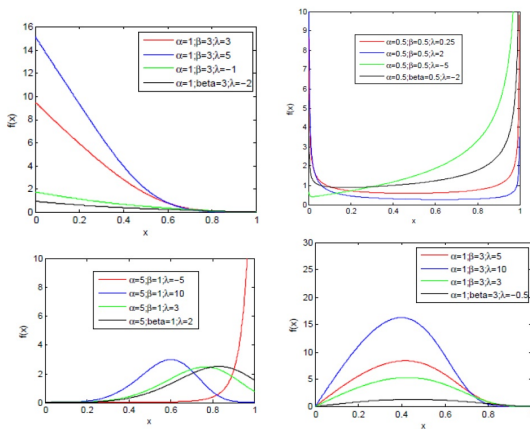
$$\begin{aligned} \gamma_2[X] &= \frac{E(X - E(X))^4}{(E(X - E(X))^2)^2} \\ &= \frac{B\left(1 + \frac{1}{\alpha}, \beta\right) - 12B\left(1 + \frac{1}{\alpha}, 2\beta\right)}{B\left(1 + \frac{1}{\alpha}, \beta\right) - 2B\left(1 + \frac{1}{\alpha}, 2\beta\right)} \\ &\quad + \frac{30B\left(1 + \frac{1}{\alpha}, 3\beta\right) - 20B\left(1 + \frac{1}{\alpha}, 4\beta\right)}{B\left(1 + \frac{1}{\alpha}, \beta\right) - 2B\left(1 + \frac{1}{\alpha}, 2\beta\right)}. \end{aligned}$$

لگاریتم درست‌نمایی ماکسیمم پارامترهای α و β به صورت

۳ توزیع کوماراسوآمی نمایی شده و ویژگی‌های آن

اگر X دارای توزیع کوماراسوآمی باشد با استفاده از تبدیل زیر می‌توان این توزیع را به صورت نمایی شده نوشت:

$$H(x) = \begin{cases} \frac{1 - e^{-\lambda F(x)}}{1 - e^{-\lambda}} & x \neq 0 \\ F(x) & x = 0 \end{cases}, \quad \lambda \in R$$



شکل ۳: تابع چگالی احتمال توزیع کوماراسوآمی نمایی شده به ازای مقادیر مختلف α و β

حال فرض کنید X دارای توزیع کوماراسوآمی نمایی شده باشد در این صورت برخی از ویژگی‌های آن از جمله تابع مولد گشتاور، ضریب چولگی و ضریب برجستگی X به ترتیب $\gamma_1[X]$ و $\gamma_2[X]$ به صورت زیر است $[\lambda]$:

$$E[X^r] = \frac{\lambda}{1 - e^{-\lambda}} \sum_{j=0}^{\infty} \frac{(-\lambda)^j}{j!} \int_0^1 \left(1 - (1-u)^{\frac{1}{\beta}}\right)^{\frac{r}{\alpha}} u^j du = \int_0^1 \frac{\alpha\beta\lambda x^{r+\alpha-1} (1-x^\alpha)^{\beta-1} e^{\lambda(1-x^\alpha)^\beta}}{e^\lambda - 1} dx,$$

با استفاده از بسط مک‌لورن $e^{\lambda(1-x^\alpha)^\beta}$ داریم:

$$E[X^r] = \frac{1}{e^\lambda - 1} \sum_{k=0}^{\infty} \frac{(\lambda)^{K+1}}{(K+1)!} EY_k^r,$$

که Y_k توزیع کوماراسوآمی با پارامترهای α و $\beta(k+1)$ است. لذا در نهایت داریم:

$$E[X^r] = \frac{\lambda\beta}{e^\lambda - 1} \sum_{k=0}^{\infty} \frac{(\lambda)^K}{(K)!} B\left(1 + \frac{r}{\alpha}, \beta(k+1)\right).$$

در این بخش تعمیمی سه پارامتری از توزیع کوماراسوآمی را ارائه می‌کنیم. این تابع بسیار انعطاف‌پذیر بوده و جایگزین مناسبی برای توزیع‌های بتا و کوماراسوآمی است. تابع چگالی احتمال، تابع توزیع تجمعی، تابع چندک و تابع نرخ شکست این توزیع به صورت زیر است:

$$F(X) = \frac{e^\lambda - e^{\lambda(1-x^\alpha)^b}}{e^\lambda - 1}, \quad (6)$$

$$f(x) = \frac{ab\lambda x^{\alpha-1} (1-x^\alpha)^{b-1} e^{\lambda(1-x^\alpha)^b}}{e^\lambda - 1}, \quad (7)$$

$$Q(u) = \left(1 - \left(1 + \frac{\log(1-u(1-e^{-\lambda}))}{\lambda}\right)^{1/\beta}\right)^{1/\alpha},$$

$$h(x) = \frac{\alpha\beta\lambda x^{\alpha-1} (1-x^\alpha)^{\beta-1}}{1 - e^{-\lambda(1-x^\alpha)^\beta}}, \quad \alpha, \beta > 0, \lambda \neq 0$$

که در آن‌ها، $0 < x < 1$.

در نمودارهای ۳ و ۴ شکل تابع چگالی احتمال و تابع نرخ شکست این توزیع برای مقادیر مختلف α ، β و λ در محیط نرم‌افزار متلب نشان داده شده است.

همان‌طور که مشخص است تابع نرخ شکست این توزیع به صورت افزایشی، کاهشی و u شکل معکوس می‌باشد.

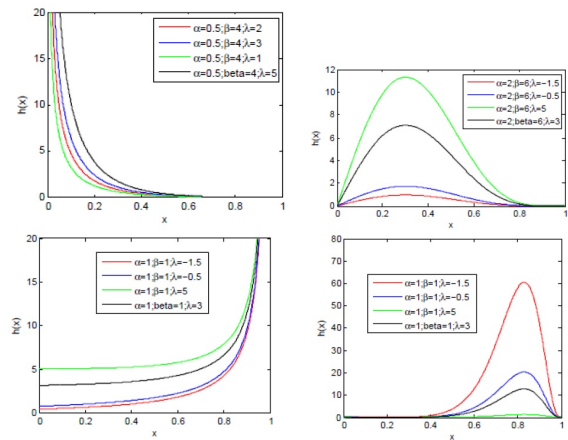
معادلات زیر به دست می‌آیند:

$$n \left(\frac{1}{\lambda} - \frac{1}{1 - e^{-\lambda}} \right) + \sum_{i=1}^n \log(1 - x_i^\alpha)^\beta = 0,$$

$$n + \sum_{i=1}^n \frac{\log x_i^\alpha}{1 - x_i^\alpha} \left(1 - \beta x_i^\alpha \left(1 + \lambda(1 - x_i^\alpha)^\beta \right) \right) = 0,$$

$$n + \sum_{i=1}^n \log(1 - x_i^\alpha) + \left(1 + \lambda(1 - x_i^\alpha)^\beta \right) = 0.$$

روابط بالا به آسانی می‌توانند با استفاده از نرم افزار R حل شوند [۹].



شکل ۴: تابع نرخ شکست توزیع کوماراسوآمی نمایی شده به ازای مقادیر مختلف α و β

۴ کاربردهایی از توزیع کوماراسوآمی و کوماراسوآمی نمایی شده

همان‌طور که قبلاً گفته شد توزیع کوماراسوآمی از ابتدا در مورد موضوعات آب و هوا مورد استفاده قرار گرفت. در این بخش دو کاربرد از توزیع آماری کوماراسوآمی و کوماراسوآمی نمایی شده را ارائه خواهیم کرد. در مثال اول نشان می‌دهیم که توزیع کوماراسوآمی توزیعی مناسبی برای داده‌های هواشناسی با توجه به نرخ شکست و نمودار TTT داده‌ها می‌باشد. برای نشان دادن مناسب بودن توزیع کوماراسوآمی برای برازش به داده‌های هواشناسی، داده‌های بارندگی روزانه ایستگاه فرودگاه رشت را در نظر گرفته و برتری این توزیع را نسبت به توزیع کوماراسوآمی نمایی شده برای این مجموعه از داده‌ها نشان می‌دهیم.

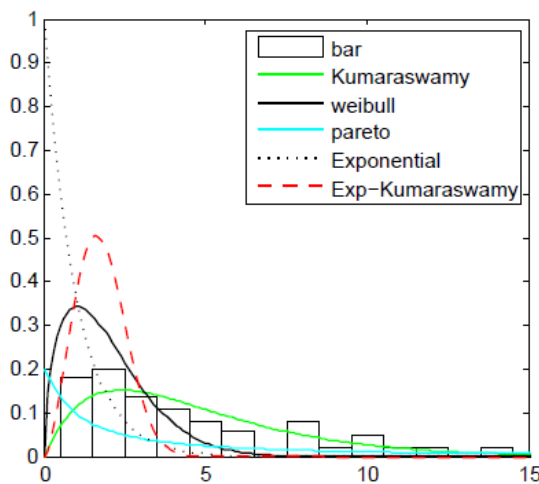
در بسیاری از کاربردها، اطلاعاتی کیفی درباره نمودارهای نرخ شکست وجود دارد که می‌تواند در انتخاب یک مدل مشخص ما را یاری کند. یک روش

ضریب چولگی و ضریب برجستگی X به ترتیب $\gamma_1[X]$ و $\gamma_2[X]$ به صورت روابط (۸) و (۹) به دست می‌آیند. لگاریتم درست‌نمایی ماکسیمم پارامترهای α ، β و λ به صورت زیر به دست می‌آید:

$$L(\lambda, \alpha, \beta) = n \log(\alpha) + n \log(\beta) + n \log \left(\frac{\lambda}{e^\lambda - 1} \right) + (\alpha - 1) \sum_{i=1}^n \log(x_i) + (\beta - 1) \sum_{i=1}^n \log(1 - x_i^\alpha) + \lambda \sum_{i=1}^n \log(1 - x_i^\alpha)^\beta. \quad (10)$$

در این صورت برآوردهای درست‌نمایی پارامترها از حل

در محیط نرم افزار متلب و همچنین رسم نمودار تابع چگالی احتمال چندین توزیع آماری، همانطور که در شکل مشخص است توزیع کوماراسوامی، نسبت به سایر توزیع‌ها، برازش بهتری را روی این داده‌ها نشان می‌دهد. در این بخش ۴۸ نمونه سنگ مخزن نفتی آورده شده است.



شکل ۶: نمودار مستطیلی و نمودار تابع چگالی احتمال توزیع‌های ارائه شده جهت یافتن بهترین توزیع برای برازش به داده‌ها

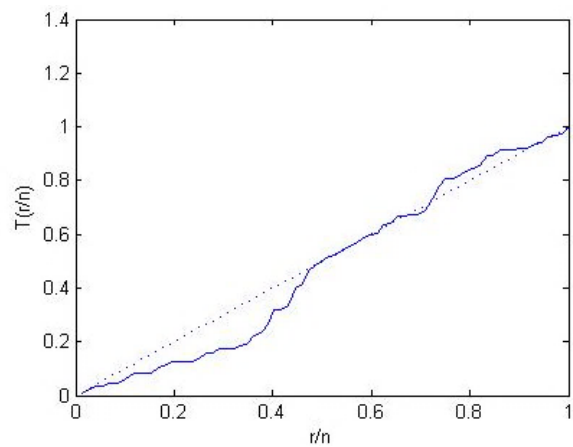
ابتدا نمودار TTT مربوط به این داده‌ها را رسم می‌کنیم. نمودار TTT مربوط به داده‌های زیر در شکل ۷ نمایش داده شده است.

0.0903296,	0.148622,	0.183312,
0.117063,	0.122417,	0.167045,
0.189651,	0.164127,	0.203654,
0.162394,	0.150944,	0.148141,
0.228595,	0.231623,	0.172567,
0.153481,	0.204314,	0.262727,
0.200071,	0.144810,	0.113852,
0.291029,	0.240077,	0.161865,

آزمون بسیار مفید، استفاده از نمودار TTT می‌باشد [۱۰]. نمودار TTT با استفاده از نمودار

$$T(r/n) = \left[\left(\sum_{i=1}^n y_{i:n} \right) + (n-r)y_{r:n} \right] / \sum_{i=1}^n y_{i:n},$$

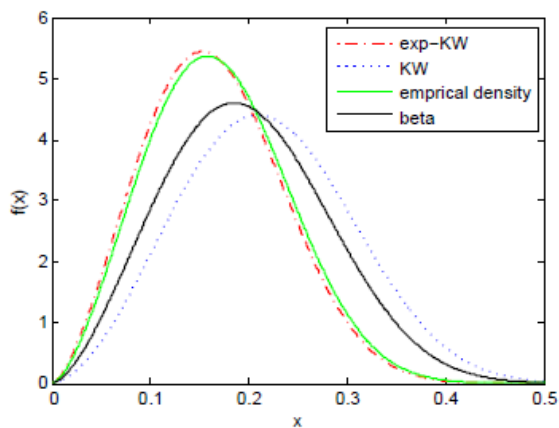
در برابر r/n به دست می‌آید که در آن $r = 1, \dots, n$ و $y_{i:n}$ به ازای $i = 1, \dots, n$ آماره‌های ترتیبی از نمونه می‌باشند. برای تابع نرخ شکست ثابت، نمودار قطری مورب، برای نرخ شکست کاهشی نمودار محدب و برای نرخ شکست افزایشی نمودار مقعر است. اگر تابع نرخ شکست u شکل باشد نمودار ابتدا مقعر و سپس محدب است. در شکل ۵ نمودار TTT مربوط به داده‌های بارندگی روزانه ایستگاه فرودگاه رشت رسم شده است.



شکل ۵: نمودار TTT مربوط به داده‌های بارندگی روزانه ایستگاه فرودگاه رشت

با توجه به نمودار TTT مربوط به داده‌های بارندگی روزانه ایستگاه فرودگاه رشت و با توجه به نرخ شکست توزیع کوماراسوامی، این توزیع، توزیع مناسبی برای این مجموعه از داده‌ها می‌باشد. با استفاده از رسم نمودار مستطیلی داده‌های بارندگی ایستگاه فرودگاه رشت

کوماراسوامی نمایی شده، برازش بهتری را به داده‌ها ارائه می‌دهد. در شکل زیر تابع چگالی احتمال تجربی و برآورد شده توزیع کوماراسوامی نمایی شده، تابع چگالی احتمال توزیع کوماراسوامی و توزیع بتا برای داده‌های ارائه شده، در محیط نرم افزار متلب رسم شده اند.



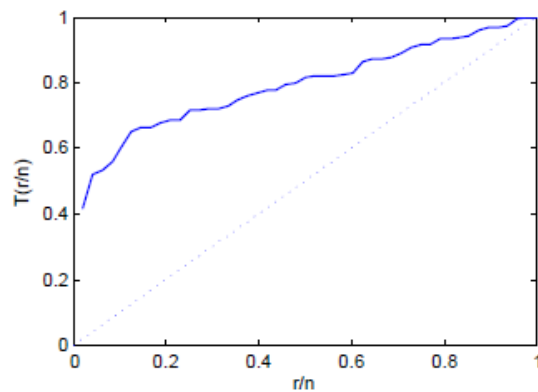
شکل ۸: مقایسه تابع چگالی احتمال چندین توزیع برای برازش بهترین توزیع به داده‌های نمونه سنگ مخزن نفتی

همان‌طور که از شکل ۸ مشخص است توزیع کوماراسوامی نمایی شده بهترین برازش را به این مجموعه از داده‌ها ارائه می‌دهد.

۵ نتایج و بحث

۱. در این مقاله توزیع کوماراسوامی و ویژگی‌های آن که در مدل‌بندی فرایندهای هیدرولوژیکی کاربرد دارد مورد بررسی قرار گرفت. با توجه به اینکه که تابع نرخ شکست توزیع کوماراسوامی و با توجه به نمودار TTT مربوط به داده‌های بارندگی روزانه ایستگاه فرودگاه رشت، این توزیع، توزیع مناسبی برای این مجموعه از داده‌ها می‌باشد.

0.280887, 0.179455, 0.191802,
0.133083, 0.225214, 0.341273,
0.311646, 0.276016, 0.197653,
0.326635, 0.154192, 0.276016,
0.176969, 0.438712, 0.163586,
0.253832, 0.328641, 0.230081,
0.464125, 0.420477, 0.200744,
0.262651, 0.182453, 0.200447.



شکل ۷: نمودار TTT مربوط به ۴۸ نمونه سنگ مخزن نفتی

همان‌طور که مشخص است با توجه به نرخ شکست توزیع کوماراسوامی نمایی شده و نمودار TTT داده‌ها، این توزیع برای این برازش به این مجموعه از داده‌ها مناسب‌تر است. برای مشخص کردن بهترین توزیع برای این مجموعه از داده‌ها، آماره‌های آزمون مختلفی را به کار می‌بریم. نتایج نیز مناسب بودن توزیع کوماراسوامی نمایی شده را بر اساس p -مقدار و آماره آزمون آکائیک برای این مجموعه از داده‌ها نشان می‌دهد. در جدول ۱، برآورد درست‌نمایی پارامترها، \hat{l} و p -مقدار و معیار آکائیک ارائه شده‌اند.

با توجه به جدول فوق مشخص است که توزیع

جدول ۱: برآورد درست‌نمایی پارامترها، \hat{t} و p -مقدار و معیار آکائیک

توزیع	برآورد پارامترها	\hat{t}	p -مقدار	آماره آزمون آکائیک
کوماراسوامی نمایی شده	$\hat{\alpha} = 0.67, \hat{\beta} = 9.15, \hat{\lambda} = -32.71$	۵۷/۸۲	۰/۶۳	-۱۰۹/۶۴
کوماراسوامی	$\hat{\alpha} = 2.71, \hat{\beta} = 44.04$	۵۲/۴۹	۰/۲۱	-۱۰۰/۹۸
بتا	$\hat{\alpha} = 5.94, \hat{\beta} = 21.20$	۵۵/۶۰	۰/۲۸	-۱۰۷/۲۰

۲. با استفاده از یک تبدیل مناسب می‌توان توزیع کوماراسوامی را به صورت توزیع کوماراسوامی نمایی شده نوشت. این توزیع بسیار انعطاف‌پذیر بوده و جایگزین مناسبی برای توزیع‌های کوماراسوامی و بتا می‌باشد.

۳. با توجه به داده‌های مربوط به ۴۸ نمونه سنگ مخزن نفتی، نشان داده شد که توزیع کوماراسوامی نمایی شده نسبت به توزیع کوماراسوامی انعطاف‌پذیرتر بوده و برازش بهتری را برای این مجموعه از داده‌ها فراهم می‌کند.

[3] Koutsoyiannis, D. and Xanthopoulos, T. (1989). On the parametric approach to unit hydrograph identification. *Water Resources Management*, 3, 107-128.

[4] Barreto-Souza W. and Simas A.B. (2013). The exp-G family of probability distributions. *Braz. Journal Probability Statics.*, 27, 84-109.

[5] McDonald, J.B. (1948). Some generalized functions for the size distribution of income. *Econometrica*, 52, 647-664.

مراجع

- [1] برهانی، ط.، فیاض، الف.، میرمصطفایی، م. و محمودی، ز.، (۱۳۹۴)، برآورد پارامترهای توزیع کوماراسوامی از طریق حل معادلات غیرخطی و شبیه‌سازی. هشتمین کنفرانس بین‌المللی انجمن ایرانی تحقیق در عملیات، دانشگاه فردوسی مشهد.
- [2] Kumarasawmy, P. (1980). A generalized probability density function for double bounded random process. *Journal of Hydrology*, 46, 79-88.
- [6] Hosking, J.R.M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society Series B*, 52, 105-124.
- [7] van Zwet, W.R. (1964). Convex Transformations of Random Variables, *Mathematisch Centrum. Amsterdam*.

- [8] Javanshiri, Z., Habibi Rad, A. and Arghami, N.R. (2015). Exp-Kumaraswamy Distributions: Some Properties and Applications. Journal of Sciences, Islamic Republic of Iran, 26(1), 57-69.
- [9] R Core Team, R. (2013). A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.
- [10] Aarset, M.V. (1987). How to identify bathtub hazard rate. IEEE Transactions on Reliability, 36, 106-108.