

ارائه مدل جدیدی از توزیع مارشال الکین نمایی تعمیم‌یافته و کاربردهای آن

آنی‌تا عبداله‌ی نانوایشه[†] و عاطفه عبداله‌ی نانوایشه[‡]

[†] دانشگاه رازی

[‡] دانشگاه آزاد اسلامی، واحد لاهیجان

چکیده. در این مقاله، ویژگی‌ها و کاربردهایی از یک توزیع سه پارامتری جدید به نام توزیع مارشال الکین نمایی تعمیم‌یافته نمایی شده ارائه شده است که کاربردهای بسیاری در هواشناسی، مسائل طول عمر، مدل‌سازی قیمت محصولات کشاورزی و قابلیت اطمینان دارد. این توزیع دارای تابع نرخ خطر افزایشی و کاهشی با توجه به مقادیر پارامترهایش است. در این مقاله ابتدا چند ویژگی‌های آماری و ریاضی این توزیع از جمله گشتاورها، برآورد ماکسیمم درست‌نمایی پارامترها بیان و یک مطالعه شبیه‌سازی در این زمینه ارائه خواهد شد. در نهایت کاربردهایی از توزیع جدید برای نشان دادن انعطاف‌پذیری توزیع، ارائه شده است. نتایج نیز مناسب بودن این مدل را برای مجموعه داده‌های واقعی در نظر گرفته‌شده، تایید می‌کنند.

واژه‌گان کلیدی: توزیع نمایی نمایی شده، توزیع مارشال الکین نمایی تعمیم‌یافته، توزیع مارشال الکین نمایی تعمیم‌یافته نمایی شده، برآورد درست‌نمایی ماکسیمم.

۱- مقدمه

توزیع نمایی تعمیم‌یافته (GE)^۱ یا توزیع نمایی نمایی شده برای اولین بار توسط ماده‌ولکار و سریواستاوا به عنوان جای‌گزینی برای توزیع‌های گاما و وایبول معرفی شد [۱۵]. بعد از

* نویسنده‌ی عهده‌دار مکاتبات
دریافت: ۱۴۰۱/۹/۱۶، پذیرش: ۱۴۰۳/۱۰/۸.

آن توزیع (GE) توسط بسیاری از محققان مانند گوپتا و کاندو [۹، ۶]، راغب و احسان‌الله [۱۶]، جائین [۱۳]، راغب و مادی [۱۷]، سران و ژنگ [۱۸] مورد بررسی قرار گرفت. گوپتا و کاندو [۷، ۱۱] نشان دادند که توزیع (GE) می‌تواند به طور گسترده در تجزیه و تحلیل بسیاری از داده‌های طول عمر، به‌خصوص به‌جای توزیع گامای دو پارامتری و توزیع وایبول دو پارامتری مورد استفاده قرار گیرد. گوپتا و کاندو [۸] همچنین روش‌های مختلفی از برآورد پارامترها را که شامل برآورد ماکسیمم درست‌نمایی، برآورد روش گشتاوری و برآورد با روش نمونه‌برداری احتمالی بر اساس یک نمونه تصادفی کامل بود، را مورد مطالعه قرار دادند. بررسی گسترده‌ای از برخی از پیشرفت‌های اخیر برای توزیع (GE) بر اساس یک نمونه تصادفی کامل توسط گوپتا و کاندو [۱۲] ارائه شده است. این توزیع دارای تابع توزیع تجمعی زیر (CDF) است:

$$(۱) \quad F(x; \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha, \quad x > 0, \quad \alpha, \lambda > 0$$

و تابع چگالی متناظر با آن نیز به صورت زیر می‌باشد:

$$(۲) \quad f(x; \alpha, \lambda) = \alpha \lambda (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}, \quad x > 0$$

اخیرا عبدالحی و همکاران توزیع نمایی واحد را ارائه و ویژگی‌های ریاضی و آماری آن را به همراه طرح بهینه مدل را به طور کامل بیان کردند [۱، ۲، ۳]. در ادامه مروری بر توزیع مارشال الکین نمایی تعمیم یافته و توزیع مارشال الکین نمایی تعمیم یافته نمایی شده و همچنین برخی از ویژگی‌های توزیع‌ها ارائه شده است. در بخش ۴ برآورد پارامترهای توزیع مورد بحث قرار گرفته است و در نهایت در بخش ۵ و ۶ یک مطالعه شبیه‌سازی کوچک و کاربردی از توزیع مارشال الکین نمایی تعمیم یافته نمایی شده برای نشان دادن این‌که توزیع پیشنهادی برازش بهتری را نسبت به سایر توزیع‌ها بر مجموعه داده‌های واقعی دارد ارائه شده است.

۲- توزیع مارشال الکین نمایی تعمیم‌یافته (MOGE) و توزیع مارشال الکین نمایی تعمیم‌یافته شده (EMOGE)

توزیع مارشال الکین نمایی تعمیم‌یافته (MOGE) برای اولین بار توسط مارشال و الکین [۱۴] پیشنهاد شد و به طور گسترده توسط آلیس و جوز [۴] مورد بحث قرار گرفت. تابع چگالی احتمال این توزیع به صورت زیر است:

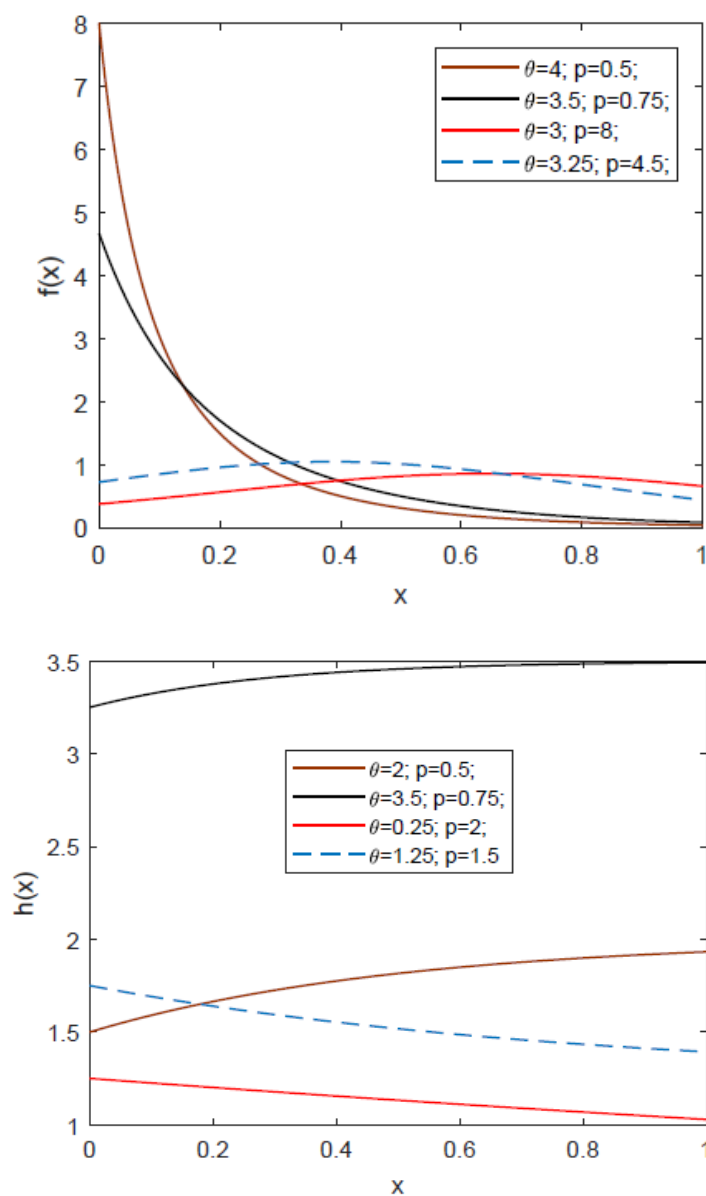
$$(۳) \quad f(x) = \frac{p\theta e^{-\theta x}}{(1 - (1-p)e^{-\theta x})^2}, \quad x > 0, \quad \theta > 0, \\ 0 < p \leq 1$$

تابع توزیع تجمعی، نرخ خط این توزیع نیز به ترتیب به صورت زیر است:

$$(۴) \quad F(x) = \frac{1 - e^{-\theta x}}{1 - (1-p)e^{-\theta x}}$$

$$(۵) \quad h(x) = \frac{\theta}{1 - (1-p)e^{-\theta x}}$$

اگر $p = 1$ باشد در این صورت توزیع MOGE به توزیع نمایی با پارامتر θ تبدیل می‌شود. شکل ۱ نمودارهای تابع چگالی احتمال و تابع نرخ خطر توزیع MOGE را به ازای مقادیر مختلف پارامترهایش نشان می‌دهد.



شکل ۱- تابع چگالی احتمال و تابع نرخ خطر توزیع MOGE به ازای مقادیر مختلف پارامترها

دوگشتاور اول، واریانس و تابع مولد گشتاور توزیع MOGE به صورت زیر است:

$$E(X) = \frac{p \log(p)}{\theta(p-1)}$$

$$E(X^r) = \frac{r p \text{Poly log}(r, 1-p)}{\theta^r (1-p)}$$

$$\text{var}(X) = \frac{-p(p \log(p))^r + r(p-1) \text{Poly log}(r, 1-p)}{\theta^r (p-1)^r}$$

$$M(t) = 1 + \frac{(1-p)^{-1+\frac{t}{\theta}} p t \text{Beta}(1-p, 1-\frac{t}{\theta}, 0)}{\theta}$$

که

$$\text{Poly log}(r, 1-p) = \sum_{k=1}^{\infty} \frac{(1-p)^k}{k^r}$$

و

$$\text{Beta}(1-p, 1-\frac{t}{\theta}, 0) = \int_0^{1-p} x^{(1-\frac{t}{\theta})-1} (1-x)^{-1} dx$$

حال در این قسمت برآورد پارامترهای توزیع را به روش حداکثر درستنمایی ارائه می‌کنیم. روش حداکثر درستنمایی یکی از مهم‌ترین روش‌ها در برآورد پارامترهای یک توزیع در آمار استنباطی است [۵]. به منظور برآورد پارامترهای توزیع، p و θ ؛ تابع لگاریتم درستنمایی بر اساس نمونه تصادفی $X_1, \dots, X_n$ به صورت زیر است:

$$l(p, \theta) = n \log p + n \log \theta - \theta \sum_{i=1}^n x_i - r \sum_{i=1}^n \log(1 - (1-p)e^{-\theta x_i})$$

(۶)

معادلات درستنمایی به صورت زیر هستند:

$$\frac{\partial l}{\partial p} = \frac{n}{p} - \sum_{i=1}^n \frac{e^{-\theta x_i}}{1 - (1-p)e^{-\theta x_i}} = 0,$$

$$\frac{\partial l}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^n x_i - \sum_{i=1}^n \frac{(1-p)x_i e^{-\theta x_i}}{1 - (1-p)e^{-\theta x_i}} = 0.$$

برآورد ماکسیم درستنمایی پارامترها با استفاده از حل تابع غیرخطی به دست می‌آید که دارای فرم بسته‌ای نیستند و با استفاده از روش‌های عددی به دست می‌آیند. حال فرض کنید X دارای تابع توزیع تجمعی رابطه (۴) باشد در این صورت تابع توزیع تجمعی توانی شده آن؛ یعنی توزیع مارشال الکین نمایی تعمیم‌یافته نمایی شده، (EMOGE) به صورت زیر است:

$$(7) \quad F(x) = \left[\frac{1-e^{-\theta x}}{1-(1-p)e^{-\theta x}} \right]^\alpha, \quad X > 0, \theta > 0, \alpha > 0, 0 < p \leq 1$$

تابع چگالی احتمال این توزیع، با مشتق‌گیری از رابطه بالا به صورت زیر به دست می‌آید:

$$f(x) = \alpha \left[\frac{(\theta e^{-\theta x})(1-(1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1-e^{-\theta x})}{(1-(1-p)e^{-\theta x})^2} \right] \left[\frac{1-e^{-\theta x}}{1-(1-p)e^{-\theta x}} \right]^{\alpha-1}$$

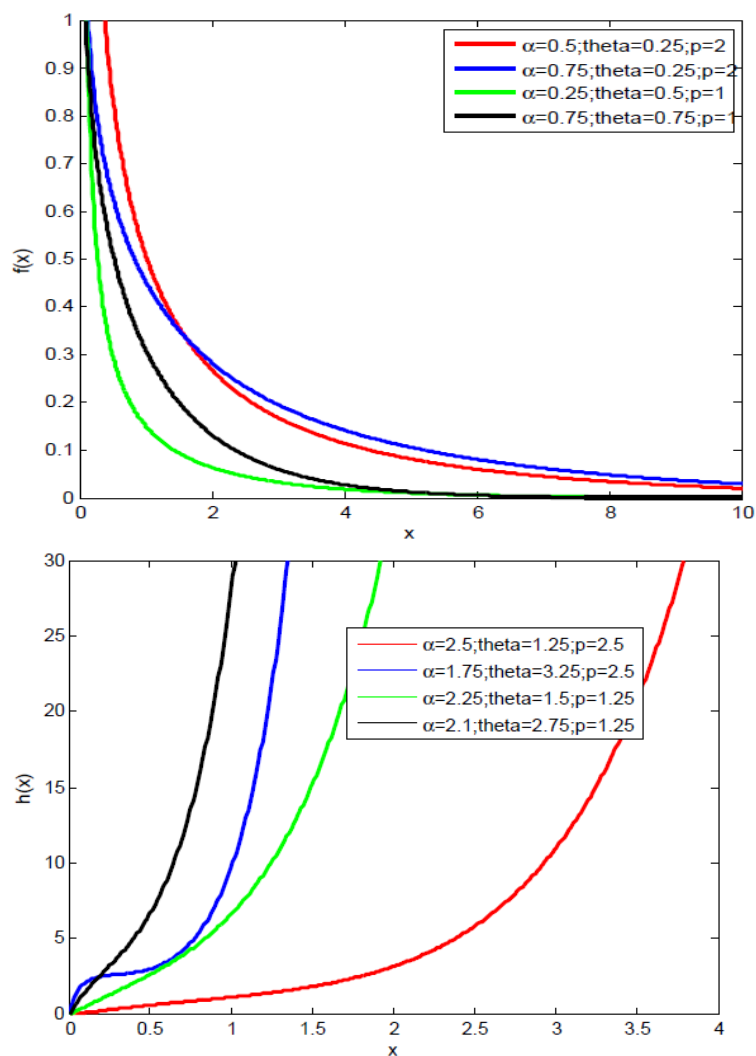
(۸)

تابع نرخ خطر این توزیع نیز به صورت زیر است:

$$h(x) = \frac{\alpha \left[\frac{(\theta e^{-\theta x})(1-(1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1-e^{-\theta x})}{(1-(1-p)e^{-\theta x})^2} \right] \left[\frac{1-e^{-\theta x}}{1-(1-p)e^{-\theta x}} \right]^{\alpha-1}}{1 - \left[\frac{1-e^{-\theta x}}{1-(1-p)e^{-\theta x}} \right]^\alpha}$$

(۹)

در شکل‌های زیر نمودار تابع چگالی احتمال و تابع نرخ خطر توزیع جدید به ازای مقادیر مختلف پارامترها رسم شده است. همان طور که مشخص است تابع نرخ خطر این توزیع به ازای مقادیر مختلف پارامترها به صورت افزایشی می‌باشد.



شکل ۲- تابع چگالی احتمال توزیع و تابع نرخ خطر توزیع (EMOGGE) به ازای مقادیر مختلف پارامترها

۳- ویژگی‌های آماری توزیع (EMOGE)

توزیع حال در این قسمت برخی از خواص آماری توزیع (EMOGE) مانند گشتاورها و تابع مولد گشتاور مدل معرفی شده را ارائه خواهیم کرد. قبل از ارائه کردن گشتاورهای مدل معرفی شده، به ارائه بسط $f(x)$ که در محاسبه گشتاورها مورد نیاز است می‌پردازیم. اگر:

$$f(x) = \alpha \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right] \\ \times \left[1 - \frac{pe^{-\theta x}}{1 - (1-p)e^{-\theta x}} \right]^{\alpha-1}$$

در این صورت با استفاده از بسط زیر:

$$(1-u)^a = \sum_{j=0}^{\infty} \binom{a}{j} (-1)^j (u)^j, \quad |u| < 1$$

$f(x)$ را بازنویسی می‌کنیم:

$$f(x) = \alpha \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right] \times \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{(-1)^j (pe^{-\theta x})^j}{(1 - (1-p)e^{-\theta x})^j} \quad (10)$$

و به طور مشابه بسط تابع توزیع آن نیز به صورت زیر است:

$$F(X) = \left[\frac{1 - e^{-\theta x}}{1 - (1-p)e^{-\theta x}} \right]^\alpha = \left[1 - \frac{pe^{-\theta x}}{1 - (1-p)e^{-\theta x}} \right]^\alpha \\ F(X) = \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{(-1)^j (pe^{-\theta x})^j}{(1 - (1-p)e^{-\theta x})^j} \quad (11)$$

در این صورت گشتاور مرتبه r ام توزیع (EMOGE) از حل انتگرال زیر محاسبه می شود:

$$E[x^r] = \int_0^{\infty} x^r \alpha \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right] \times \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{(-1)^j (p e^{-\theta x})^j}{(1 - (1-p)e^{-\theta x})^j} dx \quad (12)$$

بنابراین گشتاور اول توزیع جدید از حل انتگرال زیر به دست می آید. به همین ترتیب سایر گشتاورها نیز با استفاده از رابطه (۱۲) و با دادن مقدار به r به دست می آیند.

$$E(X) = \int_0^{\infty} x \alpha \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right] \times \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{(-1)^j (p e^{-\theta x})^j}{(1 - (1-p)e^{-\theta x})^j} dx$$

اگر X دارای توزیع رابطه (۹) باشد در این صورت تابع مولد گشتاور این توزیع نیز از حل انتگرال زیر به دست می آید:

$$M(x) = \int_0^{\infty} e^{tx} \alpha \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right] \times \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{(-1)^j (p e^{-\theta x})^j}{(1 - (1-p)e^{-\theta x})^j} dx \quad (13)$$

۴- تابع درستنمایی ماکزیم (MLE)

اگر $X_1, \dots, X_n$ نمونه ای تصادفی با اندازه n از توزیع (EMOGE) باشند در این صورت تابع ماکسیم درستنمایی به صورت زیر خواهد بود:

$$L = \prod_{i=1}^n \alpha \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right]$$

$$\prod_{i=1}^n \left[\frac{1 - e^{-\theta x}}{1 - (1-p)e^{-\theta x}} \right]^{\alpha-1}$$

در نتیجه لگاریتم تابع درستنمایی نیز به صورت زیر است:

$$l(\alpha, \theta, p) = n \log \alpha$$

$$+ \sum_{i=1}^n \log \left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right]$$

$$+ (\alpha - 1) \sum_{i=1}^n \log \left[\frac{1 - e^{-\theta x}}{1 - (1-p)e^{-\theta x}} \right]$$

(۱۴)

با مشتق‌گیری از رابطه (۱۴) نسبت به پارامترها داریم:

$$\frac{\partial l_n}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \log \left[\frac{1 - e^{-\theta x}}{1 - (1-p)e^{-\theta x}} \right],$$

$$\frac{\partial l_n}{\partial p} = \frac{\frac{((e^{-\theta x}) + (\theta e^{-\theta x})(1 - e^{-\theta x}))((1 - (1-p)e^{-\theta x})^2)}{(1 - (1-p)e^{-\theta x})^4}}{\left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right]}$$

$$- \frac{(\frac{(\frac{1}{2}e^{-\theta x}(1 - (1-p)e^{-\theta x}))((\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})))}{(1 - (1-p)e^{-\theta x})^4}}{\left[\frac{(\theta e^{-\theta x})(1 - (1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1 - e^{-\theta x})}{(1 - (1-p)e^{-\theta x})^2} \right]}$$

$$\begin{aligned}
& - \frac{\frac{(e^{-\theta x})(1-e^{-\theta x})}{(1-(1-p)e^{-\theta x})^2}}{\left[\frac{1-e^{-\theta x}}{1-(1-p)e^{-\theta x}} \right]}, \\
\partial l_n / \partial \theta = & \frac{\frac{(\theta x e^{-\theta x}(1-(1-p)e^{-\theta x}) + \theta(e^{-\theta x})^2(-e^{-\theta x} + x\theta e^{-\theta x} + p e^{-\theta x} - p\theta x e^{-\theta x}))(1-(1-p)e^{-\theta x})^2}{(1-(1-p)e^{-\theta x})^4}}{\left[\frac{(\theta e^{-\theta x})(1-(1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1-e^{-\theta x})}{(1-(1-p)e^{-\theta x})^2} \right]} \\
& - \frac{\frac{((\theta x e^{-\theta x} - p x e^{-\theta x})(1-(1-p)e^{-\theta x}))(\theta e^{-\theta x})(1-(1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1-e^{-\theta x}))}{(1-(1-p)e^{-\theta x})^4}}{\left[\frac{(\theta e^{-\theta x})(1-(1-p)e^{-\theta x}) - ((1-p)\theta e^{-\theta x})(1-e^{-\theta x})}{(1-(1-p)e^{-\theta x})^2} \right]} \\
& - \frac{\frac{(x e^{-\theta x})(1-(1-p)e^{-\theta x}) - (x e^{-\theta x} - p x e^{-\theta x})(1-e^{-\theta x})}{(1-(1-p)e^{-\theta x})^2}}{\left[\frac{1-e^{-\theta x}}{1-(1-p)e^{-\theta x}} \right]}
\end{aligned}$$

برآورد ماکسیمم درستنمایی پارامترها با برابر صفر قرار دادن عبارات بالا و با استفاده از حل تابع غیرخطی به دست می‌آیند که دارای فرم بسته‌ای نیستند. برای ارائه برآورد بازه‌ای، فاصله اطمینان و آزمون فرض روی پارامترها، نیاز به ماتریس اطلاع داریم. ماتریس اطلاع مشاهده شده به صورت زیر است:

$$J(\vartheta) = - \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha p} & I_{\alpha\theta} \\ I_{\alpha p} & I_{pp} & I_{p\theta} \\ I_{\alpha\theta} & I_{p\theta} & I_{\theta\theta} \end{bmatrix}$$

که $\vartheta' = (\alpha, \theta, p)$. عبارات واضحی برای عناصر ماتریس اطلاع مشاهده شده به صورت زیر وجود دارد:

$$I_{\alpha\alpha} = \frac{\partial^2 l}{\partial \alpha^2}, \quad I_{pp} = \frac{\partial^2 l}{\partial p^2}, \quad I_{\theta\theta} = \frac{\partial^2 l}{\partial \theta^2}, \quad I_{\alpha p} = \frac{\partial^2 l}{\partial \alpha \partial p}, \quad I_{\alpha\theta} = \frac{\partial^2 l}{\partial \alpha \partial \theta},$$

$$I_{p\theta} = \frac{\partial^2 l}{\partial p \partial \theta}$$

تحت شرایطی که برای پارامترها در خارج از فضای پارامتری انجام می‌شود نه روی مرزها، توزیع مجانبی $\sqrt{n}(\hat{\Theta} - \Theta)$ همگرا به توزیع $(\vartheta, I^{-1}(\vartheta))$ است که $I(\vartheta)$ ماتریس اطلاع مورد انتظار است. توزیع نرمال چندمتغیره $(0, I^{-1}(\vartheta))$ N_3 می‌تواند برای ایجاد بازه های اطمینان دو طرفه $(1 - \eta)100\%$ درصدی برای پارامترهای مدل به کار رود که η سطح معناداری است.

۵- یک مطالعه شبیه سازی

در این بخش یک مطالعه شبیه سازی کوچک را به منظور ارزیابی عملکرد برآورد ماکسیمم درست‌نمایی پارامترها برای توزیع معرفی شده ارائه می‌دهیم. تعداد تکرارها برابر 1000 است. ارزیابی برآوردها بر مبنای موارد زیر انجام می‌شود: میانگین توان دوم خطاهای تجربی و برآورد ماکسیمم درست‌نمایی به ازای هر داده شبیه سازی شده برای $i=1, \dots, 1000$ به صورت $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\theta})$ تعیین می‌شوند. مقدار ارببی و میانگین مربعات خطا از فرمول‌های زیر به دست می‌آیند:

$$bias_h(n) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{h}_i - h);$$

$$MSE_h(n) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{h}_i - h)^2;$$

نتایج در جدول زیر ارائه شده‌اند که نشان دهنده این است که برآوردهای ماکسیمم درست‌نمایی پارامتر θ دارای ارببی مثبت هستند. نتایج کاملاً پایدار هستند و مهم‌تر این که برای این اندازه از نمونه، به مقادیر واقعی نزدیک هستند. علاوه بر این با افزایش اندازه نمونه مقدار MSE کاهش می‌یابد.

جدول ۱- اریبی و میانگین مربعات خطای برآوردها تحت روش ماکسیمم درست‌نمایی با در نظر گرفتن $\alpha, p = 1$

Maximum Likelihood Estimation			
θ	n	Bias	MSE
۰/۵	۵۰	۰/۰۵۶۰	۰/۰۰۵۱
	۱۰۰	۰/۰۴۰۴	۰/۰۰۲۶
	۱۵۰	۰/۰۳۱۹	۰/۰۰۱۵
	۲۰۰	۰/۰۲۹۱	۰/۰۰۱۳
۲	۵۰	۰/۲۳۲۲	۰/۰۸۸۰
	۱۰۰	۰/۱۶۶۹	۰/۰۴۳۹
	۱۵۰	۰/۱۳۵۱	۰/۰۲۹۱
	۲۰۰	۰/۱۰۹۱	۰/۰۱۸۶
۵	۵۰	۰/۵۷۴۰	۰/۵۲۵۹
	۱۰۰	۰/۴۰۴۴	۰/۲۶۸۰
	۱۵۰	۰/۳۴۱۵	۰/۱۹۳۲
	۲۰۰	۰/۲۹۸۱	۰/۱۳۸۱

۶- کاربردهایی از توزیع جدید

حال در این بخش کاربردهایی از توزیع ارائه شده را بیان می‌کنیم و نشان می‌دهیم که توزیع جدید ارائه شده برآزش بهتری را نسبت به سایر مدل‌ها برای مجموعه داده‌های واقعی ارائه شده در این مقاله فراهم می‌کند. برای این منظور توزیع ذکر شده را با توزیع‌های رایلی واحد و وایبل واحد و بور واحد مقایسه می‌کنیم. همه محاسبات با استفاده از نرم افزارهای R و Matlab انجام شده‌اند.

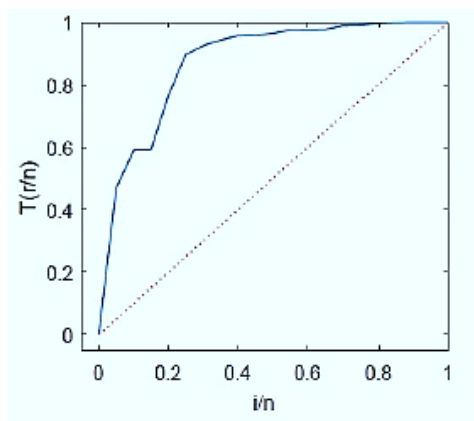
مجموعه داده‌های ارائه شده مربوط به ظرفیت آب مخزن شستا برای ماه فوریه از سال ۱۹۹۱ تا ۲۰۱۰ است که در جدول زیر نشان داده شده است.

جدول ۲- ظرفیت آب مخزن شستا برای ماه فوریه از سال ۱۹۹۱ تا ۲۰۱۰

Year	Capacity	Proportion of total capacity	Year	Capacity	Proportion of total capacity
۱۹۹۱	۱۵۴۲۸۳۸	۰/۳۳۸۹۳۶	۲۰۰۱	۳۴۹۵۹۶۹	۰/۷۶۸۰۰۷
۱۹۹۲	۱۹۶۶۰۷۷	۰/۴۳۱۹۱۵	۲۰۰۲	۳۸۳۹۵۴۴	۰/۸۴۳۴۸۵
۱۹۹۳	۳۴۵۹۲۰۹	۰/۷۵۹۹۳۲	۲۰۰۳	۳۵۸۴۲۸۳	۰/۷۸۷۴۰۸
۱۹۹۴	۳۲۹۸۴۹۶	۰/۷۲۴۶۲۶	۲۰۰۴	۳۸۶۸۶۰۰	۰/۸۴۹۸۶۸
۱۹۹۵	۳۴۴۸۵۱۹	۰/۷۵۷۵۸۳	۲۰۰۵	۳۱۶۸۰۵۶	۰/۶۹۵۹۷۰
۱۹۹۶	۳۶۹۴۲۰۱	۰/۸۱۱۵۵۶	۲۰۰۶	۳۸۳۴۲۲۴	۰/۸۴۲۳۱۶
۱۹۹۷	۳۵۷۴۸۶۱	۰/۷۸۵۳۳۹	۲۰۰۷	۳۷۷۲۱۹۳	۰/۸۲۸۶۸۹
۱۹۹۸	۳۵۶۷۲۲۰	۰/۷۸۳۶۶۰	۲۰۰۸	۲۶۴۱۰۴۱	۰/۵۸۰۱۹۴
۱۹۹۹	۳۷۱۲۷۳۳	۰/۸۱۵۶۲۷	۲۰۰۹	۱۹۶۰۴۵۸	۰/۴۳۰۶۸۱
۲۰۰۰	۳۸۵۷۴۲۳	۰/۸۴۷۴۱۳	۲۰۱۰	۳۳۸۰۱۴۷	۰/۷۴۲۵۶۳

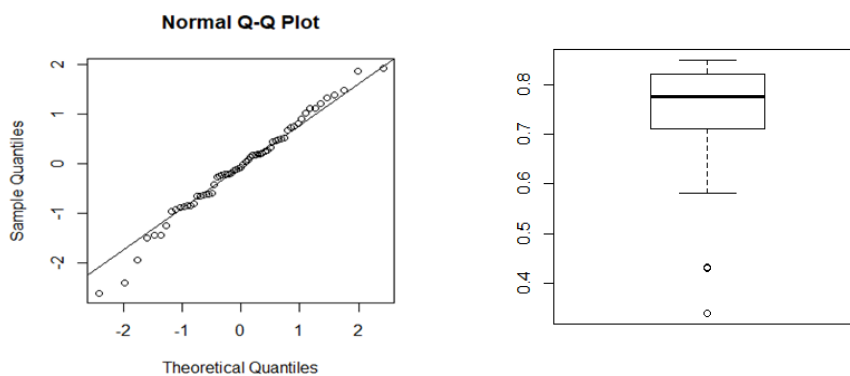
در بسیاری از کاربردها، اطلاعاتی کیفی درباره نمودارهای نرخ شکست وجود دارد که می‌توانند در انتخاب یک مدل مشخص ما را یاری کنند. یک روش آزمون بسیار مفید، استفاده از نمودار (TTT) می‌باشد. نمودار TTT با استفاده از نمودار، $T(r/n) = [\sum_{i=1}^r y_{(i)} + (n-r)y_{(r)}] / \sum_{i=1}^n y_{(i)}$ در برابر r/n به دست می‌آید که $r = 1, \dots, n$ و $y_{(i)}; (i = 1, \dots, n)$ آماره‌های ترتیبی نمونه می‌باشند. برای تابع نرخ خطر ثابت، نمودار قطری مورب، برای نرخ خطر کاهشی نمودار محدب و برای نرخ خطر افزایشی نمودار مقعر است. اگر تابع نرخ خطر u شکل باشد نمودار ابتدا محدب و سپس مقعر است و اگر تابع نرخ خطر u شکل وارون باشد نمودار ابتدا مقعر و سپس محدب است.

حال نمودار TTT داده‌ها را رسم می‌کنیم. شکل ۳ نمودار TTT مربوط به داده‌ها را نشان می‌دهد که با استفاده از آن می‌توان به ویژگی نرخ خطر داده‌ها پی برد. این نمودار نشان می‌دهد که تابع نرخ خطر مجموعه داده‌ها به صورت افزایشی است، بنابراین توزیع لگ لجستیک وایبل نمایی شده برای برازش این داده‌های واقعی مناسب است، زیرا این توزیع تابع نرخ خطر فزاینده‌ای را ارائه می‌دهد.



شکل ۳- نمودار TTT مربوط به داده‌های مخزن شستا

شکل ۴ نمودار احتمال نرمال و نمودار جعبه‌ای مجموعه داده‌ها را ارائه می‌دهد.

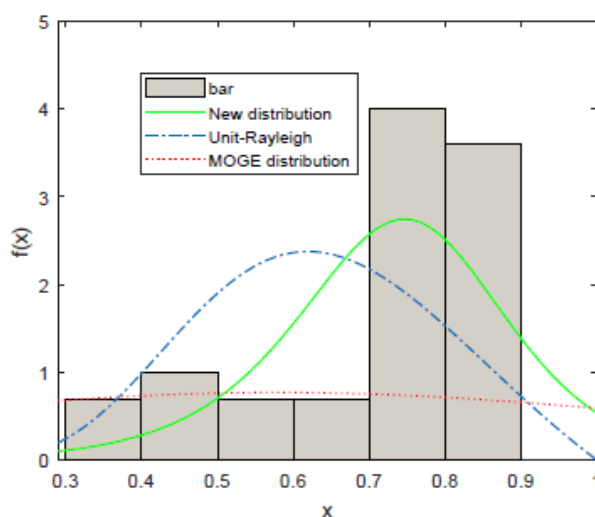


شکل ۴- احتمال نرمال و نمودار جعبه‌ای برای مجموعه داده‌های مخزن شستا

جهت بررسی نیکوئی برازش توزیع‌های آماری ارائه شده با توزیع داده‌های واقعی، از آماره‌های BIC و AIC استفاده گردید. در جدول ۱، MLE داده‌ها، مقدار BIC و AIC برای مجموعه داده‌های واقعی، ارائه شده‌اند. همان‌طور که مشخص است با توجه به معیارهای BIC, AIC (بر اساس کمترین مقدار) توزیع جدید ارائه شده بهترین برازش را برای این مجموعه از داده‌ها فراهم می‌کند.

Model	Parameter estimates			$-2 \ln L$	AIC	BIC
	α	θ	p			
New distribution	۰٫۸۱۹۷	۱۱٫۷۲۲۸	۷۶۸۸٫۵۷	-۲۱٫۳۱۷۳	-۱۵٫۳۱۷۳	-۱۲٫۳۳۰۱
Unit-Rayleigh	—	—	۵٫۲۱۵۲۷۲	-۱۷٫۲۳۰۵۱	-۱۵٫۲۳۰۵	-۱۲٫۲۳۴۷
MOGE distribution	۰٫۲		۵٫۲۵۹۹	-۴٫۴۹۴۸	-۰٫۴۹۴۸	۱٫۴۹۶۶۶

در شکل زیر نمودار تابع چگالی احتمال توزیع‌های ارائه شده و هیستوگرام داده‌های واقعی، جهت انتخاب بهترین توزیع برای برازش به داده‌ها رسم شده است. همانطور که از شکل‌ها مشخص است توزیع ذکر شده، نسبت به سایر مدل‌های رقیب، برازش بهتری را روی این مجموعه از داده‌ها فراهم می‌کند.



شکل ۵- نمودار تابع چگالی احتمال توزیع‌های در نظر گرفته شده به همراه هیستوگرام مجموعه داده‌های مخزن شستا.

۷- نتیجه‌گیری و چند پیشنهاد برای کار

در این مقاله ویژگی‌ها و کاربردهایی از توزیع مارشال ال‌کین نمای تعمیم یافته نمایشی شده بیان و ارائه شد. این مدل نسبت به سایر مدل‌ها دارای فرم‌های ساده تابع توزیع و چگالی به لحاظ ریاضی بوده و دارای ۳ پارامتر است. تابع نرخ خطر این توزیع به صورت

افزایشی و کاهشی با توجه به مقادیر مختلف پارامترهایش است. به عنوان نتیجه‌ی کلی، می‌توان گفت که جدید ارائه شده مکانیسم نسبتاً انعطاف‌پذیری برای برازش طیف وسیعی از مجموعه داده‌های واقعی مثبت فراهم می‌کند و بنابراین می‌تواند جایگزین مناسبی برای داده‌های ارایه شده در این مقاله، در الگوبندی داده‌های واقعی در زمینه‌های اقتصاد، مهندسی، تحلیل بقا و مسائل هیدرولوژیکی باشد. بیشتر محاسبات و شکل‌ها با استفاده از نرم افزار Matlab و R انجام شده است. چندین ویژگی از این توزیع در این مقاله ارائه نشده است که امیدواریم در تحقیقات آینده مورد بررسی قرار گیرند. برخی از این ویژگی‌ها عبارتند از: آماره‌های ترتیبی، گشتاورهای آماره ترتیبی، توزیع مجانبی آماره‌های اکسترم، ترتیب‌های تصادفی، آنتروپی و آنتروپی تجمعی باقیمانده، توزیع نسبت دو متغیر تصادفی. همچنین موضوع‌های استنباطی نیز می‌تواند زمینه‌ای برای تحقیق در آینده را فراهم آورد. برخی از این موضوع‌ها عبارتند از: برآوردیابی بیزی پارامترها، برآوردیابی بر اساس انواع سانسورها، برآوردیابی به کمک روش‌های دیگر مانند روش کمترین مربعات قطری، پیش‌بینی مشاهده‌های آینده از این توزیع و غیره.

توضیحات

1. generalized exponential
2. Marshall-Olkin generalized exponential

مرجع‌ها

- [1] Abdollahi, A., Jafari, H., and Khazaei, S. (2024). Locally, Bayesian and Nonparametric Bayesian optimal designs for Unit Exponential regression model. *Journal of Communications in Statistics Theory and Methods*. doi.org/10.1080/03610926.2024.2328182.
- [2] Abdollahi Nanvapisheh, A., MirMostafae, S.M.T.K., and Emrah, A. (2019). A new two-parameter distribution: properties and applications. *Journal of Mathematical Modeling*, **7.1**, 35–48.
- [3] Abdollahi, A. (2021). The comparison between Gumbel and exponentiated Gumbel distributions and their applications in

- hydrological process. *American Journal of Computer Science and Information Technology*.
- [4] Alice, T., and Jose, K.K. (1999). On Marshall- Olkin Generalized Exponential Distribution and its Applications. The Second Annual Conference of the Society of Statistics and Computer Applications held at St. Thomas College, Pala, Kerala during 28th November to 1st December 1999.
- [5] Casella, G., and Berger, R.L. (2002). *Statistical Inference*, Second edition, Duxbury Press.
- [6] Gupta, R.D., and Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, **41**, 173–188.
- [7] Gupta, R.D., and Kundu, D. (2001a). Exponentiated exponential distribution: An alternative to gamma and Weibull distributions. *Biometrical Journal*, **43**, 117–130.
- [8] Gupta, R.D., and Kundu, D. (2001b). Generalized exponential distributions: Different method of estimations. *Journal of Statistical Computation and Simulation*, **69**, 315–338.
- [9] Gupta, R.D., and Kundu, D. (2002). Generalized exponential distributions: Statistical inference. *Journal of Statistical Theory and Applications*, **1**, 101–118.
- [10] Gupta, R.D., and Kundu, D. (2003a). Closeness of gamma and generalized exponential distribution. *Communications in Statistics – Theory and Methods*, **32**, 705–721.
- [11] Gupta, R.D., and Kundu, D. (2003b). Discriminating between Weibull and generalized exponential distributions. *Computational Statistics and Data Analysis*, **43**, 179–196.
- [12] Gupta, R.D., and Kundu, D. (2007). Generalized exponential distribution: Existing results and some recent developments. *Journal of Statistical Planning and Inference*, **137**, 3537–3547.

- [13] Jaheen, Z.F. (2004). Empirical Bayes inference for generalized exponential distribution based on records. *Communications in Statistics-Theory and Methods*, **33**, 1851–1861.
- [14] 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**, 641–652.
- [15] Mudholkar, G.S., and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure data. *IEEE Transactions on Reliability*, **42**, 299–302.
- [16] Raqab, M.Z., and Ahsanullah, M. (2001). Estimation of the location and scale parameters of generalized exponential distribution based on order statistics. *Journal of Statistical Computation and Simulation*, **137**, 109–124.
- [17] Raqab, M.Z., and Madi, M.T. (2005). Bayesian inference for the generalized exponential distribution. *Journal of Statistical Computation and Simulation*, **75**, 841–852.
- [18] Sarhan, A.M. (2007). Analysis of incomplete, censored data in competing risks models with generalized exponential distributions. *IEEE Transactions on Reliability*, **56**(1), 132–138.

آنی‌تا عبداللهی نانواپیشه

دکتری آمار ریاضی

کرمانشاه، دانشگاه رازی، گروه آمار.

رایانشانی: anita.abdollahi@yahoo.com

عاطفه عبداللهی نانواپیشه

کارشناس ارشد ریاضی

گیلان، لاهیجان، دانشگاه آزاد اسلامی، واحد لاهیجان.

رایانشانی: abdollahi.atfe60@yahoo.com