

یک تعمیم جدید از توزیع گومپرتز

شهرام یعقوبزاده شهرستانی^۱

تاریخ دریافت: ۱۳۹۵/۹/۲

تاریخ پذیرش: ۱۳۹۶/۱/۳۰

چکیده:

در این مقاله، یک توزیع جدید طول عمر سه پارامتری بر اساس توزیع گومپرتز به نام مارشال-الکین گومپرتز که تعمیمی از توزیع گومپرتز و دارای نرخ شکست‌های نزولی، صعودی و وانی شکل است، معرفی شده و تابع چگالی احتمال، تابع توزیع تجمعی، تابع خطر و برخی از خصوصیات این توزیع جدید مانند گشتاورهای مرکزی، گشتاورهای آماره‌های مرتب، آنتروپی‌های رنی و شانون و تابع چندک به دست می‌آید. همچنین پارامترهای آن به روش ماکسیمم درست‌نمایی برآورد شده، به کمک یک مجموعه داده واقعی، این توزیع جدید با برخی از توزیع‌های طول عمر بر اساس توزیع گومپرتز مقایسه می‌شود.

واژه‌های کلیدی: توزیع گومپرتز، تابع شکست، برآورد ماکسیمم درست‌نمایی، توزیع مارشال-الکین گومپرتز.

۱ مقدمه

شود و نشان داده می‌شود که مدلی مناسب‌تر از بقیه توزیع‌های طول عمر معرفی شده است. در بخش دوم مقاله به معرفی توزیع جدید پرداخته و برخی از خصوصیات آن مانند تابع شکست، گشتاورهای مرکزی، تابع چگالی آماره‌های ترتیبی، آنتروپی‌های رنی و شانون و تابع چندک محاسبه و پارامترهای آن به روش ماکسیمم درست‌نمایی برآورد می‌شود. در بخش سوم با استفاده از یک مجموعه داده واقعی، توزیع جدید با برخی از توزیع‌های بر پایه توزیع گومپرتز مقایسه می‌شود.

توزیع گومپرتز^۲ [۲] نقش مهمی در مدل‌بندی زمان‌های بقا و مرگ و میر بشر دارد [۲ و ۵]. همچنین از کاربردهای دیگر توزیع گومپرتز می‌توان به کاربرد در علوم مختلفی مانند زیست‌شناسی [۳]، تحلیل بقا [۸]، جامعه‌شناسی [۲۳]، کامپیوتر [۱۶]، بازاریابی [۱] اشاره کرد. از آنجا که تابع خطر توزیع گومپرتز ثابت و افزایشی است، نمی‌تواند به عنوان یک توزیع طول عمر برای مدل‌بندی پدیده‌های با تابع خطر غیر ثابت و غیر افزایشی مؤثر باشد. به همین علت توزیع‌هایی بر پایه توزیع گومپرتز توسط نویسندگان مختلف به عنوان توزیع جدید طول عمر معرفی شد. از جمله این توزیع‌ها می‌توان به توزیع بتا-گومپرتز [۹]، توزیع گومپرتز تعمیم یافته [۴]، توزیع مک دونالد گومپرتز [۱۸] و گومپرتز-پوآسن [۲۱] که یکی از توزیع‌های متعلق به خانواده توزیع‌های سری توانی گومپرتز [۱۰] است، اشاره کرد. چون مدل‌بندی زمان‌های بقا و مرگ و میر بشر حائز اهمیت بوده و توزیع‌های کمی به عنوان توزیع طول عمر معرفی شدند، هدف اصلی این مقاله آن است که یک توزیع جدید سه پارامتری بر اساس توزیع گومپرتز با نرخ‌های شکست نزولی، صعودی و وانی شکل به عنوان توزیع طول عمر معرفی

۲ توزیع مارشال-الکین گومپرتز

تابع توزیع تجمعی و تابع چگالی احتمال توزیع گومپرتز که با نماد $Gom(\alpha, \beta)$ نشان داده می‌شود به ترتیب به صورت زیر است.

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

$$f(x) = \alpha e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x} - 1)}, \quad x > 0, \alpha > 0, \beta > 0 \quad (2)$$

قضیه ۱.۲. می‌باشد. اگر X متغیری تصادفی با تابع چگالی احتمال (۲) باشد، گشتاور مرکزی مرتبه r ام آن به صورت زیر

^۱ عضو هیئت علمی گروه آمار، دانشگاه پیام نور، ایران

^۲ Gompertz

^۳ generalized integro-exponential function

است.

و تابع خطر توزیع مارشال-الکین گومپرتز نیز به صورت

$$h(x) = \frac{\alpha p e^{\beta x}}{1 - (1-p)e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}}$$

می باشد که نمودار آن به ازای مقادیر مختلف پارامترهایش در شکل ۲ نشان داده شده است. شکل ۱ که نمودار توزیع مارشال-الکین گومپرتز به ازای مقادیر مختلف پارامترهایش است، بیان کننده انعطاف پذیر بودن آن و شکل ۲ بیان کننده صعودی، نزولی و وانی شکل بودن تابع خطر این توزیع می باشد. البته می توان به کمک قضیه گلاسز [۷] شرایط صعودی، نزولی، وانی شکل و تک مدی بودن تابع خطر هر توزیعی را بررسی کرد. برای جزئیات بیشتر در باره کاربرد قضیه گلاسز به [۱۵] رجوع شود که رفتار تابع خطر توزیع لیندلی-لگاریتمی را بررسی کرده است. قضیه گلاسز به صورت زیر بیان می شود:

قضیه ۲.۲. (قضیه گلاسز) فرض کنید $\eta(t) = -\frac{f'(t)}{f(t)}$ و $g(t) = \frac{1}{h(t)}$ باشند که $h(t)$ تابع خطر متناظر با تابع چگالی احتمال $f(t)$ است.

۱. اگر به ازای همه مقادیر $t > 0$ داشته باشیم $\eta'(t) = \frac{d\eta(t)}{dt} > 0$ ، تابع خطر، تابعی صعودی است.

۲. اگر به ازای همه مقادیر $t < 0$ ، $\eta'(t) < 0$ باشد آن گاه تابع خطر، تابعی نزولی است.

۳. فرض کنید یک مقدار $t_0 > 0$ وجود دارد به طوری که به ازای همه مقادیر $t \in (0, t_0)$ داریم $\eta'(t) < 0$ و $\eta'(t_0) = 0$ و به ازای همه مقادیر $t > t_0$ داریم $\eta'(t) > 0$. آن گاه تابع خطر، وانی شکل است اگر $y_0 > 0$ وجود داشته باشد به طوری که $g'(y_0) = 0$ باشد و در غیر این صورت، تابع خطر صعودی است.

۴. فرض کنید یک مقدار $t_0 > 0$ وجود دارد به طوری که به ازای همه مقادیر $t \in (0, t_0)$ داریم $\eta'(t) > 0$ و $\eta'(t_0) = 0$ و به ازای همه مقادیر $t > t_0$ داریم $\eta'(t) < 0$. آن گاه تابع خطر، تک مدی است اگر $y_0 > 0$ وجود داشته باشد به طوری که $g'(y_0) = 0$ باشد و در غیر این صورت تابع خطر، نزولی است.

$$E(X^r) = \frac{r!}{\beta^r} e^{\frac{\alpha}{\beta}} E_1^{r-1} \left(\frac{\alpha}{\beta} \right)$$

که $E_s^n(z) = \int_1^\infty (\ln x)^n x^{-s} e^{-zx} dx$ تابع انتگرال نمایی-تعمیم یافته^۳ نام دارد.

اثبات. برای اثبات به [۱۲] مراجعه شود. □

با استفاده از روش های مختلف، خانواده توزیع های با تابع خطر انعطاف پذیرتر معرفی می شوند که به عنوان توزیع های طول عمر در تحلیل بقا می توانند نقش مهم داشته باشند. یکی از این روش ها، روش مارشال-الکین [۱۳] است که یک خانواده توزیع جدید با یک پارامتر اضافی معرفی می کند که تابع توزیع آن به صورت

$$F(x) = \frac{1 - \bar{G}(x)}{1 - (1-p)\bar{G}(x)}, \quad p > 0, -\infty < x < \infty \quad (3)$$

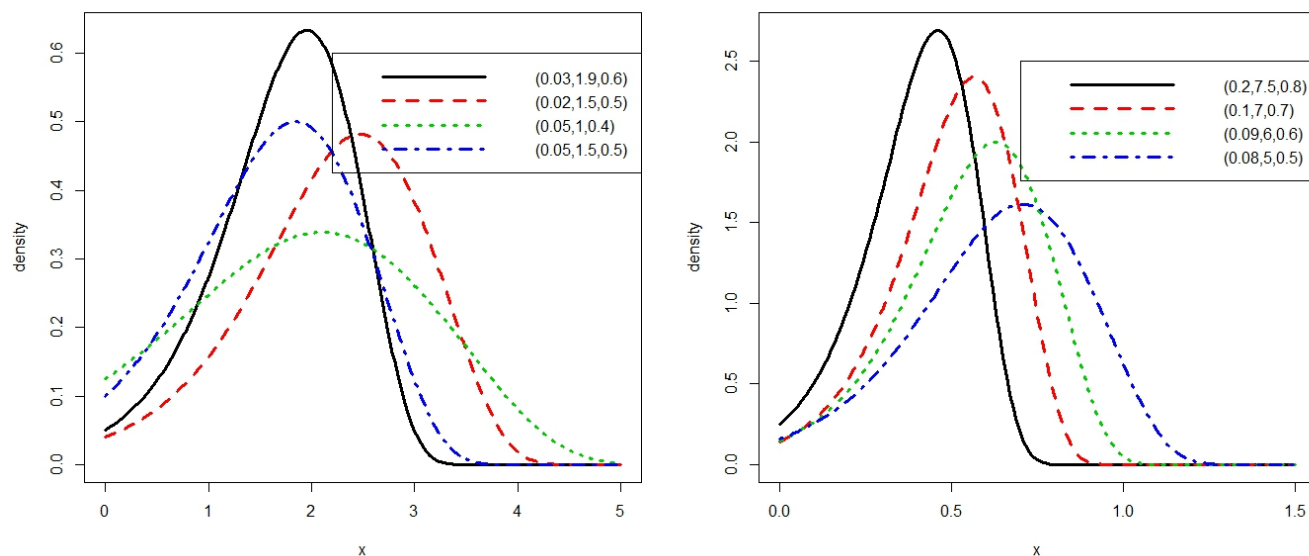
و تابع چگالی احتمال آن به صورت

$$f(x) = \frac{pg(x)}{(1 - (1-p)\bar{G}(x))^2}, \quad p > 0, -\infty < x < \infty \quad (4)$$

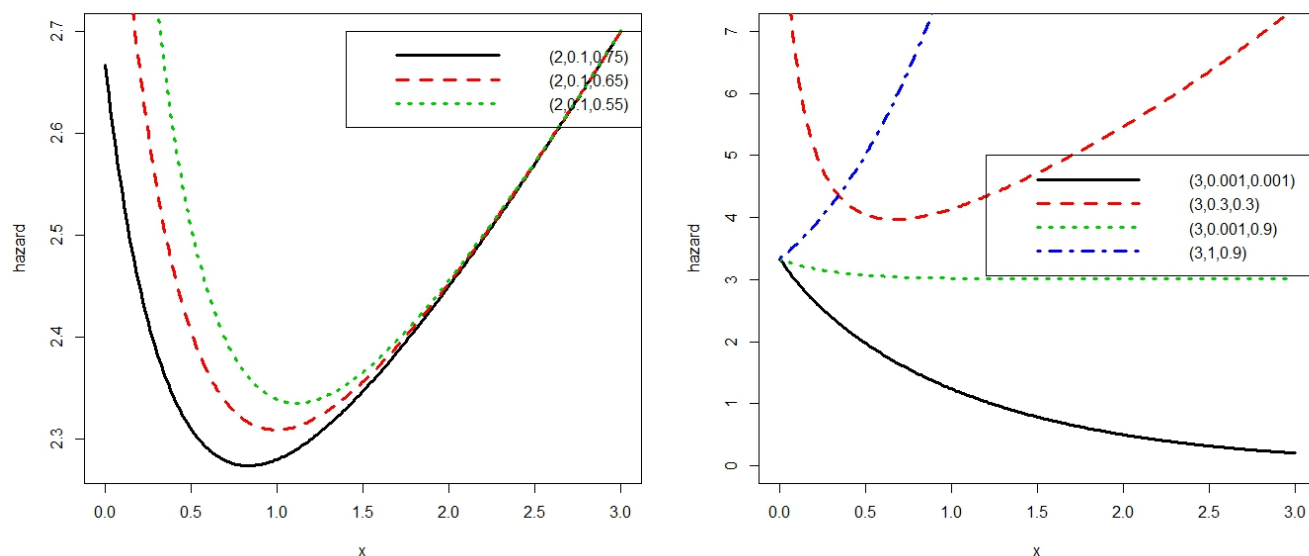
می باشد و $\bar{G}(x) = 1 - G(x)$ که $G(x)$ یک تابع توزیع پیوسته دلخواه و $g(x)$ نیز تابع چگالی احتمال متناظر با آن است. خانواده توزیع با تابع توزیع داده شده در رابطه (۳) و تابع چگالی احتمال داده شده در رابطه (۴)، خانواده توزیع مارشال-الکین-G نامیده می شود که در این مقاله با نماد MOG نشان داده می شود. اگر در روابط (۳) و (۴)، G یک توزیع نمایی باشد، یک خانواده توزیع دو پارامتری تولید می شود که انعطاف پذیرتر از توزیع های دو پارامتری وایبول، گاما و لگ نرمال است. در این مقاله با جایگذاری روابط (۱) و (۲) به ترتیب به جای $G(x)$ و $g(x)$ در روابط (۳) و (۴)، توزیع مارشال-الکین گومپرتز^۴ تولید می شود که تابع توزیع و تابع چگالی احتمال آن به ترتیب به صورت زیر است.

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

$$f(x) = \frac{\alpha p e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}}{\left[1 - (1-p)e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}\right]^2}, \quad x > 0, \alpha, \beta, p > 0$$



شکل ۱. نمودارهای تابع چگالی احتمال توزیع مارشال-الکین گومپرتز به ازای مقادیر مختلف (α, β, p)



شکل ۲. نمودارهای تابع خطر مارشال-الکین گومپرتز به ازای مقادیر مختلف (α, β, p)

۱.۲ گشتاورهای مرکزی

اکنون به رابطه‌ای اشاره می‌شود که در این بخش و بخش‌های بعدی از آن استفاده می‌شود.

$$(1-z)^\delta = \sum_{j=0}^{\infty} (-1)^j \binom{\delta}{j} z^j, \quad \delta \in R, |z| < 1 \quad (5)$$

اگر متغیر تصادفی X دارای توزیع با تابع چگالی احتمال داده شده در رابطه (۴) باشد، $f(x)$ با توجه به کاربرد رابطه (۵) برای

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

$$\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}$$

به صورت زیر بازنویسی می‌شود:

$$f(x) = \alpha p e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \times \sum_{j=0}^{\infty} (-1)^j (1-p)^j \binom{-2}{j} e^{-\frac{\alpha}{\beta j}(e^{\beta x}-1)}.$$

۳.۲ آنتروپی‌های رنی و شانون

اگر X یک متغیر تصادفی با تابع چگالی احتمال $f(x)$ و تابع توزیع تجمعی $F(x)$ باشد، آنتروپی رنی [۱۷] به صورت زیر تعریف می‌شود:

$$I(r) = \frac{1}{1-r} \log \left\{ \int_0^\infty f^r(x) dx \right\}, \quad r > 0, r \neq 1$$

اگر $f(x)$ تابع چگالی احتمال توزیع مارشال-الکین گومپرتز باشد به کمک رابطه (۵) داریم:

$$\begin{aligned} f^r(x) &= (\alpha p)^r e^{r\beta x} \\ &\times \sum_{j=0}^{\infty} \binom{r+j-1}{j} (p-1)^j e^{-\frac{(j+r)\alpha}{\beta} (e^{\beta x}-1)}, \\ \int_0^\infty f^r(x) dx &= (\alpha p)^r \\ &\times \sum_{j=0}^{\infty} \left\{ \binom{r+j-1}{j} (p-1)^j e^{-\frac{(j+r)\alpha}{\beta}} \right. \\ &\times \left. \int_0^\infty e^{r\beta x} e^{-\frac{(j+r)\alpha}{\beta} e^{\beta x}} dx \right\} \\ &= \int_0^\infty u^{r-1} e^{-\frac{(j+r)\alpha}{\beta} u} du. \end{aligned}$$

به کمک تغییر متغیر $u = e^{\beta x}$ داریم:

$$I_1 = \int_0^\infty e^{r\beta x} e^{-\frac{(j+r)\alpha}{\beta} e^{\beta x}} dx = \int_1^\infty u^{r-1} e^{-\frac{(j+r)\alpha}{\beta} u} du.$$

اگر $r < 1$ باشد، خواهیم داشت $I_1 = E_{(1-r)} \left(\frac{(j+r)\alpha}{\beta} \right)$ که $E_s(z) = \int_1^\infty x^{-s} e^{-zx} dx$ برای اطلاع بیشتر در باره $E_s(z)$ به [۱۲] رجوع شود. اگر $r > 1$ باشد، I_1 به صورت زیر محاسبه می‌شود:

$$I_1 = \left[\frac{\beta}{(r+j)\alpha} \right]^r \left(r! - \Gamma\left(\frac{(j+r)\alpha}{\beta}, r\right) \right) = h(\alpha, \beta, j),$$

به طوری که $\Gamma(a, r) = \int_a^\infty x^{r-1} e^{-x} dx$. بنا بر این با فرض

$$I = \int_0^\infty f^r(x) dx$$

داریم:

$$I = \begin{cases} (\alpha p)^r \sum_{j=0}^{\infty} S(\alpha, \beta, p, j) E_{(1-r)} \left(\frac{(j+r)\alpha}{\beta} \right), & r < 1 \\ (\alpha p)^r \sum_{j=0}^{\infty} S(\alpha, \beta, p, j) h(\alpha, \beta, j), & r > 1 \end{cases}$$

به طوری که $S(\alpha, \beta, p, j) = \binom{r+j-1}{j} (p-1)^j e^{-\frac{(j+r)\alpha}{\beta}}$ بنا بر این به کمک I آنتروپی رنی به راحتی به دست می‌آید. آنتروپی

گشتاور مرکزی مرتبه r ام متغیر تصادفی X به صورت زیر به دست می‌آید.

$$\begin{aligned} E(X^r) &= \sum_{j=0}^{\infty} \left[\frac{(p-1)^j}{j+1} \binom{j+1}{j} \right. \\ &\times \left. \int_0^\infty \left(x^r (j+1) \alpha p e^{\beta x} e^{-\frac{(j+1)\alpha}{\beta} (e^{\beta x}-1)} \right) dx \right]. \end{aligned}$$

چون $\int_0^\infty \left(x^r (j+1) \alpha p e^{\beta x} e^{-\frac{(j+1)\alpha}{\beta} (e^{\beta x}-1)} \right) dx$ گشتاور مرکزی مرتبه r ام توزیع گومپرتز با پارامترهای $(j+1)\alpha$ و β است، با توجه به قضیه ۱.۲ داریم:

$$\begin{aligned} E(X^r) &= \frac{r! p}{\beta^r} \sum_{j=0}^{\infty} \left[\binom{j+1}{j} \frac{(p-1)^j}{j+1} e^{\frac{(j+1)\alpha}{\beta}} \right. \\ &\times \left. E_1^{(r-1)} \left(\frac{(j+1)\alpha}{\beta} \right) \right]. \end{aligned}$$

۲.۲ گشتاورهای آماره‌های مرتب

اگر $X_{i:n}$ آماره ترتیبی یک نمونه تصادفی n تایی از توزیع با تابع چگالی احتمال داده شده در رابطه (۴) باشد، تابع چگالی احتمال آن $f_{i:n}(x)$ به صورت

$$f_{i:n}(x) = \frac{f(x)(F(x))^{i-1}(1-F(x))^{n-i}}{B(i, n-i+1)}, \quad x > 0 \quad (6)$$

محاسبه می‌شود که در آن $f(x)$ و $F(x)$ به ترتیب تابع چگالی احتمال و تابع توزیع تجمعی توزیع مارشال-الکین گومپرتز می‌باشند. با جایگذاری روابط (۳) و (۴) به ترتیب به جای $F(x)$ و $f(x)$ در رابطه (۶) و با استفاده از بسط دوجمله‌ای خیام-نیوتن، تابع چگالی احتمال آماره ترتیبی $X_{i:n}$ به صورت زیر به دست می‌آید.

$$\begin{aligned} f_{i:n}(x) &= \frac{\alpha p e^{\beta x}}{B(i, n-i+1)} \\ &\times \sum_{j=0}^{n-i} \sum_{k=0}^{i+j-1} \frac{(-1)^{j+k} e^{-\frac{(k+1)\alpha}{\beta} (e^{\beta x}-1)}}{\left[1 - (1-p) e^{-\frac{(j+1)\alpha}{\beta} (e^{\beta x}-1)} \right]^{i+j+1}}, \end{aligned}$$

اکنون به کمک قضیه ۱.۲ و بعد از کمی محاسبات جبری، گشتاور مرکزی مرتبه r ام آماره ترتیبی $X_{i:n}$ به صورت زیر محاسبه می‌شود:

$$\begin{aligned} E(X_{i:n}^r) &= \frac{r! p}{B(i, n-i+1)} \sum_{j=0}^{n-i} \sum_{k=0}^{i+j-1} \sum_{s=0}^{\infty} \left[\frac{(-1)^{j+k} (1-p)^s}{i+j+s} \right. \\ &\times \left. \binom{i+j+s}{s} e^{\frac{(i+j+s)\alpha}{\beta}} E_1^{(r-1)} \left(\frac{(i+j+s)\alpha}{\beta} \right) \right] \end{aligned}$$

برآورد ماکسیمم درست‌نمایی Θ که با نماد $\hat{\Theta}$ نشان داده می‌شود از حل هم‌زمان معادله‌های $\frac{\partial \ell_n}{\partial \alpha} = 0$ و $\frac{\partial \ell_n}{\partial \beta} = 0$ و $\frac{\partial \ell_n}{\partial p} = 0$ به دست می‌آید که به ترتیب به صورت زیر می‌باشند.

$$\frac{2(1-p)}{\beta} \sum_{i=1}^n \frac{(e^{\beta x_i} - 1) e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{1 - (1-p)e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)}} + \frac{1}{\beta} \sum_{i=1}^n (e^{\beta x_i} - 1) - \frac{n}{\alpha} = 0,$$

$$2(p-1) \sum_{i=1}^n \frac{[(e^{\beta x_i} - 1) - \beta x_i e^{\beta x_i}] e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{1 - (1-p)e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)}} + \sum_{i=1}^n (e^{\beta x_i} - 1) - \beta \sum_{i=1}^n x_i e^{\beta x_i} = 0,$$

$$2 \sum_{i=1}^n \frac{e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{1 - (1-p)e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)}} - \frac{n}{p} = 0.$$

البته برآوردها به صورت بسته به دست نمی‌آیند و با استفاده از بسته maxLik در نرم‌افزار R قابل حل می‌باشند. همچنین توزیع مجانبی $\hat{\Theta}$ با توجه به [۶] توزیع $N(\Theta, \mathbf{I}^{-1}(\hat{\Theta}))$ است که در آن داریم:

$$\mathbf{I}^{-1}(\Theta) = \begin{bmatrix} -I_{\alpha\alpha} & -I_{\alpha\beta} & -I_{\alpha p} \\ -I_{\beta\alpha} & -I_{\beta\beta} & -I_{\beta p} \\ -I_{p\alpha} & -I_{p\beta} & -I_{pp} \end{bmatrix}^{-1}$$

به‌طوری که $I_{\theta_i \theta_j} = \frac{\partial^2 \ell_n}{\partial \theta_i \partial \theta_j}$ است. با جایگذاری $\hat{\Theta}$ به جای Θ داریم:

$$\mathbf{I}^{-1}(\hat{\Theta}) = \begin{bmatrix} \text{Var}(\hat{\alpha}) & \text{Cov}(\hat{\alpha}, \hat{\beta}) & \text{Cov}(\hat{\alpha}, \hat{p}) \\ \text{Cov}(\hat{\beta}, \hat{\alpha}) & \text{Var}(\hat{\beta}) & \text{Cov}(\hat{\beta}, \hat{p}) \\ \text{Cov}(\hat{p}, \hat{\alpha}) & \text{Cov}(\hat{p}, \hat{\beta}) & \text{Var}(\hat{p}) \end{bmatrix}$$

که نقش مهمی در محاسبه فاصله اطمینان و آزمون فرض پارامترهای α و β و p دارد.

۳ تحلیل داده‌های واقعی

در این بخش با استفاده از یک مجموعه داده‌های واقعی، توزیع مارشال-الکین گومپرتز با توزیع‌های مک‌دونالد گومپرتز، گومپرتز تعمیم‌یافته، بتا-گومپرتز، گومپرتز-پوآسن و گومپرتز-

شانون [۱۹] به صورت $E(-\log f(X))$ تعریف می‌شود. البته آنتروپی شانون را نیز می‌توان از آنتروپی رنی به کمک رابطه $\lim_{r \rightarrow 1} I(r)$ به دست آورد.

۴.۲ تابع چندک

اگر X یک متغیر تصادفی با توزیع $\text{MOG}(\alpha, \beta, p)$ باشد، تابع چندک آن به ازای $u \in (0, 1)$ به صورت زیر تعریف می‌شود.

$$Q(u) = F^{-1}(u) = \frac{1}{\beta} \log \left[1 - \frac{1}{\beta} \log \left(\frac{1-u}{1-u(1-p)} \right) \right], \quad (7)$$

که در آن $F^{-1}(u)$ وارون تابع توزیع تجمعی مارشال-الکین گومپرتز می‌باشد. رابطه (۷) در شبیه‌سازی توزیع مارشال-الکین گومپرتز بسیار مفید است، بدین صورت که اگر U متغیری تصادفی و دارای توزیع یکنواخت پیوسته در بازه $(0, 1)$ باشد، متغیر تصادفی $X = Q(U)$ دارای توزیع $\text{MOG}(\alpha, \beta, p)$ می‌شود. همچنین با استفاده از تابع چندک، میزان کشیدگی توزیع مارشال-الکین گومپرتز با توجه به [۱۴] به صورت

$$K = \frac{Q(\frac{9}{10}) - Q(\frac{5}{10}) + Q(\frac{3}{10}) - Q(\frac{1}{10})}{Q(\frac{6}{10}) - Q(\frac{2}{10})},$$

و میزان چولگی آن با توجه به [۱۱] به صورت

$$S = \frac{Q(\frac{7}{8}) - 2Q(\frac{1}{2}) + Q(\frac{1}{8})}{Q(\frac{3}{4}) - Q(\frac{1}{4})},$$

حاصل می‌شود که برای توزیع نرمال استاندارد و توزیع تی استودنت با ۱۰ درجه آزادی مقدار K ، برابر با صفر، مقدار S به ترتیب ۱/۲۳۳۱ و ۱/۲۷۷۰۵ می‌باشد.

۵.۲ برآورد پارامترها

در این بخش بر اساس نمونه تصادفی $X_1, \dots, X_n$ از توزیع مارشال-الکین گومپرتز، پارامترهای آن یعنی α و β و p به روش ماکسیمم درست‌نمایی برآورد می‌شوند. بنا بر این لگاریتم تابع درست‌نمایی با فرض $\Theta = (\alpha, \beta, p)$ به صورت زیر است:

$$\ell_n = \log L(x; \Theta)$$

$$= n \log \alpha + n \log p + \sum_{i=1}^n x_i - \frac{\alpha}{\beta} \sum_{i=1}^n (e^{\beta x_i} - 1) - 2 \sum_{i=1}^n \log \left[1 - (1-p)e^{-\frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right].$$

لگاریتمی با تابع‌های چگالی احتمال

و اطلاع آکائیکه سازگار^۶ است و جدول ۳ شامل آماره‌های نیکویی برازش کرامر-فونمیزس^۷، واتسون^۸ و کولموگوروف-اسمیرنوف^۹ همراه با p -مقدار مربوط می‌باشد که کوچک‌تر بودن مقادیر معیارهای اطلاع و آماره‌های آزمون نیکویی برازش برای هر توزیعی در مقایسه با چند توزیع دیگر دلیل بر بهتر بودن آن توزیع می‌باشد. معیارهای اطلاع آکائیکه و آکائیکه سازگار به‌ترتیب عبارت‌اند از:

$$AIC = -2\ell(\hat{\theta}) + 2k$$

$$CAIC = -2\ell(\hat{\theta}) + \frac{2kn}{n-k-1}$$

که در آن $\ell(\hat{\theta})$ مقدار لگاریتم تابع درست‌نمایی به‌ازای برآورد ماکسیمم درست‌نمایی پارامترها، k تعداد پارامترها و n تعداد مشاهدات است. آماره‌های نیکویی برازش کرامر-فونمیزس، واتسون و کولموگوروف-اسمیرنوف که به‌ترتیب با نمادهای T_{CM} و T_W و T_{K-S} نشان داده می‌شود عبارت‌اند از:

$$T_{CM} = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i, \hat{\alpha}, \hat{\beta}, \hat{p}) - \frac{2i-1}{2n} \right]^2,$$

$$T_W = T_{CM} + \sum_{i=1}^n \left[\frac{F(x_i, \hat{\alpha}, \hat{\beta}, \hat{p})}{n} - \frac{1}{2} \right]^2,$$

$$T_{K-S} = \max_{1 \leq i \leq n} \left[\frac{i}{n} - F(x_i, \hat{\alpha}, \hat{\beta}, \hat{p}), F(x_i, \hat{\alpha}, \hat{\beta}, \hat{p}) - \frac{i-1}{n} \right],$$

که F تابع توزیع مارشال-الکین گومپرتز است.^۵

$$f_{McG}(x) = \frac{\alpha p e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}}{B(\lambda/p, r)} \left(1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \right)^{\lambda-1} \times \left[1 - \left(1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \right)^p \right]^{\gamma-1},$$

$$f_{GG}(x) = \alpha p e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \left(1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \right)^{\lambda-1},$$

$$f_{BG}(x) = \frac{p e^{\beta x} e^{-\frac{\lambda \alpha}{\beta}(e^{\beta x}-1)}}{B(p, \lambda)} \left(1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \right)^{p-1},$$

$$f_{GP}(x) = \frac{\alpha p e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}}{1 - e^{-p}} e^{-p e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)}},$$

$$f_{GL}(x) = \alpha p e^{\beta x} e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} (-\log(1-p))^{-1} \times \left\{ 1 - p \left[1 - e^{-\frac{\alpha}{\beta}(e^{\beta x}-1)} \right] \right\}^{-1},$$

مقایسه می‌شود که در آنها $x > 0$ و $\alpha, \beta, \gamma, \lambda, p > 0$ است. این مجموعه داده‌ها شامل ۶۳ مقدار ثبت‌شده از مقاومت الیاف شیشه‌های به ضخامت ۱/۵ سانتی‌متری توسط یک موسسه فیزیک در انگلستان می‌باشد که در [۲۰] گزارش شده است. داده‌ها به‌صورت جدول ۱ است.

جدول ۲ شامل مقادیر برآورد ماکسیمم درست‌نمایی پارامترهای توزیع‌های مورد مقایسه، معیارهای اطلاع آکائیکه^۵

جدول ۱. داده‌های مربوط به مقاومت الیاف شیشه‌های به ضخامت ۱/۵ سانتی‌متری

۱/۱۳	۱/۲۹	۱/۸۱	۲/۰۱	۰/۷۷	۱/۷۳	۱/۸۱	۲/۰۰	۰/۷۴	۱/۴۲
۲/۲۴	۰/۸۱	۱/۶۶	۱/۶۸	۱/۷۶	۱/۵۲	۱/۵۸	۱/۶۱	۱/۶۴	۱/۶۸
۱/۷۶	۱/۸۴	۱/۵۳	۱/۵۹	۱/۶۱	۰/۹۳	۱/۲۵	۱/۳۶	۱/۳۶	۱/۴۹
۱/۶۶	۱/۶۹	۱/۴۸	۱/۵۰	۱/۲۷	۱/۳۹	۱/۴۹	۱/۱۱	۱/۲۸	۰/۵۵
۱/۵۵	۱/۶۱	۱/۶۱	۱/۶۳	۱/۶۷	۱/۷۰	۱/۷۸	۱/۸۹	۱/۶۰	۱/۶۲
۱/۴۸	۱/۵۱	۱/۷۰	۱/۷۷	۱/۸۴	۰/۸۴	۱/۲۴	۱/۳۰		

^۵ Akaike information criterion (AIC)

^۶ Consistent Akaike information criterion (CAIC)

^۷ Cramer-VonMisses (CM)

^۸ Watson (W)

^۹ Kolmogorov-Smirnov (K-S)

```
&FFF<-function(x,a,b,p){\
&(1-exp((-a/b)*(exp(b*x)-1)))*(1+(p-1)\
&*exp((-a/b)*(exp(b*x)-1)))^{\wedge}(-1)\
&}\}
```

```
&ks.test(x,"FFF",0.7983691, 1.310359,58.98452)
```

به همین روش می توان آماره های T_{CM} و T_W و T_{K-S} با p -مقدار مربوط را برای سایر توزیع ها نیز به دست آورد. البته باید دقت کرد که قسمت های F و FF و FFF برنامه به کمک تابع توزیع مربوط نوشته شود.

بنا بر این با مشاهده جدول های ۲ و ۳، نتیجه گرفته می شود که توزیع مارشال-الکین گومپرتز تعمیم یافته برای برازش به داده ها بهتر از سایر توزیع ها است. همچنین با مقایسه نمودارهای تابع چگالی احتمال توزیع های مورد مقایسه در شکل ۳، برتری توزیع مارشال-الکین گومپرتز تعمیم یافته نسبت به سایر توزیع های انتخابی برای برازش به داده ها مشخص می شود.

۴ نتیجه گیری

در این مقاله، یک توزیع جدید طول عمر سه پارامتری به نام توزیع مارشال-الکین گومپرتز معرفی شد که تعمیمی جدید از توزیع گومپرتز است. برخی از ویژگی های این توزیع جدید مانند گشتاورهای مرکزی، گشتاورهای آماره های مرتب، آنتروپی های رنی و شانون و تابع چندک به دست آورده شد. همچنین پارامترهای آن به روش ماکسیمم درست نمایی برآورد شد و به کمک یک مجموعه داده های واقعی و با استفاده از آماره های آزمون نیکویی برازش و معیارهای اطلاع، این توزیع جدید با برخی از تعمیم های توزیع گومپرتز مقایسه شد و نشان داده شد که در مقایسه با سایر توزیع ها، می تواند مدلی مناسب برای توزیع طول عمر در تحلیل بقا باشد.

برنامه محاسبه آماره های آزمون نیکویی برازش در نرم افزار R برای توزیع MOG به صورت زیر است.

۱. محاسبه آماره T_{CM}

```
&x<-c(1.27, 1.04, 0.74, 2 , 1.81, 1.73, 1.68,\
& 1.64, 1.61, 1.58, 1.52, 1.49, 1.32, 1.25, 0.93,\
& 0.55, 1.50, 1.54, 1.42, 1.28, 1.11, 0.77, 2.01,\
& 1.81, 1.76, 1.68, 1.66, 1.61, 1.59, 1.53, 1.49,\
&1.39, 1.66, 1.62, 1.61, 1.55, 1.50, 1.48, 1.29,\
&1.13,0.81, 2.24, 1.84, 1.76, 1.69, 1.66, 1.62,\
&1.6, 1.89, 1.78, 1.7, 1.67, 1.63, 1.61, 1.55,\
&1.51, 1.48, 1.30, 1.24, 0.84, 1.84, 1.77, 1.70)\
&y<-sort(x)\
&n<-length(x)\
&z<-seq(1,n)\
&F<-function(y,z,a,b,p){\
&\left[(1-exp((-a/b)\ast(exp(b*y[z])-1)))\
&*(1+(p-1)*exp((-a/b)*(exp(b*y[z])-1)))\
&^{\wedge}(-1)-((2*z-1)/(2*n))]^{\wedge}(2)\
&w<-function(y,z,a,b,p){\
&(1/(12*n))+sum(F(y,z,a,b,p))\
&w(y,z,0.7983691, 1.310359 ,58.98452)
```

۲. محاسبه آماره T_W

```
&FF<-function(y,z,a,b,p){\
&(1-exp((-a/b)*(exp(b*y[z])-1)))*(1+(p-1)\
&*exp((-a/b)*(exp(b*y[z])-1)))^{\wedge}(-1)\
&U<-function(y,z,a,b,p){\
&w(y,z,a,b,p)+sum(((FF(y,z,a,b,p)/n)-(1/2))\
&^{\wedge}(2))\
&U(y, z, 0.7983691, 1.310359 ,58.98452)
```

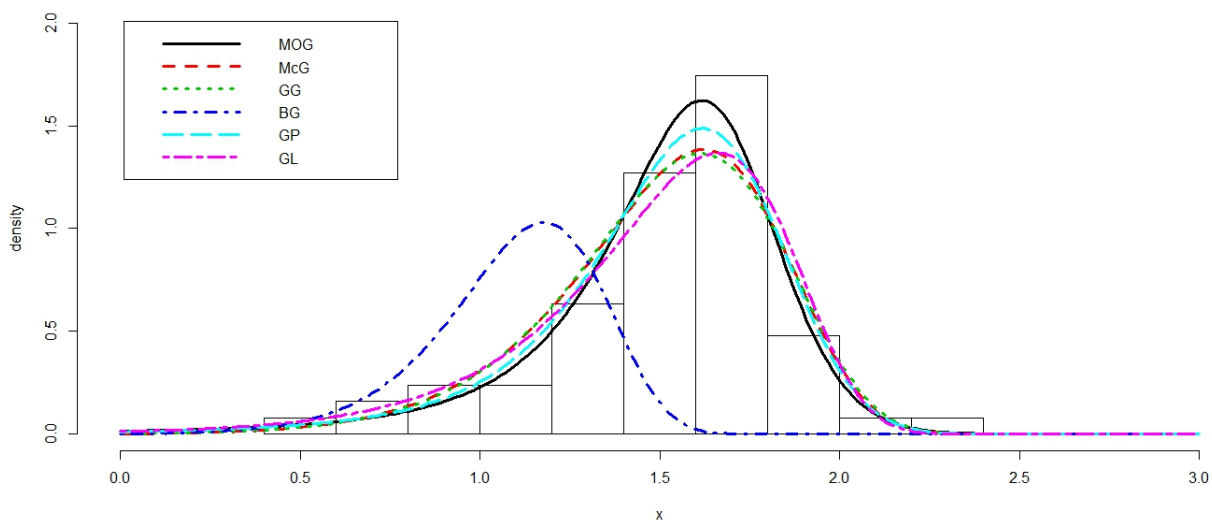
۳. محاسبه آماره T_{K-S} با p -مقدار مربوط

جدول ۲. برآورد ماکسیمم درست‌نمایی پارامترها و مقادیر معیارهای اطلاع برای داده‌های مقاومت الیاف شیشه

توزیع‌ها	$\hat{\alpha}$	$\hat{\beta}$	$\hat{p}$	$\hat{\lambda}$	$\hat{\gamma}$	AIC	CAIC
MOG	۰/۷۹۸۳۶	۱/۳۱۰۳	۵۸/۹۸۴	—	—	۳۰/۶۴	۳۳/۱۸
McG	۰/۰۶۹۹۰	۲/۱۲۷۹	۴/۸۱۸۰	۲/۰۶۱۲	۲/۶۶۹۲	۳۷/۷۱	۳۸/۵۹
GG	۰/۰۳۶۸۰	۲/۸۶۴۷	۱/۶۲۵۳	—	—	۳۴/۲۹	۳۴/۵۹
BG	۰/۰۳۵۷۰	۲/۸۴۸۷	۱/۶۳۵۷	۱/۰۶۵۰	—	۳۶/۲۹	۳۶/۸۴
GP	۰/۱۰۲۱۱	۲/۳۷۳۰	۳/۵۶۷۸	—	—	۳۲/۷۰	۳۵/۲۴
GL	۰/۰۱۶۴۶	۳/۳۶۳۹	۰/۵۳۳۰۱	—	—	۳۵/۵۸	۳۵/۹۹

جدول ۳. مقادیر آماره‌های آزمون نیکویی برازش برای داده‌های مقاومت الیاف شیشه

توزیع‌ها	کرامر-فون میزس	واتسون	کولموگوروف-اسمیرنوف	p -مقدار
MOG	۰/۰۷۵۰۸۶	۱۵/۳۳۰۳	۰/۰۹۷۴	۰/۵۸۹۰
McG	۰/۰۸۸۳۲۴	۱۵/۳۹۷۲	۰/۱۳۰۸	۰/۲۳۱۳
GG	۰/۱۵۶۸۰۹	۱۵/۴۰۵۱	۰/۱۳۲۲	۰/۲۲۰۸
BG	۸/۴۸۸۹۸	۲۳/۴۳۷۲	۰/۶۲۷۰	۰
GP	۰/۱۰۶۵۵۱	۱۵/۵۲۸۴	۰/۴۰۹۳	۰/۱۱۱۹
GL	۰/۱۵۶۷۶۱۸	۱۵/۴۰۶۲۸	۰/۱۲۷۵	۰/۲۶۷۰



شکل ۳. نمودار برآورد تابع چگالی احتمال توزیع‌های مورد مقایسه برای داده‌های مقاومت الیاف شیشه‌ها

مراجع

- [1] Bemmaor, A.C. and Glady, N, (2012), Modeling purchasing behavior with sudden death: A flexible customer lifetime model, *Management Science*, **58**(5), 1012-1021.
- [2] Gompertz, B, (1825), On the nature of the function expressive of the law of human mortality and on the new mode of determining the value of life contingencies, *Philosophical Transactions of the Royal Society American*, **115**, 513-580

- [3] Economos, A.C, (1982), Rate of aging, rate of dying and the mechanism of mortality, *Archives of Gerontology and Geriatrics*, **1**(1), 46-51.
- [4] El-Gohary, A. and Al-Otaibi, A.N, (2013), The generalized Gompertz distribution, *Applied Mathematical Modelling*, **37**(2), 13-24.
- [5] Gavrilov, L. and Govrilova, N, (1991), *The biology of Life Span: A Quantitative approach*: Chur: Harwood.
- [6] Gupta, R.D. and Kundu, D, (1999), Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, **41**, 173-188.
- [7] Glaser, R.E, (1980), Bathtub and Related failure rate characterizations, *Journal of the American Statistical Association*, **75**, 667-672.
- [8] Johnson, N.L, Kotz, S. and Balakrishnan, N, (1995), *Continuous Univariate Distributions*, Vol. 2, 2nd ed., John Wiley and Sons.
- [9] Jafari, A.A., Tahmasebi, S. and Alizadeh, M, (2014), The beta-Gompertz distribution, *Revista Colombiana de Estadística*, **37**, 141-158.
- [10] Jafari, A.A. and Tahmasebi, S, (2015), Gompertz-power series distributions, *Communications in Statistics-Theory and Methods*, **45**(13), 3761-3781.
- [11] Kenney, J.F. and Keeping, E, (1962), *Mathematics of Statistics*, D. Van Nostrand Company.
- [12] Milgaram, M, (1985), The generalized integro-exponential function, *Mathematics and Computation*, **44**(170), 443-458.
- [13] 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.
- [14] Moors, J.J.A, (1988), A quantile alternative for kurtosis, *Journal of the Royal Statistical Society, Series D (The Statistician)*, **37**(1), 25-32.
- [15] Mahmoodi, E. and Abolhasani, S, (2016), The Lindley-Logarithmic distribution: Models and Properties, *Journal of Statistical Sciences*, **10**(1), 139-158.
- [16] Ohishi, K, Okamura, H. and Dohi, T, (2009), Gompertz software reliability model: Estimation algorithm and empirical validation, *Journal of Systems and Software*, **82**(3), 535-543.
- [17] Renyi, A. (1961), *On measures of entropy and information*, In: Proceedings of Berkeley symposium, statistics, probability, vol1, pp 547-561.
- [18] Roozegar, R., Tahmasebi, S., Jafari, A. A. (2017). The McDonald Gompertz distribution: properties and applications. *Communications in Statistics-Simulation and Computation*, **46**(5), 3341-3355.
- [19] Shannon, C.E, (1951), Prediction and entropy of printed English, *Bell System Technical Journal*, **30**, 50-64.
- [20] Smith, R.L, Naylor, J.C, (1987), A comparison of maximum likelihood and bayesian estimators for the three-parameter Weibull distribution, *Applied Statistics*, **36**, 358-369.

- [21] Yaghoobzadeh Shahrastani, S, Shadrokh, A. and Yarmohammadi, M, (2015), Introduced a new Lifetime distribution of Power Series distribution of the family Gompertz, *Mosaheb Mathematical Researches*, **1**(2), 72-95.
- [22] Wetterstrand, W.H, (1981), Parametric models for life insurance mortality data: Gompertz's law over time, *Transactions of Society of Actuaries*, **33**, 159-179.
- [23] Willemse, W. and Koppelaar, H, (2000), Knowledge elicitation of Gompertz law of mortality, *Scandinavian Actuarial Journal*, **2**, 168-179.