توزیع لجستیک-نرمال: مشخصات و کاربردها

فتانه نظام‌پور،[۱] علیرضا سلیمانی[۲]

تاریخ دریافت: ۱۳۹۴/۹/۱۰
تاریخ پذیرش: ۱۳۹۶/۱/۳۰

چکیده:
در این مقاله برخی ویژگی‌های خانوادهٔ لجستیک-X و عضوی از این خانواده، توزیع لجستیک-نرمال، در جزئیات مورد مطالعه قرار گرفته است. میانگین انحرافات، تابع خطر و مد برای توزیع لجستیک-نرمال به‌دست‌آمده است. همچنین در این مقاله از روش درست‌نمایی ماکسیمم برای برآورد پارامترها و از یک مجموعه‌داده برای نشان دادن برنامه‌های کاربردی، توزیع لجستیک-نرمال استفاده شده است.

واژه‌های کلیدی: توزیع T-X، توزیع لجستیک، تابع خطر، معیار اطلاع آکائیکه، معیار اطلاع بیزی، برآورد درست‌نمایی ماکسیمم.

۱ مقدمه

برای تولید توزیع‌های پیوسته روش‌هایی وجود دارند. اخیراً [۱] روشی جدید برای تولید خانواده‌ای از توزیع‌ها را ارائه دادند و آن را خانوادهٔ T-X نامیدند. برای بررسی روش‌های تولید توزیع‌های پیوسته یک‌متغیره می‌توان به [۳]مراجعه کرد. در این مقاله نخست برخی ویژگی‌های کلّی خانوادهٔ لجستیک-X را مطالعه می‌کنیم و پس از آن جزئیات توزیع لجستیک-نرمال را که یک عنصر از خانوادهٔ لجستیک-X است، مطالعه می‌کنیم. فرض کنیم $F(x)$ تابع توزیع تجمعی متغیر تصادفی X و $r(t)$ تابع چگالی احتمال متغیر تصادفی Tتعریف‌شده روی $(-\infty,\infty)$ باشد. تابع توزیع تجمعی خانوادهٔ T-X از توزیع‌های تعریف شده توسط [۱] به‌صورت زیر است:

$$G(x)=\int_{\cdot}^{-log(\text{۱}-F(x))} r(t)dt \qquad (\text{۱})$$

خانوادهٔ توزیع‌های تعریف‌شده در (۱) را خانوادهٔ مبدل تبدیل‌شده یا خانوادهٔ T-X نامیدند. تابع چگالی احتمال خانواده T-X به قرار زیر است:

$$r(t)=\frac{\frac{\text{۱}}{\beta}e^{-\frac{t-\alpha}{\beta}}}{(\text{۱}+e^{-\frac{t-\alpha}{\beta}})^{\text{۲}}} \qquad (\text{۲})$$

تعریف (۲) منجر به توزیع لجستیک-X با تابع چگالی احتمال زیر می‌گردد:

$$\begin{aligned} g(x)&=\frac{f(x)}{\text{۱}-F(x)}r(-log(\text{۱}-F(x)))\\ &=\frac{f(x)}{\text{۱}-F(x)}\frac{\frac{\text{۱}}{\beta}e^{\frac{log(\text{۱}-F(x))+\alpha}{\beta}}}{(\text{۱}+e^{\frac{log(\text{۱}-F(x))+\alpha}{\beta}})^{\text{۲}}} \qquad (\text{۳})\end{aligned}$$

تابع توزیع لجستیک-X در (۳) به‌صورت زیر نوشته می‌شود:

$$\begin{aligned} G(x)&=\int_{\cdot}^{-log(\text{۱}-F(x))}\frac{\frac{\text{۱}}{\beta}e^{-\frac{t-\alpha}{\beta}}}{(\text{۱}+e^{-\frac{t-\alpha}{\beta}})^{\text{۲}}}dt\\ &=\frac{\text{۱}}{\text{۱}+e^{\frac{log(\text{۱}-F(x))+\alpha}{\beta}}}-\frac{\text{۱}}{\text{۱}+e^{\alpha/\beta}} \qquad (\text{۴})\end{aligned}$$

۲ خانوادهٔ لجستیک-X

در این بخش، برخی خواص عمومی لجستیک-X در (۳) را مورد بحث قرار می‌دهیم.

قضیهٔ ۱.۲. اگر متغیر تصادفی Y از توزیع لجستیک با پارامترهای α و β پیروی کند وF یک تابع توزیع دلخواه باشد، متغیر تصادفی $X=F^{-\text{۱}}(\text{۱}-e^{-Y})$ از توزیع لجستیک-X با پارامترهای α و β پیروی می‌کند.

اثبات. با استفاده از روش تبدیل‌ها نتیجه حاصل می‌شود. □

[۱] دانش‌آموختهٔ کارشناسی ارشد آمار، دانشگاه حکیم سبزواری خراسان رضوی، ایران
[۲] عضو هیئت علمی گروه آمار، دانشگاه حکیم سبزواری خراسان رضوی، ایران

تابع خطر خانوادهٔ لجستیک-X به‌صورت زیر است:

$$\begin{aligned} h(x) &= \frac{g(x)}{۱-G(x)} \\ &= \frac{(۱+e^{\alpha/\beta})f(x)e^u}{\beta(۱-F(x))(۱+e^u)(۱+e^u(۲+e^{\alpha/\beta}))} \end{aligned} \quad (۵)$$

که در آن

$$e^u = e^{(log(۱-F(x))+\alpha)/\beta}$$

و $-\infty < X < \infty$ است. برخی از حالت‌های خاص این خانواده به قرار زیرند:

۱. وقتی $\alpha = ۰$ و $\beta = ۱$ باشد، خانوادهٔ لجستیک-X در (۳) به‌صورت زیر تقلیل می‌یابد:

$$g(x) = f(x)(۲-F(x))^{-۲} \quad (۶)$$

۲. وقتی $\alpha = ۰$ و $\beta = ۱$ باشد، خانوادهٔ لجستیک-X در (۴) به‌صورت زیر تقلیل می‌یابد:

$$G(x) = \frac{۱}{۲-F(x)} - \frac{۱}{۲} \quad (۷)$$

۳. وقتی $\alpha = ۰$ و $\beta = ۱$ باشد، تابع چندک خانوادهٔ لجستیک-X به‌صورت زیر تقلیل می‌یابد:

$$Q(p) = F^{-۱}\left(\frac{۴p}{۲p+۱}\right) \quad (۸)$$

۴. وقتی $\alpha = ۰$ و $\beta = ۱$ باشد، تابع خطر خانوادهٔ لجستیک-X به‌صورت زیر تقلیل می‌یابد:

$$h(x) = \frac{۲f(x)}{(۲-F(x))(۴-۳F(x))} \quad (۹)$$

۳ توزیع لجستیک-نرمال وچند ویژگی آن

اکنون چند خاصیت از خانوادهٔ لجستیک-X را ارائه می‌کنیم.

قضیهٔ ۱.۳. مُد توزیع لجستیک-نرمال جواب معادلهٔ زیر است:

$$X = \mu + \sigma^۲ h_N(x)\left(۱ - \frac{۱}{\beta} - \frac{۲}{\beta(۱+e^{\frac{\alpha - H_N(x)}{\beta}})}\right) \quad (۱۰)$$

که $h_N(x) = \frac{\phi(x)}{۱-\Phi(x)}$ و $H_N(x) = -log(۱-\Phi(x))$ به‌ترتیب تابع خطر و تابع تجمعی خطر برای توزیع نرمال‌اند.

اثبات. با استفاده از $\Phi'(x) = -\sigma^{-۲}(x-\mu)\phi(x)$، مشتق نسبت به X از (۵) می‌تواند به‌صورت زیر حاصل شود.

$$g'(x) = \frac{\frac{e^{\alpha/\beta}}{\beta}\phi(x)(۱-\Phi(x))^{۱/\beta - ۱}}{(۱+e^{\alpha/\beta}(۱-\Phi(x))^{۱/\beta})^۲}K(x) \quad (۱۱)$$

که

$$\begin{aligned} K(x) = &-\sigma^۲(X-\mu) + \phi(۱-\Phi(x))^{-۱} \\ &\left(۱ - \frac{۱}{\beta} - \frac{۲}{\beta(۱+e^{\frac{\alpha + log(۱-\Phi(x))}{\beta}})}\right) \end{aligned} \quad (۱۲)$$

نتیجه قضیه با توجه به $K(x) = ۰$ حاصل می‌شود که هم‌ارز رابطهٔ (۱۰) است. □

وقتی $\mu = ۰$ باشد، رابطهٔ (۱۱) نشان می‌دهد که مد در حالت $\alpha \geq ۱$ و $\beta \geq ۱$، نامنفی و در حالت $\alpha < ۱$ و $\beta < ۱$، منفی است. همچنین مُد یک تابع افزایشی از β و μ و σ و یک تابع کاهشی از α است. قضیهٔ زیر تابع چندک را برای توزیع لجستیک-نرمال ارائه می‌کند.

قضیهٔ ۲.۳. اگر $Q(p)$، $۰ < p < ۱$ نمایانگر تابع چندک برای توزیع لجستیک-نرمال باشد آن‌گاه:

$$Q(p) = \Phi^{-۱}\left(۱ - \left(\frac{(۱+e^{\alpha/\beta})(۱-p) - ۱}{e^{\alpha/\beta}((۱+e^{\alpha/\beta})p+۱)}\right)^{\beta}\right) \quad (۱۳)$$

اثبات. نتیجه فوراً با استفاده از $G(Q(p)) = p$ و (۴) با جایگزینی $F(x)$ با $\Phi(x)$ به دست می‌آید. وقتی $X \to +\infty$، حد تابع خطر لجستیک-نرمال، $+\infty$ است. نمودارهای مختلف از $h(x)$ وقتی $\mu = ۰$ و $\sigma = ۱$ باشد، برای مقادیر مختلف α و β ارائه شده است. □

بررسی شکل ۱ نشان می‌دهد که تابع خطر لجستیک-نرمال به‌طور یکنوا افزایش می‌یابد.

۴ برآورد پارامتر برای توزیع لجستیک-نرمال

فرض کنیم نمونه‌ای تصادفی به اندازهٔ n از توزیع لجستیک-نرمال گرفته شده باشد. تابع لگاریتم درست‌نمایی برای توزیع

لجستیک-نرمال در (۳) به قرار زیر است:

$$\begin{aligned} logL(\alpha,\beta,\mu,\sigma) =& \frac{n\alpha}{\beta} - nlog\beta - nlog\alpha \\ & - \frac{n}{2}log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\ & + (\frac{1}{\beta}-1)\sum_{i=1}^{n}log(1-\Phi(x_i)) \\ & - 2\sum_{i=1}^{n}log(1+e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}}) \end{aligned} \tag{14}$$

با استفاده $\frac{\partial\Phi(x)}{\partial\mu} = -\phi(x)$ و $\frac{\partial\Phi(x)}{\partial\sigma} = -(\frac{x-\mu}{\sigma})\phi(x)$، مشتق (۱۴) نسبت به α و β و μ و σ به‌صورت زیر است:

$$\frac{\partial logL}{\partial\alpha} = \frac{n}{\beta} - \frac{2}{\beta}e^{\frac{\alpha}{\beta}}\sum_{i=1}^{n}\frac{1}{1+e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}}} \tag{15}$$

$$\begin{aligned} \frac{\partial logL}{\partial\beta} =& -\frac{n}{\beta} - \frac{n\alpha}{\beta^2} - \frac{1}{\beta^2}\sum_{i=1}^{n}log(1-\Phi(x_i)) \\ & + \frac{2}{\beta^2}\sum_{i=1}^{n}\frac{(log(1-\Phi(x_i))+\alpha)e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}}}{1+e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}}} \end{aligned} \tag{16}$$

$$\begin{aligned} \frac{\partial logL}{\partial\mu} =& \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \\ & + \sum_{i=1}^{n}\frac{\phi(x_i)}{1-\Phi(x_i)}\Big\{\frac{1}{\beta}-1 \\ & - \frac{2e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}}}{\beta(1+e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}})}\Big\} \end{aligned} \tag{17}$$

$$\begin{aligned} \frac{\partial logL}{\partial\sigma} =& -\frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\ & + \frac{1}{\sigma}\sum_{i=1}^{n}\frac{(x_i-\mu)\phi(x_i)}{1-\Phi(x_i)}\Big\{\frac{1}{\beta}-1 \\ & - \frac{2e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}}}{\beta(1+e^{\frac{log(1-\Phi(x_i))+\alpha}{\beta}})}\Big\} \end{aligned} \tag{18}$$

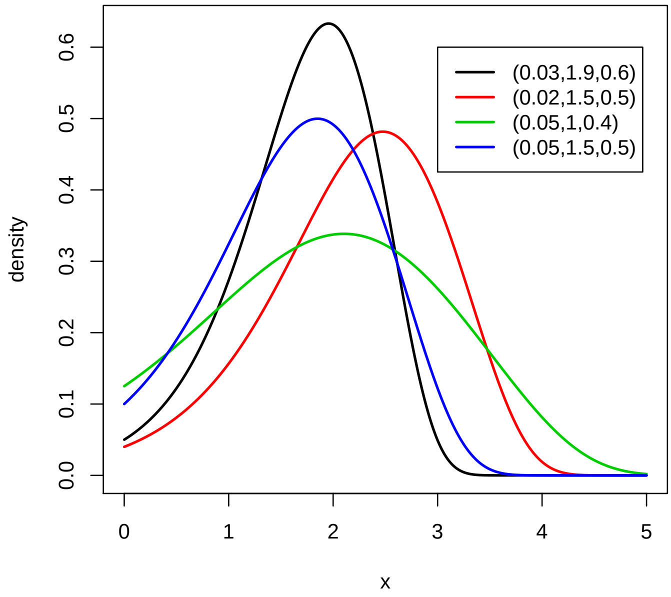
برآورد درست‌نمایی ماکسیمم $\hat{\alpha}$ و $\hat{\beta}$ و $\hat{\mu}$ و $\hat{\sigma}$، می‌تواند با قرار دادن رابطه‌های (۱۵) تا (۱۸) برابر با صفر و حل آنها به‌طور مکرر حاصل شود. مقادیر اوّلیه برای μ و σ میانگین نمونه، $\bar{x}$ و انحراف استاندارد نمونه S می‌باشد. مقادیر اوّلیه برای پارامترهای α و β با فرض کردن نمونهٔ تصادفی $y_i = -log(1-\Phi(x_i))$، $i = 1,2,3,...,n$ و حل نمودن نسبت به α و β، به‌صورت $\alpha. = \bar{y}$ و $\beta. = \frac{\sqrt{3}}{\pi}S_y$ به دست می‌آید که در آن $\bar{y}$ و S_y^2 به‌ترتیب میانگین و واریانس نمونهٔ $y_1, y_2, ..., y_n$ هستند.

۵ کاربردهای توزیع لجستیک- نرمال

در این بخش، توزیع لجستیک-نرمال برای مجموعه دادهٔ فرضی به کار می‌رود. از آن‌جا که داده‌های حقیقی در دسترس نبود، از داده‌های فرضی استفاده کرده‌ایم. این داده‌ها در جدول ۲، به‌صورت فرضی، ارائه شده‌اند. برآورد درست‌نمایی ماکسیمم، مقدار لگاریتم درست‌نمایی، معیار اطلاع آکائیکه (AIC) و معیار اطلاع بیزی (BIC) در جدول ۳ داده شده است. نتایج را برای برازش توزیع لجستیک نرمال به مجموعه‌داده‌ها در جدول ۳ گزارش کرده‌ایم و نتایج را با توزیع لجستیک مقایسه کرده‌ایم. نتایج در جدول ۳ نشان می‌دهد که توزیع لجستیک-نرمال برازش بهتری نسبت به توزیع لجستیک برای این داده‌ها دارد.

۶ خلاصه و نتیجه‌گیری

در این بررسی خواص خانوادهٔ لجستیک-X ارائه شد. حالت خاصی از خانوادهٔ لجستیک-X (لجستیک-نرمال)مطالعه شد. توزیع لجستیک-نرمال تعمیمی از توزیع نرمال است، به‌طور کلّی توزیع لجستیک-X تعمیمی از توزیع X است. ویژگی‌های مختلفی از توزیع لجستیک-نرمال از جمله میانگین انحرافات، تابع خطر و مد برررسی شد. در نهایت یک مجموعه‌داده از توزیع لجستیک-نرمال با توزیع شناخته‌شدهٔ لجستیک مقایسه شد. نتایج نشان داد که نوزیع لجستیک-نرمال برازش بهتری برای این داده‌ها ارائه می‌کند.



شکل ۱. تابع خطر لجستیک-نرمال وقتی پارامترهای نرمال استاندارد

جدول ۱. توزیع لجستیک-X برای برخی از متغیرهای تصادفی معین X

X	تابع چگالی احتمال X	تابع چگالی احتمال لجستیک-X
یکنواخت	$\frac{1}{b-a}$	$\frac{e^{\frac{log((b-X)/(b-a))+\alpha}{\beta}}}{\beta(b-X)(1+e^{\frac{log((b-X)/(b-a))+\alpha}{\beta}})^2}$
وایبل	$\frac{c}{\gamma}(\frac{X}{\gamma})^{c-1}e^{-(\frac{X}{\gamma})^c}$	$\frac{c}{\beta\gamma}(\frac{X}{\gamma})^{c-1}\frac{e^{\frac{\alpha-(\frac{X}{\gamma})^c}{\beta}}}{(1+e^{\frac{\alpha-(\frac{X}{\gamma})^c}{\beta}})^2}$
نمایی به توان رسیده	$\frac{c\lambda e^{-\lambda X}}{(1-e^{-\lambda X})^{1-c}}$	$\frac{c\lambda e^{-\lambda X}u^{c-1}e^{\frac{log(1-u^c)+\alpha}{\beta}}}{\beta(1-u)(1+e^{\frac{log(1-u^c)+\alpha}{\beta}})^2}$، $u = 1 - e^{-\lambda X}$
گامبل نوع ۲	$abX^{-a-1}e^{-bX^{-a}}$	$\frac{abX^{-a-1}(1-u)e^{\frac{log(u)+\alpha}{\beta}}}{\beta u(1+e^{\frac{log(u)+\alpha}{\beta}})^2}$، $u = 1 - e^{-bX}$
گاما	$\frac{1}{\Gamma(n)m^n}X^{n-1}e^{-X/m}$	$\frac{m^n X^{n-1}e^{-X/m}e^{(log(u)+\alpha)/\beta}}{\beta u(1+e^{(log(u)+\alpha)/\beta})^2}$، $u = 1 - \frac{\gamma(n,X/m)}{\Gamma(n)}$

جدول ۲. داده‌های فرضی

−۰/۳۴۸۶۱۱۷۲	۰/۷۸۹۲۰۳۵۶	−۰/۸۳۸۵۵۷۴۳	−۱۰/۴۵۳۷۳۸۵۵	۰/۲۸۹۵۸۵۸۴
۱/۱۶۰۹۲۷۸۴	۰/۵۳۰۷۹۲۳۹	۲/۲۶۴۸۸۰۹۳	−۰/۶۱۶۵۶۶۶۶	۰/۰۴۳۳۶۲۰۸
−۲۰/۷۶۲۱۳۳۵۸	۰/۳۸۶۵۸۷۷۴	−۰/۳۲۷۸۳۲۶۰	۰/۴۱۶۳۶۲۰۷	−۰/۰۸۰۴۰۵۰۶
۰/۱۲۱۷۶۰۵۵	−۰/۵۸۶۰۲۴۵۸	−۱/۶۵۲۲۸۳۲۹	۰/۹۲۲۵۷۶۷۰	۴/۲۵۶۸۴۵۸۶
۰/۲۹۲۸۸۸۷۷	−۰/۵۲۶۱۷۹۶۸	−۱/۰۸۵۹۸۷۴۰	−۰/۶۵۷۶۹۳۰۴	−۰/۷۵۹۰۳۱۹۳
−۰/۳۹۹۷۶۳۷۰	۲/۰۰۸۲۸۹۷۳	۰/۲۴۱۸۵۶۰۳	−۲/۴۸۵۸۸۰۲۸	−۰/۴۰۳۲۱۱۹۷
−۰/۶۸۸۴۷۴۰۶	۰/۰۶۷۸۶۱۵۱	−۱۳/۳۰۶۳۹۵۷۲	−۳۰/۹۶۵۳۶۲۲۷	−۰/۹۴۳۵۱۲۰۱
−۱/۶۸۶۲۶۳۰۰	۳/۱۴۹۲۲۸۱۱	۰/۶۲۴۴۳۹۱۳	۰/۱۴۳۲۹۴۶۴	۰/۶۶۶۹۷۶۹۸
۱/۱۶۱۸۷۷۲۷	−۰/۴۴۱۳۱۴۵۷	۲/۰۳۵۳۷۰۴۴	−۰/۳۷۰۱۴۰۱۱	−۰/۰۸۰۱۹۶۶۷
۱/۱۵۴۱۵۲۵۸	۰/۵۱۹۹۷۷۴۰	۱/۲۰۹۳۹۵۰۸	−۰/۹۸۵۸۸۵۷۰	۰/۳۶۵۰۱۱۱۵

جدول ۳. برآورد پارامترها

مدل	لجستیک	لجستیک-نرمال
برآورد پارامترها	$\begin{cases} \hat{\alpha} = -۰/۱۸۵ \\ \hat{\beta} = ۱/۸۳۵ \end{cases}$	$\begin{cases} \hat{\mu} = -۱/۳۳۳ \\ \hat{\sigma} = ۵/۷۵۶ \\ \hat{\alpha} = ۰/۸۸۳ \\ \hat{\beta} = ۰/۷۴۳ \end{cases}$
لگاریتم درست‌نمایی	۱۳۹/۸۹۷۸	۱۸۶/۲۷۹
معیاراطلاع آکائیکه	−۲۷۵/۷۹۶	−۳۶۴/۵۵۷۸
معیاراطلاع بیزی	−۲۶۴/۱۴۸	−۳۵۶/۹۱

مراجع

[1] Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, **71(1)**, 63-79.

[2] Alzaatreh, A., Lee, C. and Famoye, F. (2012). On the discrete analogues of continuous distributions. *Statistical Methodology*, **9(6)**, 589-603.

[3] Lee, C., Famoye, F. and Alzaatreh, A. Y. (2013). Methods for generating families of univariate continuous distributions in the recent decades. *Wiley Interdisciplinary Reviews: Computational Statistics*, **5(3)**, 219-238.