مجله علوم آماری، پاییز و زمستان ۱۳۹۶
جلد ۱۱، شماره ۲، ص ۲۸۵ – ۲۹۵
DOI: 10.18869/acadpub.jss.11.2.285

توزیع بتای اریب اندازه آماسیده

سید محمدرضا علوی، صفورا علی‌بابایی و رحیم چینی‌پرداز

گروه آمار، دانشگاه شهید چمران اهواز

تاریخ دریافت: ۱۳۹۵/۱۰/۱ تاریخ آخرین بازنگری: ۱۳۹۶/۷/۳

چکیده: توزیع بتا برای مدل‌بندی داده‌هایی که به‌صورت نسبت هستند، توزیع مناسبی است. در موارد زیادی که داده‌های نسبت شامل تعداد زیادی صفر و یک هستند، توزیع‌های بتای آماسیده مناسب هستند. در صورتی که احتمال ثبت چنین مشاهده‌هایی متناسب با یک تابع وزن نامنفی از آن‌ها باشد، آن مشاهده‌ها دارای توزیع موزون بتای آماسیده هستند. تمرکز این مقاله روی توزیع اریب اندازه بتای آماسیده به عنوان یک حالت خاص از توزیع موزون بتای آماسیده با وزن توانی است. خواصی از این توزیع مطالعه و پارامترهای آن به روش‌های ماکسیمم درستنمایی و گشتاوری برآورد و با استفاده از مطالعه شبیه‌سازی دو روش مقایسه می‌شوند. در پایان مدل مفروض به داده‌های واقعی نسبت مرگ و میر برازش داده می‌شوند.

واژه‌های کلیدی: توزیع بتا، توزیع بتای آماسیده، توزیع موزون، توزیع بتای آماسیده اریب اندازه.

۱ مقدمه

توزیع بتا یک توزیع پیوسته برای توصیف داده‌هایی است که در بازه (۱و۰) اندازه‌گیری می‌شوند. متغیر تصادفی Y دارای توزیع بتا با پارامترهای a و b است، اگر تابع چگالی احتمال آن به‌صورت

$$f(y;a,b)=\frac{1}{B(a,b)}y^{a-1}(1-y)^{b-1}I_{(0,1)}(y)$$

باشد، که در آن $a>0$، $b>0$، $B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ و

آدرس الکترونیکی نویسنده مسئول مقاله: سید محمدرضا علوی، alavi-m@scu.ac.ir

کد موضوع‌بندی ریاضی (۲۰۱۰): 62E10، 62E15

$$I_A(y) = \begin{cases} ۱ & y \in A \\ ۰ & y \notin A \end{cases}$$

یک تابع نشانگر است. با تغییر پارامترهای $\mu = \frac{a}{a+b}$ و $\varphi = a+b$ فراری وکریباری نتو (۲۰۰۴) توزیع بتا را به‌صورت

$$f(y;\mu,\varphi) = \frac{۱}{B(\mu\varphi,(۱-\mu)\varphi)} y^{\mu\varphi-۱}(۱-y)^{(۱-\mu)\varphi-۱} I_{(۰,۱)}(y)$$

بازنویسی کردند، که در آن $۰ < \mu < ۱$ میانگین توزیع و $\varphi > ۰$ پارامتر دقت است. این توزیع با $Beta(\mu\varphi,(۱-\mu)\varphi)$ نمایش داده می‌شود. تابع چگالی احتمال توزیع بتا به‌دلیل انعطاف پذیری زیاد و پذیرفتن شکل‌های مختلف از جمله یکنواخت، چوله به چپ و راست، نمایی، نمایی معکوس و Uشکل، توزیع پرکاربردی است. چون مقادیر توزیع بتا در بازه $(۰,۱)$ قرار دارد، برای توصیف داده‌های نسبت می‌توان از این توزیع استفاده کرد. برخی اوقات ممکن است یک نسبت، مقادیر زیادی از صفر و یک را با احتمال مثبت اختیار کند. در چنین مواردی توزیع بتا، انتخاب مناسبی نیست زیرا مجموعه $\{۰,۱\}$ در تکیه‌گاه توزیع بتا قرار ندارد. یک راه حل مناسب برای مدل‌بندی این داده‌ها، استفاده از توزیع‌های آماسیده است که توسط تیو (۲۰۰۲) معرفی شده‌اند. توزیع آماسیده بتا آمیخته‌ای از یک توزیع گسسته و یک توزیع پیوسته است. مدل‌های آماسیده در صفر برای داده‌های پیوسته، علاقه محققان را با توجه به کاربردهای تجربی به خود جلب کرده است. ایده آمیخته کردن یک توزیع تباهیده در یک نقطه با یک توزیع پیوسته، توسط اتکیسون (۱۹۵۵)، با معرفی آمیخته‌ای از یک توزیع تباهیده در صفر و یک توزیع لگ‌نرمال به نام توزیع دلتا، مطرح شد. کوک و همکاران (۲۰۰۶) توزیع بتای آماسیده در صفر را بررسی کردند. کوک و همکاران (۲۰۰۸) یک مدل رگرسیونی برای داده‌های منابع مالی پیشنهاد دادند، که در آن متغیر پاسخ دارای توزیع بتای آماسیده در صفر است. میانگین این توزیع یک پیش‌بین خطی با استفاده از تابع پیوند لوجیت است. هوف (۲۰۰۷) این کار را برای داده‌های نسبت با توزیع بتا آماسیده در یک انجام داد. اُسپینا و فراری (۲۰۱۰) توزیع بتای آماسیده در یک و صفر را پیشنهاد دادند که تعمیم دو مدل فوق است.

متغیر تصادفی Y دارای توزیع بتای آماسیده در صفر و یک است، اگر تابع چگالی آن به‌صورت

$$bic(y;\alpha,\mu,\varphi) = \begin{cases} \alpha & y = c \\ (۱-\alpha)f(y;\mu,\varphi) & y \in (۰,۱) \end{cases}$$

باشد، که در آن $c \in \{0, 1\}$ و $0 < \alpha < 1$ پارامتر آمیختگی است. اگر $c = 0$ توزیع بتای آماسیده در صفر و به‌ازای $c = 1$ توزیع بتای آماسیده در یک حاصل می‌شود. این توزیع با نماد $BIC(\alpha, \mu, \varphi)$ نمایش داده می‌شود (اُسپینا و فراری، ۲۰۱۰).

در صورتی که احتمال ثبت مشاهده‌ها از توزیعی، متناسب با یک تابع نامنفی از آن‌ها باشد، چنین مشاهده‌هایی یک نمونه موزون از آن توزیع هستند. نمونه موزون در حقیقت یک نمونه تصادفی از یک توزیع جدید است که به آن توزیع موزون متناظر با آن توزیع گفته می‌شود. فرض کنید متغیر تصادفی Y دارای تابع چگالی احتمال $f(y;\theta)$ باشد. اگر ساز و کار نمونه‌گیری به‌گونه‌ای باشد که مشاهده y با احتمالی متناسب با یک تابع نامنفی $w(y,\omega)$ ثبت شود، در آن‌صورت مشاهده ثبت شده یک مقدار از نسخه موزون Y است که با Y_w نمایش داده می‌شود. تابع چگالی احتمال Y_w به‌صورت

$$f_w(y;\theta,\omega) = \frac{w(y,\omega) f(y;\theta)}{E_\theta[w(Y,\omega)]}$$

است که در آن $E_\theta[w(Y,\omega)] = \int_{-\infty}^{\infty} w(y,\omega) f(y;\theta) dy$ ضریب نرمال ساز است. به توزیع Y_w توزیع موزون Y تحت وزن $w(y,\omega)$به ساز و کار نمونه‌گیری، نمونه‌گیری موزون [1] تحت وزن $w(y,\omega)$ گفته می‌شود (رائو، ۱۹۶۵). پارامتر ω، پارامتر وزن نام دارد. اگر تابع وزن ثابت باشد یعنی $w(y,\omega) = k$، توزیع های اصلی و موزون یکسان هستند و نمونه‌گیری موزون منجر به نمونه تصادفی می‌شود. اگر $w(y,\omega) = y^\omega$، توزیع موزون را توزیع اریب اندازه[2]و نمونه گیری را اریب اندازه مرتبه ωگویند. گاهی نمونه‌گیری اریب اندازه مرتبه اول را اریب طول نیز می‌گویند (کاکس، ۱۹۶۲). ایده اصلی توزیع موزون توسط فیشر (۱۹۳۴) هنگامی که اثر روش‌های تعیین داده‌ها را روی برآورد پارامترهای جامعه مطالعه می‌کرد، شکل گرفت و کاکس (۱۹۶۲) حالت خاصی از نمونه گیری موزون را تحت نام نمونه‌گیری اریب طول یا اریب اندازه به کار برد. رائو (۱۹۶۵) اولین کسی است که به طور رسمی توزیع های موزون را معرفی کرد. محققین زیادی توزیع‌های موزون را در زمینه‌های مختلف مطالعه کرده‌اند. گوپتا و کوندا (۲۰۰۹)، علوی و چینی پرداز (۲۰۰۹)، چاکرابراتی (۲۰۱۰)، کریمی و علوی (۲۰۱۴) و علوی (۲۰۱۶) از توزیع های موزون استفاده کرده اند. تمرکز اصلی این مقاله بر توزیع اریب اندازه بتای آماسیده در صفر و یک است.

در بخش ۲ توزیع بتای اریب اندازه آماسیده معرفی می‌شود و مشخصه های توزیع، تابع امتیاز، ماتریس اطلاع فیشر بیان و پارامترهای مدل برآورد می‌شوند. هنگامی که شکل بسته برای برآوردها به‌دست نمی‌آید،

[1] Weighted sampling
[2] Size-biased

با روش‌های عددی برآوردها محاسبه می‌شوند. در بخش ۳ با استفاده از شبیه سازی به ارزیابی این برآوردها پرداخته می‌شود. در بخش ۴ داده‌های نسبت مرگ و میر کودکان پرورشگاه‌های واجد شرایط در مناطق شهری برزیل به این توزیع برازش داده می‌شوند. در نهایت در بخش ۵ ارائه نتایج بیان می‌شوند.

۲ توزیع بتای اریب اندازه آماسیده در صفر و یک

اگر توزیع جامعه اصلی $Beta(\mu\varphi, (1-\mu)\varphi)$ باشد، توزیع اریب اندازه مرتبه ω آن با نماد $SBETA(\omega)$ نشان داده می‌شود و تابع چگالی احتمال آن به‌صورت

$$f_w(y;\mu,\varphi,\omega) = \frac{y^{\omega}}{\mu_\omega B(\mu\varphi,(1-\mu)\varphi)} y^{\mu\varphi-1}(1-y)^{(1-\mu)\varphi-1} I_{(0,1)}(y) \qquad (1)$$

است، که در آن $\mu_\omega = \frac{B(\mu\varphi+\omega,(1-\mu)\varphi)}{B(\mu\varphi,(1-\mu)\varphi)}$ گشتاور غیرمرکزی مرتبه ω است. بنابراین

$$f_w(y;\mu,\varphi,\omega) = \frac{1}{B(\mu\varphi+\omega,(1-\mu)\varphi)} y^{\mu\varphi+\omega-1}(1-y)^{(1-\mu)\varphi-1} I_{(0,1)}(y)$$

در نتیجه توزیع بتای اریب اندازه مرتبه ω نیز یک توزیع بتاست یعنی دارای خاصیت پایایی است (علوی و چینی پرداز، ۲۰۰۹). میانگین و واریانس توزیع بتای اریب اندازه مرتبه ω به‌ترتیب به‌صورت $\mu_w = \frac{\mu\varphi+\omega}{\varphi+\omega}$ و $\sigma_w^2 = \frac{(\mu\varphi+\omega)(1-\mu)\varphi}{(\varphi+\omega)^2(\varphi+\omega+1)}$ هستند. وقتی $\omega = 0$ باشد توزیع‌های اصلی و موزون یکی هستند. اگر $\omega = 1$ باشد توزیع موزون را توزیع بتای اریب طول می‌نامند و با نماد $LBETA$ نمایش داده می‌شود. اگر $SBETA(\omega)$ در صفر و یک آماسیده شود توزیع حاصل بتای اریب اندازه آماسیده در صفر و یک نامیده و با $SBIC(\omega)$ نمایش داده می‌شود. تابع چگالی احتمال آن به‌صورت

$$sbic(y;\alpha,\mu,\varphi) = \begin{cases} \alpha & y = c \\ (1-\alpha) f_w(y;\alpha,\mu,\varphi) & y \in (0,1) \end{cases} \qquad (2)$$

است، که در آن $c \in \{۰, ۱\}$. این توزیع، آمیخته‌ای از یک توزیع بتای اریب اندازه و یک توزیع تباهیده در c است. میانگین توزیع بتای اریب اندازه آماسیده در صفر و یک برابر

$$E(Y_w) = \alpha c + (۱ - \alpha)\mu_w$$

است، که در آن μ_w میانگین توزیع بتای اریب اندازه است. گشتاور غیرمرکزی مرتبه دوم برابر

$$E(Y_w^۲) = \alpha c + (۱ - \alpha)\mu_۲^w = \alpha c + (۱ - \alpha)\frac{(\mu\varphi + \omega)(\mu\varphi + \omega + ۱)}{(\varphi + \omega)(\varphi + \omega + ۱)}$$

است، که در آن $\mu_۲^w$ گشتاور غیرمرکزی مرتبه دو توزیع بتای اریب اندازه است. در نتیجه

$$Var(Y_w) = E(Var(Y_w|I_{\{c\}}(Y_w))) + Var(E(Y_w|I_{\{c\}}(Y_w)))$$

که در آن

$$E(Var(Y_w|I_{\{c\}}(Y_w))) = (۱ - \alpha)\tfrac{(\mu\varphi+\omega)(۱-\mu)\varphi}{(\varphi+\omega)^۲(\varphi+\omega+۱)}$$

$$\begin{aligned} Var(E(Y_w|I_{\{c\}}(Y_w))) &= c^۲\alpha + \mu_w^۲(۱ - \alpha) - (\alpha c + \mu_w(۱ - \alpha))^۲ \\ &= \alpha(۱ - \alpha)(c - \mu_w)^۲ \end{aligned}$$

بنابراین

$$Var(Y_w) = (۱ - \alpha)\sigma_w^۲ - \alpha(۱ - \alpha)(c - \mu_w)^۲.$$

تابع درستنمایی برای برآورد $\theta = (\alpha, \mu, \varphi)$ براساس نمونه موزون به‌صورت

$$L_w(\theta) = \Pi_{i=۱}^n sbic(y; \alpha, \mu, \varphi) = L_{۱w}(\alpha)L_{۲w}(\mu, \varphi)$$

است، که در آن

$$L_{1w}(\alpha) = \Pi_{i=1}^{n} \alpha^{I_{\{c\}}(y_i)} (1-\alpha)^{1-I_{\{c\}}(y_i)} = \alpha^{T_1}(1-\alpha)^{n-T_1}$$
$$L_{2w}(\mu,\varphi) = \Pi_{i=1}^{n} f_w(y_i;\mu,\varphi)^{1-I_{\{c\}}(y_i)}$$

لگاریتم تابع درستنمایی توزیع بتای اریب اندازه آماسیده در صفر و یک به‌صورت

$$\ell_w(\theta) = \log L_w(\theta) = \ell_{1w}(\alpha) + \ell_{2w}(\mu,\varphi)$$

است، که در آن $\ell_{1w}(\alpha) = T_1 \log\alpha + T_2 \log(1-\alpha)$ و

$$\begin{aligned}\ell_{2w}(\mu,\varphi) &= T_2 \log\{\frac{\Gamma(\varphi+\omega)}{\Gamma(\mu\varphi+\omega)\Gamma((1-\mu)\varphi)}\} + T_3(\mu\varphi+\omega-1)\\ &\quad + T_4((1-\mu)\varphi)-1)\end{aligned}$$

که T_1 تعداد مقادیر c، $T_2 = n - T_1$ تعداد مقادیر مخالف c در نمونه، $T_3 = \sum_{i,y_i\in(0,1)} \log(y_i)$ و $T_4 = \sum_{i,y_i\in(0,1)} \log(1-y_i)$ است. تابع‌های امتیاز برای α، μ و φ عبارتند از:

$$\begin{aligned}U_\alpha(\alpha) &= \frac{T_1}{\alpha} + \frac{n-T_1}{1-\alpha}\\ U_\mu(\mu,\varphi) &= -\varphi\{(n-T_1)[\Psi((1-\mu)\varphi) - \Psi(\mu\varphi+\omega)] + T_3 - T_4\}\\ U_\varphi(\mu,\varphi) &= -(n-T_1)\{\Psi(\varphi+\omega) - \mu\Psi(\mu\varphi+\omega)\\ &\quad -(1-\mu)\Psi((1-\mu)\varphi)\} - \mu T_3 + (1-\mu)T_4T_3 + (1-\mu)T_4 \qquad (3)\end{aligned}$$

ماتریس اطلاع فیشر برای توزیع بتای اریب اندازه آماسیده در صفر و یک به‌صورت

$$K(\theta) = \begin{pmatrix} k_{\alpha\alpha} & 0 & 0 \\ 0 & k_{\mu\mu} & k_{\mu\varphi} \\ 0 & k_{\varphi\mu} & k_{\varphi\varphi} \end{pmatrix}$$

است، که در آن مؤلفه‌های آن عبارتند از :

$$
\begin{aligned}
k_{\alpha\alpha} &= \frac{n}{\alpha(1-\alpha)} \\
k_{\mu\mu} &= n(1-\alpha)\varphi^{2}\{\Psi'(\mu\varphi+\omega)+\Psi'((1-\mu)\varphi)\} \\
k_{\varphi\varphi} &= n(1-\alpha)\{\mu^{2}\Psi'(\mu\varphi+\omega)+(1-\mu)^{2}\Psi'((1-\mu)\varphi)-\Psi'(\varphi+\omega)\} \\
k_{\mu\varphi} &= n(1-\alpha)\varphi\{\mu\Psi'(\mu\varphi+\omega)-(1-\mu)\Psi'((1-\mu)\varphi)\}
\end{aligned}
$$

که در آن‌ها

$$\Psi'(z)=\frac{d^{2}\log\Gamma(z)}{dz^{2}}=\frac{d\Psi(z)}{dz}$$

تابع تری گاما است و $\Psi(z)=\frac{\Gamma'(z)}{\Gamma(z)}$. برآورد ماکسیمم درستنمایی α برابر $\widehat{\alpha}=\frac{T_1}{n}$ است. برآورد ماکسیمم درستنمایی μ و φ از حل دو معادله $U_\mu(\mu,\varphi)$ و $U_\varphi(\mu,\varphi)$ به‌روش عددی با استفاده از الگوریتم $BFGS$ به‌دست می‌آیند. این الگوریتم نوعی از روش های موسوم به شبه نیوتن[3] است که بهترین عملکرد را بین روش‌های بهینه‌سازی غیرخطی دارد (میتلهامر و همکاران، ۲۰۰۰).

برآورد گشتاوری (MM) پارامترهای μ و φ توزیع بتای اریب اندازه آماسیده در صفر و یک به‌صورت

$$\tilde{\mu}=\frac{s^{2}(\bar{y}-\omega)}{\bar{y}(1-\bar{y})-(\omega+1)s^{2}},\quad \tilde{\varphi}=\frac{\bar{y}(1-\bar{y})}{s^{2}}-(\omega+1)$$

هستند، که در آن $\bar{y}=\sum_{i,y_i\in(0,1)}\frac{y_i}{n-T_1}$ و $s^{2}=\sum_{i,y_i\in(0,1)}\frac{(y_i-\bar{y})^{2}}{n-T_1}$ به ترتیب میانگین و واریانس مقادیری از نمونه هستند که صفر یا یک نیستند.

۳ مطالعه شبیه‌سازی

باتوجه به این‌که برای برآورد ماکسیمم درستنمایی μ و φ شکل بسته‌ای حاصل نمی‌شود، برای ارزیابی برآوردهای ماکسیمم درستنمایی با استفاده از نرم افزارR، شبیه سازی مونت کارلو به‌کار برده می‌شود. بدون از دست دادن کلیت مسئله مرتبه اریب اندازه یک در نظر گرفته می‌شود. از توزیع بتای اریب طول آماسیده در صفر ($LBEZI$) به‌ازای $\alpha=0/1$، $\mu=0/1$ و $\varphi=2$ نمونه‌هایی به حجم n ۱۰۰۰ بار تولید

[3] Quasi-Newton algorithm

نموده، برآوردهای ماکسیمم درستنمایی در هر بار با روش $BFGS$ به‌دست آورده شده‌است. میانگین و میانگین توان دوم خطای برآوردها در ۱۰۰۰ بار تکرار به‌عنوان اُمید ریاضی و میانگین توان دوم خطای برآوردکننده‌ها محاسبه شده‌اند. همچنین با روشی مشابه، برآورد به روش گشتاوری نیز محاسبه گردیده‌اند. جدول ۱ نتایج شبیه‌سازی پارامترهای توزیع بتای اریب طول آماسیده در صفر را برای $\alpha = ۰/۱$، $\mu = ۰/۱$ و $\varphi = ۲$ به ازای مقادیر مختلف حجم نمونه $n = ۱۰, ۲۰, ۵۰, ۱۰۰, ۵۰۰, ۱۰۰۰$ نشان می‌دهد. از آنجا که برآورد ماکسیمم درستنمایی و گشتاوری α یکسان و نااریب با واریانس $V(\hat{\alpha}) = \frac{\alpha(۱-\alpha)}{n}$ است در نتایج شبیه‌سازی این پارامتر آورده نشده است. با ملاحظه جدول ۱ اریبی برآورد ماکسیمم درستنمایی μ با افزایش حجم نمونه به صفر میل می‌کند. همچنین ریشه میانگین توان دوم خطای برآورد ماکسیمم درستنمایی μ با افزایش حجم نمونه از ریشه میانگین توان دوم خطای برآورد گشتاوری μ کمتر است که بیانگر آن است که روش ماکسیمم درستنمایی کاراتر است. با این حال برای نمونه های کوچک، برآوردهای ماکسیمم درستنمایی و گشتاوری φ به‌طور قابل توجهی اریب هستند. اریبی برآورد ماکسیمم درستنمایی به جز برای $n = ۱۰$ کمتر از اریبی برآورد گشتاوری است. با توجه به ریشه میانگین توان دوم خطای برآورد φ در دو روش، برآورد روش ماکسیمم درستنمایی بهتر از برآورد روش گشتاوری است.

۴ مثال کاربردی

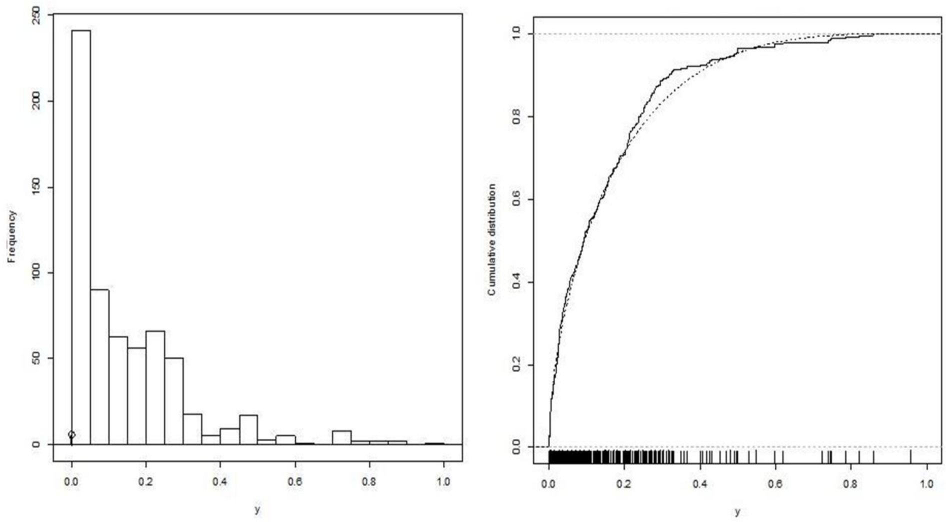
یک مجموعه داده از شاخص‌های اجتماعی در پرورشگاه های برزیل شامل ۵۵۶۱ مشاهده است که هر داده بیانگر نسبت مرگ و میر کودکان کمتر از یک سال پرورشگاه در سال ۲۰۰۰ است که توسط عوامل نامعلوم می‌میرند. داده‌ها در سایت $www.datasus.org.br$موجود است. یک نمونه اریب طول از آنها انتخاب شد. بافت‌نگار داده‌های نمونه اریب طول در شکل ۱ نمایش داده شده است. برای این مجموعه داده‌ها از توزیع بتای اریب طول آماسیده در صفر با پارامترهای (α, μ, φ) استفاده شده است. برآوردهای ماکسیمم درستنمایی به‌صورت $\hat{\alpha} = ۰/۰۰۹۳(۰/۰۰۷۲)$، $\hat{\mu} = ۰/۱۵۱۷(۰/۰۴۶)$ و $\hat{\varphi} = ۴/۰۰۵۰(۰/۰۰۵۰)$ هستند که اعداد داخل پرانتز برآورد ریشه میانگین توان دوم خطای برآورد را نشان می‌دهند دهند. توزیع تجربی (خط ممتد) و توزیع تجمعی برازشی (نقطه چین) داده ها در شکل ۱ بیانگر آن است که توزیع بتای اریب طول آماسیده در صفر مدل مناسب برای داده‌ها است.

جدول ۱. اریبی و جذر میانگین توان دوم خطای برآورد پارامترهای توزیع $LBEZI$ برای $\alpha = ۰/۱$

پارامتر	اندازه نمونه	میانگین		اریبی		جذر میانگین توان دوم خطا	
		گشتاوری	درستنمایی	گشتاوری	درستنمایی	گشتاوری	درستنمایی
$\mu = ۰/۱$	۱۰	−۰/۱۵۵۹	۰/۱۵۹۴	−۰/۲۵۶۰	۰/۰۵۹۴	۶/۰۷۹۸	۰/۱۴۳۹
	۲۰	۰/۰۸۶۴	۰/۰۸۶۴	۰/۱۳۷۱	−۰/۰۱۳۶	۰/۰۳۷۱	۰/۲۳۸۷
	۵۰	۰/۰۹۵۰	۰/۱۱۰۹	−۰/۰۰۵۰	۰/۰۱۰۹	۰/۱۰۸۸	۰/۰۷۸۵
	۱۰۰	۰/۰۹۵۲	۰/۱۰۲۴	−۰/۰۰۴۷	۰/۰۰۲۴	۰/۰۷۶۴	۰/۰۵۹۸
	۵۰۰	۰/۰۹۹۶	۰/۰۹۹۹	−۰/۰۰۰۴	−۰/۰۰۰۱	۰/۰۳۲۵	۰/۰۳۱۲
	۱۰۰۰	۰/۰۹۸۵	۰/۰۹۸۹	−۰/۰۰۱۴	−۰/۰۰۱۱	۰/۰۲۳۴	۰/۰۲۱۴
$\varphi = ۲$	۱۰	۳/۱۶۴۴	۳/۳۳۵۰	۱/۱۶۴۴	۱/۳۳۵۰	۳/۳۹۲۵	۳/۲۸۶۱
	۲۰	۲/۴۹۶۲	۲/۶۲۵۰	۰/۴۹۶۲	۰/۶۲۵۰	۱/۴۰۴۱	۱/۴۰۴۶
	۵۰	۲/۱۴۴۶	۲/۱۸۶۵	۰/۱۴۴۶	۰/۱۸۶۵	۰/۶۸۲۰	۰/۶۱۹۲
	۱۰۰	۲/۰۷۸۳	۲/۰۹۶۲	۰/۰۷۸۳	۰/۰۹۶۲	۰/۴۷۶۲	۰/۴۲۵۷
	۵۰۰	۲/۰۱۸۰	۲/۰۱۸۲	۰/۰۱۸۰	۰/۰۱۸۲	۰/۲۰۸۷	۰/۱۹۹۶
	۱۰۰۰	۲/۰۰۳۰	۲/۰۰۳۸	۰/۰۰۳۰	۰/۰۰۳۸	۰/۱۵۲۷	۰/۱۴۴۰

بحث و نتیجه‌گیری

در این مقاله توزیع بتای اریب اندازه آماسیده در صفر و یک معرفی و میانگین، واریانس و خصوصیات دیگری از آن ارائه شد. برآورد ماکسیمم درستنمایی و گشتاوری پارامترهای با استفاده از شبیه سازی مقایسه شدند. توزیع بتای اریب طول آماسیده به یک نمونه اریب طول از داده های نسبت مرگ و میر کودکان پرورشگاه‌های واجد شرایط در مناطق شهری برزیل برازش داده شد و برآورد پارامترهای مدل به روش ماکسیمم درستنمایی $\hat{\alpha} = ۰/۰۰۹۳(۰/۰۰۷۲)$ ، $\hat{\mu} = ۰/۱۵۱۷(۰/۰۴۶)$ و $\hat{\varphi} = ۴/۰۰۵۰(۰/۰۰۵۰)$ برآورد شدند که اعداد داخل پرانتز خطای معیار برآوردها را نشان می‌دهند.



(ب) (الف)

شکل ۱. نمودار الف توزیع تجمعی داده‌ها (خط ممتد) و توزیع برازشی بتای اریب طول (خط چین) و نمودار ب بافت‌نگار داده‌ها

تقدیر و تشکر

نویسندگان از پیشنهادهای ارزنده داوران ، هیئت تحریریه و ویراستار محترم مجله علوم آماری که باعث ارتقای مقاله شد، کمال تشکر و قدردانی را دارند.

مراجع

Cressie, N. (1993), *Statistics for Spatial Data*, Revised Edition, John Wiley, New York.

Aitchison, J. (1955), On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin, *Journal of the American Statistical Association*, **50**, 901–908.

Alavi, S. M. R. (2016), A Generalized Class of From-Invariant Bivariate Weighted Distributions, *Communications in Statistics-Theory and Methods*, **46**, 2193–2201.

Alavi, S. M. R. and Chinipardaz, R. (2009), Form-Invariance under Weighted Sampling , *Statistics*, **43**, 81-90.

Chakraborty, S. (2010), On Some Distributional Properties of the Family of Weighted Generalized Poisson Distribution, *Communications in Statistics-Theory and Methods*, **39**, 2767-2788.

Cook, D. O., Kieschnick, R., McCullough, B. D. (2006), On the Heterogeneity of Corporate Capital Sructures and its Implications, *Disponivel em http://ssrn.com/abstract=671061.*

Cook, D. O., Kieschnick, R., McCullough, B. D. (2008), Regression Analysis of Proportions in Finance with Self Selection, *Journal of Empirical Finance*, **15**, 860–867.

Cox, D. R. (1962),*Renewal Theory*, Methuen, London.

Ferrari, S. L. P., Cribari-Neto, F. (2004), Beta Regression for Modeling Rates and Proportions, *Journal of Applied Statistics*, **3**, 799–815.

Fisher, R. A. (1934), The Effects of the Methods of Ascertainment Upon the Estimation of Frequencies, *Annal Eugen*, **6**, 13-26.

Gupta, R. D. and Kundu, D. (2009), A New Class of Weighted Exponential Distributions, *Statistics*, **43**, 621-634.

Hoff, A. (2007), Second Stage DEA: Comparison of Approaches for Modeling the DEA Score, *European Journal of Operational Research*, **181**, 425–435.

Karimi, M. and Alavi, S. M. R. (2014), The Effect of Weight Function on Hypothesis Testing in Weighted Sampling, *Journal of Applied Statistics*, **41**, 2493-2503.

Mittelhammer, R. C., Judge, G. G. and Miller, D.J. (2000), *Econometric Foundations*, Cambridge University Press, New York.

Ospina, R. and Ferrari, S.L.P. (2010). Inflated Beta Distributions, *Statistical Papers*, **51**, 111–126.

Rao, C. R. (1965), *On Discrete Distribution Arising out of Methods of Ascertainment. Classical and Contagious Distribution.; 320-332, in G. P. Patil (ed.)*, Statistical Publishing Society, Calcutta.

Tu, W. (2002), Zero Inflated Data, *Encyclopedia of Environmetrics*, **4**, 2387–2391.