

توزیع‌های چوله و کاربرد آن در تحلیل داده‌پزشکی کارخانه فولاد مبارکه اصفهان

محمد غلامی فشارکی^۱، انوشیروان کاظم نژاد^۲، فرید زایری^۳، محسن روضاتی^۴

چکیده:

امروزه استفاده از توزیع‌های منعطف مانند توزیع‌های چوله که توانایی بیشتری را در بیان رفتار داده‌های مشاهده شده دارند افزایش چشمگیری یافته است. این توزیع‌ها، بیشتر در علوم پزشکی و علوم رفتاری که مقادیر واقعی متغیرهای تصادفی دارای توزیعی غیر متقاضی است مورد استفاده قرار می‌گیرد. با توجه به کاربرد زیاد توزیع‌های چوله، در این مقاله بعد از مروری مختصر در مورد توزیع‌های چوله معروف، توزیع‌های نرمال، تی، نرمال و تی چوله به داده‌های استخراج شده از داده‌های پزشکی شرکت فولاد مبارکه برآنش داده شده و سپس با استفاده از شاخص AIC بهترین مدل، از بین مدل‌های موجود مورد انتخاب قرار می‌گیرد.

واژه‌های کلیدی: روش بررسی پنهان، توزیع‌های نرمال چوله-مستقل، نرمال چوله، تی چوله.

۱ مقدمه

و هسو (۲۰۰۵)؛ جارا، کیونتنا و سن مارتین (۲۰۰۸)، پنهان شدن فرآیند تولید داده‌ها و کاهش اطلاعات (جارا و همکاران (۲۰۰۸)) و یکتا بودن تبدیل در مطالعات مختلف می‌گردد (برای مثال عدم محاسبه تبدیل لگاریتمی در داده‌هایی با دامنه منفی) (بندیوپدهایی، لاصوس، آبانتولی و قوش (۲۰۱۰)). استفاده از روش توزیع‌های غیر نرمال نیز علی‌رغم اینکه مشکلات قبلی را تا حدود زیادی مرتفع می‌کند اما محققان را از دسترسی به خصوصیات مناسب توزیع‌های بیضوی مانند نرمال دور می‌سازد. روش دیگر در حل این معضل به کار بردن توزیع‌هایی است که دارای خصوصیات بیضوی بوده و از لحاظ نظری قادر به تبیین تغییرات مشاهدات بدون استفاده از فرض نرمالیتی هستند. به این ترتیب استفاده از خانواده‌هایی از توزیع‌های نامتقاضی که قابلیت

یکی از فرض‌های اساسی در بسیاری از مدل‌سازی‌های آماری فرضیه نرمال بودن توزیع مشاهدات است. از طرفی این فرضیه در بسیاری از حالات برقرار نیست، از این رو محققان جهت بررسی‌های خود به روش‌هایی چون تبدیل داده‌ها (به ویژه استفاده از تبدیل کاکس-باکس)، استفاده از روش‌های ناپارامتری و استفاده از توزیع‌های غیر نرمال دست می‌زنند. استفاده از روش‌های ناپارامتری علی‌رغم کاهش اطلاعات از لحاظ اجرایی دارای محدودیت‌های زیادی برای محققان بوده و از طرف دیگر استفاده از روش تبدیل متغیرها نیز باعث مشکلاتی چون تفسیر نامناسب پارامترهای تحت بررسی، احراز نشدن فرض نرمالیتی توأم متغیرها در حالات چند متغیره (تسونگ و لی (۲۰۰۳)؛ لی، لین، لی

^۱ گروه آمار زیستی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، mohammad.gholami@modares.ac.ir

^۲ گروه آمار زیستی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس، Kazem_an@modares.ac.ir

^۳ گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی.

^۴ پزشک واحد پیگیری بیماری‌ها، بخش طب صنعتی، شرکت فولاد مبارکه اصفهان، مبارکه.

در این معادله در صورتی که $g(x) = h(x)$ انتخاب گردد، خانواده توزیع‌های چوله-متقارن^۶ (گپتا، چانگ و هوانگ، ۲۰۰۲) به دست می‌آید. که از آن جمله می‌توان به توزیع نرمال چوله، تی چوله، یکنواخت چوله، لاپلاس چوله و لجستیک چوله اشاره نمود. از این میان، توزیع نرمال چوله از جمله اولین، همانند و خوش تعریف‌ترین توزیع‌ها نسبت به توزیع نرمال است که خصوصیات آن بصورت وسیع در حالت تک متغیره توسط آزالینی (۱۹۸۵) و در حالت چند متغیره توسط آزالینی و دال ول (۱۹۹۶) مورد بررسی قرار گرفته است. اما آزالینی نشان داده است که میزان پارامتر چولگی این توزیع در بازه $(0, 0.995)$ تغییر می‌کند و این بدان معنی است که از توزیع نرمال چوله تنها در موقعي می‌توان استفاده نمود که چولگی داده‌ها زیاد نباشد. این موضوع را از دیدگاه دیگری نیز می‌توان توضیح داد. در صورتی که پارامتر چولگی λ از مقدار 5 بیشتر شود، توزیع نرمال چوله به سرعت به توزیع نیم نرمال^۷ میل نموده و عملاً برای مقادیر بالای چولگی (از مقدار 5 به بالا) امکان تشخیص توزیع نرمال چوله بسیار دشوار می‌گردد (اسپینوسا، ۲۰۰۸). همچنین از معايب دیگر توزیع نرمال چوله می‌توان به برآورده بی نهایت پارامتر اشاره نمود (جتنون، ۲۰۰۴؛ دی، ۲۰۱۰). البته محققان مختلف به ارائه راهکارهایی جهت حل این مشکل پرداخته‌اند (پوسی، ۲۰۰۰).

در صورتی که $g(x) = \varphi$ (که φ تابع چگالی نرمال استاندارد است) و تابع $H(\cdot)$ دلخواه انتخاب شود خانواده توزیع‌های چوله نرمال-متقارن^۸ (نادریج و کوتر، ۲۰۰۳) به دست می‌آید در صورتی که $H(x) = \Phi(x)$ (که $\Phi(\cdot)$ تابع توزیع نرمال استاندارد است) و $g(\cdot)$ دلخواه انتخاب گردد، خانواده توزیع‌های متقارن-نرمال^۹ (گپتا و همکاران (۲۰۰۲)، ونگاس و بولفارین، ۲۰۰۷) حاصل می‌شود. گپتا و همکاران (۲۰۰۲) با انتخاب توزیع‌های لاپلاس، یکنواخت و لجستیک نشان دادند که خانواده توزیع‌های چوله نرمال-متقارن برخلاف خانواده توزیع‌های نرمال چوله توانایی پوشش میزان چولگی بیشتری را دارا هستند، اما یکی از مشکلات در استفاده از آنها سنگینی محاسبات توزیع‌های این خانواده‌ها

تبیین چولگی و کشیدگی داده‌ها را دارا باشد، برای مدل بندي داده‌ها و تحلیل‌های آماری نیرومند نسبت به تخطی از فرض نرمالیتی می‌تواند مفید و کارساز باشد. از این رو امروزه استفاده از تحلیل‌های رگرسیونی و مدل‌های خطی با استفاده از توزیع‌ی بیضوی چوله به طور قابل ملاحظه‌ای مورد توجه پژوهشگران قرار گرفته است (آزالینی و کاپتانیو (۱۹۹۹)؛ فرناندر و استیل (۱۹۹۸)؛ ساهو، دی و برانکو (۲۰۰۳)). از این رو با توجه به ضرورت استفاده از توزیع‌های چوله در مدل‌سازی و علی الخصوص تحلیل داده‌های پزشکی، در این مقاله برآن شدیم تا بصورت مختصر به مروری در مورد توزیع‌های چوله پرداخته (بخش ۲) و سپس با استفاده از یکسری داده‌های پزشکی با درجات مختلف چولگی و کشیدگی (بخش ۳) به برآش $\mathbf{4}$ مدل نرمال، تی، نرمال چوله و بالاخره تی چوله (بخش ۴) بپردازیم و درنهایت با استفاده از نتایج حاصل، به بینشی نو در استفاده از توزیع‌های چوله دست یابیم.

۲ مروری بر فرآیندهای تولید توزیع‌های چوله

جهت ایجاد توزیع‌های چوله روش‌های متعددی وجود دارد. یکی از متداول ترین این روش‌ها استفاده از روش برشی پنهان^{۱۰} (معادله ۱) است. (ابطحی، توحیدی و بهبودیان، ۲۰۰۹) آزالینی و کاپتانیو (۲۰۰۳) و آزالینی (۲۰۰۵) ثابت کردند در صورتی که $w(x)$ یک تابع فرد و $H(\cdot)$ یک تابع توزیع با تابع چگالی متقارن حول صفر (\cdot) و $g(x)$ یک تابع چگالی متقارن باشد، آنگاه تابع چگالی $f_x(X)$ که به فرم معادله (۱) تعریف می‌شود، یک تابع چگالی نامتقارن خواهد بود.

$$f_x(X) = \frac{1}{2} g(x) H(w(x)), \quad (1)$$

$$-\infty < x < \infty.$$

در صورتی که $w(x) = \lambda x$ انتخاب شود، معادله (۱) به فرم معادله (۲) در می‌آید.

$$f_x(X) = \frac{1}{2} g(x) H(\lambda x), \quad (2)$$

$$-\infty < x < \infty.$$

⁵Hidden Truncation Method

⁶Skew Symmetric

⁷Half Normal

⁸Skew Normal Symmetric

⁹Skew Symmetric Normal

نموده‌اند. همچنین خردمندی و همکاران(۲۰۱۰) با استفاده از تبدیل توزیع نرمال به تی با v درجه آزادی در معادلات (۳-۵) به جای تابع چگالی نرمال به ارائه توزیع‌های تعمیم یافته جدید دست یافتند (معادلات ۶-۹).

$$f_{\lambda_1, \lambda_2, v}(x) = 2t_v(x) \Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \quad (6)$$

$$-\infty < x < \infty, \quad \lambda_2 \geq 0,$$

$$f_{\lambda, \gamma, v}(x) = 2t_v(x)\Phi(\lambda x + \gamma x^3), \quad (7)$$

$$-\infty < x < \infty,$$

$$f_{\lambda, n}(x) = 2c_n(\lambda) \varphi(x)\Phi^n(\lambda x), \quad (8)$$

$$c_n(\lambda) = \left(\int_{-\infty}^{+\infty} 2\varphi(x)\Phi^n(\lambda x) dx \right)^{-1},$$

$$f_{\lambda_1, \lambda_2}(x) = 2c_{n, \lambda_1, \lambda_2}(\lambda) \times \varphi(x)\Phi^n\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \quad (9)$$

$$-\infty < x, \quad \lambda_1 < \infty, \quad \lambda_2 \geq 0,$$

از دیگر تحقیقات صورت گرفته در مورد توزیع بالاکریستان می‌توان به موارد زیر اشاره کرد. یادگاری و همکاران(۲۰۰۸) و نیز آرلانوول و همکاران(۲۰۰۴)، به تعمیم دیگری از این توزیع دست یافتند (معادله ۱۰). حسنعلی پور و شرفی (۲۰۱۰) به بررسی برخی خصوصیات این توزیع پرداختند. نکوخو و علامت‌ساز (۲۰۱۱) با در نظر گرفتن توزیع لاپلاس در عوض تابع چگالی نرمال(معادله ۱۱) در توزیع تعمیم یافته آرلنوند و همکاران (۲۰۰۴) به بررسی برخی از خصوصیات این توزیع پرداخته و اصطلاحاً این توزیع را توزیع تعمیم یافته متقارن چوله لاپلاس نرمال نامگذاری کردند.

$$f_{\lambda_1, \lambda_2}(x) = 2c_{n, \lambda_1, \lambda_2}(\lambda) \times \varphi(x)\Phi^n\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \quad (10)$$

$$-\infty < x, \quad \lambda_1 < \infty, \quad \lambda_2 \geq 0,$$

که در آن

$$c_{n, \lambda_1, \lambda_2}(\lambda) = \left(\int_{-\infty}^{+\infty} 2\varphi(x) \times \Phi^n\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right) dx \right)^{-1},$$

¹⁰Skew-Generalized Normal (SGN)

¹¹Skew-Curved Normal

است. همچنین در این راستا و با توجه به مشکلات خانواده توزیع نرمال و همچنین تک مدی بودن خانواده توزیع نرمال چوله و لزوم وجود خانواده‌هایی با پارامتر چولگی و کشیدگی وسیعتر و همچنین ارائه توزیع‌های دومدی، محققان با تغییر (x) به ارائه توزیع نرمال تعییم یافه^{۱۰} پرداختند که در این زمینه می‌توان به کارهای ما و جنتن (۲۰۰۴) با در نظر گرفتن معادله (۱) و با پارامترهای زیر

$$w(x) = \lambda x + \gamma x^3,$$

$$g(x) = h(x) = \phi(x),$$

(معادله (۳)، بالاکریستان (۲۰۰۲) با در نظر گرفتن پارامترهای

$$w(x) = \lambda x,$$

$$g(x) = \varphi(x),$$

$$H(x) = c_n(\lambda)\Phi^n(\lambda x),$$

(معادله (۴))

$$f_{\lambda, \gamma}(x) = 2\varphi(x)\Phi(\lambda x + \gamma x^3), \quad (3)$$

$$-\infty < x < \infty,$$

$$f_{\lambda, n}(x) = 2c_n(\lambda) \varphi(x)\Phi^n(\lambda x), \quad (4)$$

$$c_n(\lambda) = \left(\int_{-\infty}^{+\infty} 2\varphi(x)\Phi^n(\lambda x) dx \right)^{-1},$$

آرلانوول و همکاران (۲۰۰۴) با در نظر گرفتن معادله (۱) و با پارامترهای $g(x) = h(x) = \phi(x)$ و $w(x) = \frac{\lambda_1 x}{\sqrt{(1+\lambda_2 x^2)}}$ (معادله (۵))

$$f_{\lambda_1, \lambda_2}(x) = 2\varphi(x)\Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \quad (5)$$

$$-\infty < x < \infty, \quad \lambda_2 \geq 0,$$

اشارة نمود.

در توزیع تعمیم یافته ول و همکاران در حالتی که $\lambda_1 = \lambda_2$ آنگاه توزیع حاصله توزیع نرمال چوله خمیده^{۱۱} نامیده می‌شود. یکی از مزایای استفاده از این تبدیل سادگی بدست آوردن توزیع مزدوج در محاسبه برآوردهای بیزی است [۳]. شرفی و بهبودیان (۲۰۰۶) در مورد خصوصیات توزیع بالاکریستان در حالتی که n برابر با ۲ است بحث

توزیع تی چوله به توزیع کوشی چوله و در صورتی که v به بینهایت میل کند آنگاه این توزیع به توزیع نرمال چوله میل می‌کند که کاربرد این توزیع‌ها در مدل‌سازی در منابع مختلف بیان شده است (آزالینی، ۲۰۰۸؛ لین و لی، ۲۰۰۸). همچنین در صورتی که توزیع U را توزیع بتا با پارامترهای قرار دهیم، توزیع X به توزیع چوله بریده^{۱۵} تبدیل می‌شود. کاربرد این توزیع‌ها در مدل‌سازی در مقاله ونگ و جنتن (۲۰۰۶) بیان شده است. همچنین در صورتی که U یک توزیع گستته که دو مقدار γ و 1 را با احتمال v و $1-v$ اختیار می‌کند انتخاب شوند، توزیع X به توزیع نرمال آلدود چوله (SCN)^{۱۶} (به فرم معادله ۱۶) تبدیل خواهد شد. در این فرمول هرگاه $\gamma = 1$ آنگاه SCN به توزیع نرمال چوله تبدیل می‌شود.

$$f(x) = 2 \left\{ v \varphi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\frac{\sqrt{\gamma}(x-\mu)}{\sigma}\right) + (1-v) \varphi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\frac{(x-\mu)}{\sigma}\right) \right\}. \quad (16)$$

همچنین در راستای دو مدی کردن چگالی توزیع‌های چوله اولیویرا و همکاران (۲۰۰۹) با اضافه کردن $\kappa = \int_{-\infty}^{+\infty} x^{\kappa} f(x) dy < \infty$ ، خانواده ارائه شده در معادله (۱) را به یک خانواده دومدی تبدیل نموده‌اند. نحوه تعریف این معادله در معادله (۱۷) و بصورت نمونه یک خانواده توزیع چوله دو مدی تعریف می‌شود. در این معادله هر چه مقدار α بیشتر باشد دومدی بودن تابع چگالی بیشتر و هر چه مقدار آن به صفر نزدیکتر باشد، میزان دومدی بودن تابع چگالی کمتر می‌شود.

$$f(x) = 2 \left(\frac{1+\alpha x^{\alpha}}{1+\alpha \kappa} \right) f_{\circ}(x) H(w(x)), \quad (17)$$

$$-\infty < x < \infty \alpha \geq 0,$$

همچنین باید گفت علاوه بر دو روش بیان شده قبلی محققان از راه‌کارهای دیگری نیز جهت تعریف توزیع‌های چوله بهره می‌جویند که از آن جمله می‌توان به روش آماره ترتیبی^{۱۷} اشاره نمود. در این روش توزیع چوله به شکل معادله (۱۸) و با دو پارامتر نامعلوم a و b تعریف می‌شود. خصوصیات این توزیع بصورت مفصل بوسیله گپتا و نداریج (۲۰۰۴) و ایوجن و همکاران (۲۰۰۲) مورد بررسی قرار گرفته است. اصطلاحاً به

¹²Skew-Normal/Independent Distributions

¹³Normal/Independent (NI)

¹⁴Mixture Normal

¹⁵Skew-Slash

¹⁶Skew-Contaminated Normal(SCN)

¹⁷Order Statistics Method

$$f_{\lambda_1, \lambda_2}(x) = e^{-|x|} \Phi\left(\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}\right), \quad (11)$$

$$-\infty < x, \lambda_1 < \infty, \lambda_2 \geq 0.$$

روش دیگر در ساخت توزیع‌های تی، کوشی چوله استفاده از راه‌کاری است که اصطلاحاً به آن کلاس توزیع‌های نرمال چوله-مستقل (SNI)^{۱۲} (SNI) می‌گویند (معادله (۱۲))

$$X = \mu + \frac{Z}{\sqrt{U}}. \quad (12)$$

در عبارت بالا μ پارامتر موقعیت بوده، U یک متغیر تصادفی مثبت با توزیع تجمعی $H(u|v)$ و تابع چگالی $h(u|v)$ است و v یک بردار که مشخص کننده توزیع متغیر تصادفی U می‌باشد در نظر گرفته می‌شود. Z نیز یک توزیع نرمال چوله با میانگین صفر، واریانس σ^2 و پارامتر چولگی λ است. ($Z \sim N(0, \sigma^2, \lambda)$) (آلانوول، بولفارین و لاچوس، ۲۰۰۵؛ لاچوس، دی و کانچوک، ۲۰۰۹) با دانستن U توزیع X بصورت معادله (۱۳) و توزیع حاشیه‌ای به شکل معادله (۱۴) تبدیل می‌شود.

$$X | u \tilde{SN}(\mu, u^{-1}\sigma^2, \lambda), \quad (13)$$

$$f(x) = 2 \int_0^\infty \Phi\left(\frac{\sqrt{u}\lambda(x-\mu)}{\sigma}\right) \times \phi\left(\frac{\sqrt{u}(x-\mu)}{\sigma}\right) dH(u|v). \quad (14)$$

در این معادله در صورتی که $\lambda = 0$ باشد آنگاه کلاس X به کلاس توزیع نرمال-مستقل (NI)^{۱۳} با تابع چگالی به شکل معادله (۱۵) تبدیل خواهد شد.

$$f(x) = \int_0^\infty \phi\left(\frac{\sqrt{u}(x-\mu)}{\sigma}\right) dH(u|v). \quad (15)$$

تابع چگالی کلاس NI در واقع یک کلاس از توزیع‌های نرمال آمیخته^{۱۴} است. همچنین در کلاس SNI در صورتی که به جای توزیع U ، توزیع کای دو با v قرار دهیم آنگاه توزیع X به توزیع تی چوله با v درجه آزادی که توسط آزالینی و کاپتانیو (۲۰۰۳) و گوبتا (۲۰۰۳) مورد تحقیق قرار گرفته است تبدیل می‌شود. همچنین در صورتی که v برابر ۱ باشد،

$$f_{\lambda}(x) = \frac{1}{\pi + 4\tan^{-1}(\lambda)} \varphi(x) \Phi(\lambda|x|), \quad (22)$$

$$-\infty < x, \lambda < +\infty,$$

و

داده‌های مورد بررسی در این مقاله براساس اطلاعات جمع‌آوری شده از افراد داطلب که طی یک مطالعه مقطعی در سال ۱۳۸۹ در مجتمع فولاد مبارکه اصفهان بصورت تصادفی انتخاب شده‌اند جمع‌آوری شده است (داده‌ها در ضمیمه ارائه شده است). در این مطالعه از میان ۶۴۰ نفر افراد ۱۶۶ مبتلا به فشار خون بالا با استفاده روش نمونه‌گیری تصادفی ساده نفر انتخاب و سپس این افراد فرم پرسشنامه استرس شغلی و اطلاعات دموگرافیک را تکمیل نمودند. در این مطالعه جهت اندازه‌گیری استرس شغلی از فرم استاندارد پرسشنامه GHQ12 (پرسشنامه سلامت عمومی) که جهت بررسی سلامت عمومی و مشکلات روانی در جامعه طراحی گردیده است استفاده شده است. منظری و همکاران (۲۰۰۳) روانی و پایانی نسخه فارسی این پرسشنامه را مورد تائید قرار داده‌اند. در این مطالعه اطلاعات جمعیت شناختی مانند سن و سابقه کاری با استفاده پرونده پزشکی جمع‌آوری شده و برای اندازه‌گیری فشار خون از دستگاه فشار خون جیوه‌ای کالیبره شده از دو دست و پس از ۵ دقیقه استراحت استفاده شده است. در این مطالعه همچنین متغیرهای دور کمر و شاخص چاقی^{۲۱} (BMI) نیز توسط پزشک اندازه‌گیری شده است.

۴ نتایج

در جدول ۴ میزان میانگین، میانه، انحراف معیار هر متغیر و همچنین میزان چولگی و کشیدگی متغیرهای موجود در مطالعه و همچنین میزان اختلاف این چولگی و کشیدگی از توزیع نرمال به ترتیب با استفاده از آزمون‌های دی آگوستینو (۱۹۷۰) و آزمون آنسکاپ-گلیان (۱۹۸۳) آن بسته "moments" و با استفاده از نرم افزار R به آزمون پرداختیم. همچنین از آنجایی که تعداد داده‌ها در این مطالعه تقریباً بزرگ بود

توزیع تشکیل یافته به شکل معادله (۱۸)، توزیع بتا-نرمال گفته می‌شود. از آنجایی که میزان چولگی توزیع بتا نرمال تحت دو پارامتر نامعلوم a و b یکنواخت نبوده و امکان دو مدی شدن و همچنین قابلیت پوشش توزیع‌های با چولگی زیاد را ندارد، برخی از محققین (فربا و استیل، ۲۰۰۴) ترجیح می‌دهند تا با قرار دادن شرط $b = a$ این توزیع را به توزیعی با یک پارامتر محدود کنند و به خصوصیات بهتری از این توزیع دست یابند. از دیگر راه‌کارهای تولید داده‌های چوله می‌توان به روش مقیاس عامل معکوس^{۱۸} (فرناندز و استیل، ۱۹۹۸) (معادله (۱۹)) و روش یکپارچه^{۱۹} (معادله (۲۰)) (فربا و استیل، ۲۰۰۶) اشاره کرد.

$$f_{a,b}(x) = [\Phi(x)]^{a-1} \times [1 - \Phi(x)]^{b-1} / 2B(a, b), \quad (18)$$

$$f_{\gamma}(x) = \frac{1}{\gamma + \frac{1}{\gamma}} \varphi(x^{\gamma - \text{sign}(x)}), \quad (19)$$

$$\gamma > 0,$$

$$f_p(x) = h(x)p[H(x)], \quad (20)$$

که در آن p یک تابع چگالی دلخواه است. شایان ذکر است که به غیر از روش‌های بیان شده، محققان مختلف بصورت مشخص به تعریف توزیع‌های چوله پرداخته‌اند که از آن جمله می‌توان به کارهای نادریج (۲۰۰۸) (معادله (۲۱)), کیم (۲۰۰۵) تحت عنوان توزیع نرمال چوله دو تکه‌ای^{۲۰} (معادله (۲۲)) و یا حالت تعیین یافته آن (جمالزاده، علیپور و بالاکریشنان، ۲۰۰۹) (معادله (۲۳)) و یا حالت ساده شده آن وقتی ($\gamma = 0$ معادله (۲۴)) اشاره کرد. در جداول ۱ تا ۳ توزیع‌های چوله متداول به همراه روش تعریف و خصوصیات آن ارائه شده است

$$f_{\alpha,a,b}(x) = Cx^{\alpha-1} \exp(-ax^{\alpha} - bx), \quad (21)$$

$$x > 0, a > 0, -\infty < b < +\infty.$$

که در آن

$$C = \int x^{\alpha-1} \exp(-ax^{\alpha} - bx),$$

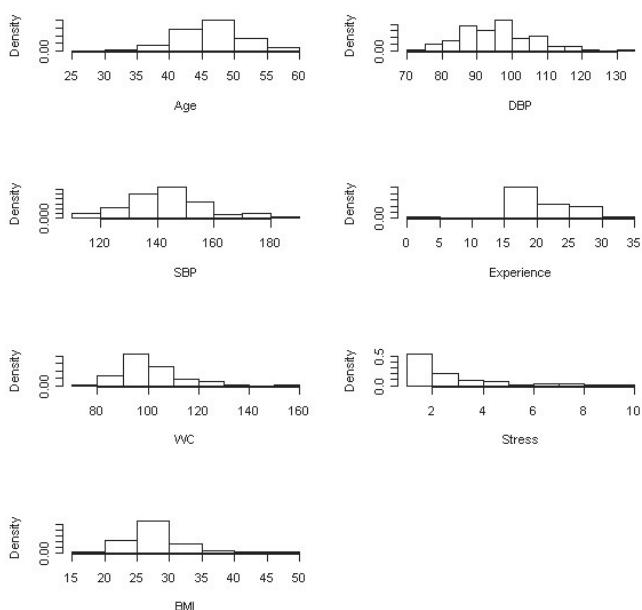
¹⁸Inverse Scale Factors Method

¹⁹Unified Method

²⁰Two-Piece Skew-Normal

²¹Body Mass Index (BMI)

نشود ولی میزان چولگی و کشیدگی داده مورد مطالعه در محدوده نرمال باشد باز توزیع نرمال چوله برآش بهتری را نسبت به توزیع نرمال نشان می‌دهد. این موضوع را از دیدگاه دیگر می‌توان اینگونه بیان نمود که از آنجایی که اثر عدم تقارن بر فرض نرمالیتی در کل جدی تر از دم سنگین بودن توزیع بواسطه وجود نقاط پرت است (مدھولکر و هیتسون، ۲۰۰۰). در این موارد توزیع نرمال چوله برآش بهتری را نسبت به نرمال نشان خواهد داد. همچنین در مورد متغیرهای سابقه کار، دور کمر، استرس و شاخص BMI از آنجایی که میزان چولگی و کشیدگی هر دو فاصله زیادی با توزیع نرمال داشتند از این رو استفاده از توزیعی که علاوه بر قابلیت مدل‌سازی چولگی، قابلیت بالایی در مدل‌سازی کشیدگی دارد مورد پیش‌بینی بود. از این رو توزیع تی چوله گزینه مناسبی برای این متغیرها به شمار می‌رفت که نتایج حاصل از برآش مدل (جدول ۵) نیز نشان دهنده این برآش بهتر در توزیع تی چوله نسبت به سه توزیع دیگر بود. در پایان و براساس نتایج این مطالعه می‌توان راهکار زیر را جهت مدل‌سازی به فرم جدول ۶ ارائه نمود. به این ترتیب که ابتدا براساس آزمون‌های شاپیرو، دی آگوستینو و آنسکاپ-گلیان به ترتیب به بررسی نرمال بود، چولگی و کشیدگی متغیر مورد مطالعه با توزیع نرمال پرداخته و سپس با استفاده از جدول ۶ به انتخاب توزیع مناسب از میان ۴ توزیع نرمال، تی، نرمال چوله و تی چوله دست می‌زنیم.



شکل ۱. نمودار هیستوگرام برای متغیرهای مورد مطالعه

²²Shapiro test

(n = ۱۶۶) جهت بررسی همتوزیعی متغیرهای مورد مطالعه با توزیع نرمال از آزمون شاپیرو^{۲۲} (پاتریک، ۱۹۸۲) و با استفاده از نرم افزار R به این مهم پرداخته و نتایج را در جدول ۴ ارائه نمودیم. نمودار هیستوگرام متغیرهای مورد مطالعه نیز در شکل ۱ نمایش داده شده است. با توجه به نمودار ۱ و نتایج جدول ۴ می‌توان به خوبی به تخطی توزیع داده‌ها از فرضیه نرمال بودن مشاهدات بی‌برد. در جدول ۵ خلاصه نتایج برآش و همچنین AIC مدل که یک شاخص کفایت مدل است. (مقادیر کمتر آن نشان دهنده برآش بهتر است) حاصل از ۴ توزیع نرمال، تی، نرمال چوله و تی چوله ارائه شده است. برای برآش توزیع‌های نرمال و تی با استفاده از بسته "bbmle"^{۲۳} و جهت برآش توزیع‌های نرمال چوله و تی چوله با استفاده از بسته "sn"^{۲۴} و با استفاده از نرم افزار R به این منظور دست یافتیم.

۵ بحث و نتیجه‌گیری

یکی از مسائل فرا روی محققان در استفاده از توزیع‌های آماری، انتخاب درست این توزیع‌ها است. در این مقاله به منظور آشکار ساختن کاربرد توزیع‌های چوله به برآش ۴ توزیع نرمال، تی، نرمال چوله و بالاخره تی چوله در یکسری داده پزشکی با درجات مختلف چولگی و کشیدگی دست زدیم. همانگونه که در جدول ۴ مشاهده می‌نمایید از آنجایی که میزان چولگی متغیر سن در حد توزیع نرمال و میزان کشیدگی آن در حد این توزیع نیست می‌توان پیش از برآش مدل به مناسب بودن توزیع تی به عنوان بهترین توزیع جهت برآش مدل مشکوک بود. این فرضیه با استفاده از معیار شاخص AIC مورد بررسی قرار گرفت و براساس نتایج جدول ۵ این فرضیه مورد تأیید قرار گرفت. همچنین برای متغیرهای فشار خون سیستولیک و دیاستولیک با وجودی که میزان چولگی و کشیدگی در حد توزیع نرمال گزارش گردید اما آزمون نرمال بودن داده‌ها تأیید نشده و براساس مقادیر AIC توزیع نرمال چوله بهترین گزینه برای برآش برای این متغیرها در نظر گرفته شد. آزالینی نشان داده است که میزان پارامتر چولگی توزیع نرمال چوله در بازه زیر (-۰/۹۹۵, ۰/۹۹۵)

تغییر می‌کند و این یعنی اینکه از توزیع نرمال چوله در موقعي می‌توان استفاده کرد که چولگی داده‌ها زیاد نباشد (آزالینی، ۱۹۸۵؛ اسپینوسا، ۲۰۰۸). بنابراین در موقعي که هم توزیعی داده‌ها با توزیع نرمال تائید

جدول ۱: برخی از توزیع‌های چوله معروف براساس نوع انتخاب و (معادله (۱))

$H(.)$	$g(.)$	$w(.)$	نام توزیع
نرمال	نرمال	λx	نرمال چوله یا نرمال-نرمال چوله
کوشی	کوشی	λx	کوشی چوله
تی	تی	λx	t -چوله
یکنواخت	یکنواخت	λx	چوله یکنواخت
لایاس	لایاس	λx	لایاس چوله
لجستیک	لجستیک	λx	لجستیک چوله
کوشی	نرمال	λx	کوشی-نرمال چوله
تی	نرمال	λx	نرمال- t چوله
یکنواخت	نرمال	λx	یکنواخت-نرمال چوله
لایاس	نرمال	λx	لایاس-نرمال چوله
لجستیک	نرمال	λx	نرمال-لجستیک چوله
نرمال	*تی	λx	تی - نرمال چوله
نرمال	یکنواخت	λx	یکنواخت - نرمال چوله
نرمال	لایاس	λx	لایاس - نرمال چوله
نرمال	لجستیک	λx	لجستیک - نرمال چوله
$c_n(\lambda) \Phi^n(.)$	نرمال	λx	بالاکریشنان چوله
نرمال	نرمال	$\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}$	توزیع چوله نرمال تعمیم یافته نوع ۱
نرمال	نرمال	$\lambda x + \gamma x^3$	توزیع چوله نرمال تعمیم یافته نوع ۲
نرمال	نرمال	$\frac{\lambda_1 x + \lambda_2 x^3}{\sqrt{1+\lambda_2 x^2}}$	توزیع چوله نرمال تعمیم یافته نوع ۴
نرمال	تی	$\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}$	توزیع چوله تی-نرمال تعمیم یافته نوع ۱
نرمال	تی	$\lambda x + \gamma x^3$	توزیع چوله تی-نرمال تعمیم یافته نوع ۲
$c_{n,v}(\lambda) \Phi^n(.)$	تی	λx	توزیع چوله تی-نرمال تعمیم یافته نوع ۳
نرمال	تی	$\frac{\lambda_1 x + \lambda_2 x^3}{\sqrt{1+\lambda_2 x^2}}$	توزیع چوله تی-نرمال تعمیم یافته نوع ۴
نرمال	لایاس	$\frac{\lambda_1 x}{\sqrt{1+\lambda_2 x^2}}$	توزیع چوله لایاس - نرمال تعمیم یافته

جدول ۲: برخی از توزیع‌های چوله معروف براساس نوع انتخاب U (معادله (۱۱))

U	نام توزیع
$\chi_{(n)}^2, (n > 1)$	تی چوله
$\chi_{(1)}^2$	کوشی چوله
$Beta(v, 1)$	چوله بربیده شده
$v I(\gamma) + (1-v)I(1)$	چوله آلدوده به نرمال

جدول ۳: تابع برخی از توزیع‌های چوله معروف دیگر

نام توزیع	تابع چگالی
توزیع بتا نرمال دو پارامتری	$\frac{[\Phi(.)]^{a-1}[(1-\Phi(.))^{b-1}}{\sqrt{B(a,b)}}$
توزیع بتا نرمال تک پارامتری	$\frac{[\Phi(.)]^{a-1}[(1-\Phi(.))^{\frac{1}{a}-1}}{\sqrt{B(a,\frac{1}{a})}}$
توزیع عامل معکوس	$f(x) = \frac{1}{\gamma + \frac{1}{x}} \varphi(x\gamma^{-sign(x)})$
توزیع چوله نادریج	$Cx^{\alpha-1} \exp(-ax^\alpha - bx)$
توزیع نرمال چوله دو تکه‌ای	$\frac{1}{\pi + \sqrt{tan^{-1}(\lambda)}} \varphi(x) \Phi(\lambda x)$
توزیع نرمال چله دو تکه‌ای تعمیم یافته	$\frac{1}{\pi + \sqrt{tan^{-1}(\lambda_1)}} \varphi(x) \Phi(\lambda_1 x) \Phi(\lambda_2 x)$

جدول ۴: میزان میانگین، میانه، انحراف معیار، چولگی و کشیدگی متغیرهای موجود در مطالعه و میزان اختلاف این چولگی و کشیدگی از توزیع نرمال

BMI++	استرس	WC++	سابقه کار	SBP**	DBP*	سن	پارامتر
۲۷,۶۶	۲,۸۷	۱۰۰,۴۷	۲۱,۳۰	۱۴۴,۶۵	۹۶,۸۹	۴۶,۴۷	میانگین
۲۷,۰۳	۲	۹۹	۲۰	۱۴۴	۹۶	۴۷	میانه
۴,۰۷	۲	۱۱,۴۴	۴,۷۱	۱۴,۵	۱۰,۵۹	۵,۴۴	انحراف معیار
۱,۶۱	۱,۴۷	۱,۳۵	-۰,۹۸	۰,۵۳	۰,۴۰	-۰,۲۵	چولگی
۴,۳۲	۴,۰۸	۳,۸۴	-۳,۰۵	۱,۸۱	۱,۴۰	۰,۸۹۹	آماره آزمون \dagger
۰,۰۰۰۱	۰,۰۰۰۱	۰,۰۰۰۱	۰,۰۰۰۲	۰,۰۷	۰,۱۶	۰,۳۶۸	\ddagger Pvalue
۷,۷۳	۵,۰۲	۵,۹۲	۵,۹۸	۳,۲۷	۳,۲۹	۳,۹۸	کشیدگی
۵,۰۱	۳,۳۶	۴,۰۵	۴,۰۹	۰,۹۴	۰,۹۷	۲,۱۹۷	آماره آزمون \ddagger
۰,۰۰۰۱	۰,۰۰۰۸	۰,۰۰۰۱	۰,۰۰۰۴	۰,۳۴۸۸	۰,۳۳۰۹	۰,۰۲۸	\ddagger Pvalue
۰,۸۹۳	۰,۸۹۲	۰,۹۰۹	۰,۸۷۳	۰,۹۷۰	۰,۹۸۳	۰,۹۸۱	آماره آزمون \circ
۰,۰۰۰۱	۰,۰۰۰۱	۰,۰۰۰۱	۰,۰۰۰۱	۰,۰۰۱	۰,۰۴۱	۰,۰۱۸	\circ Pvalue

* : فشار خون دیاستولیک، ** : فشار خون سیستولیک، +: دور کمر، ++: شاخص توده بدنی،

\dagger : آماره آزمون و همچنین Pvalue برای آزمون دی آگوستینو جهت مقایسه چولگی متغیر با چولگی توزیع نرمال

\ddagger : آماره آزمون و همچنین Pvalue برای آزمون آنسکاپ-گلیان جهت مقایسه کشیدگی متغیر با کشیدگی توزیع نرمال

\circ : آماره آزمون و همچنین Pvalue برای آزمون شاپیرو جهت مقایسه توزیع داده‌ها با توزیع نرمال

جدول ۵: پارامترهای مدل برای متغیرهای مورد مطالعه براساس ۴ مدل نرمال، تی، نرمال چوله و تی چوله

BMI++	استرس	WC+	ساقه کار	SBP**	DBP*	سن	بارامتر	مدل
۲۷.۶۶ (±۰.۳۱)	۲/۸۷ (±۰.۱۵)	۱۰۰/۴۷ (±۰.۸۸)	۲۱/۳۰ (±۱/۱۲)	۱۲۴.۶۵ (±۰.۸۲)	۹۶.۸۹ (±۰.۲۲)	۴۶/۴۷ (±۰/۱۵)	μ	نرمال
۷.۰۶ (±۰.۲۲)	۱/۸۹ (±۰.۱۰)	۱۱/۴۱ (±۰.۶۳)	۶.۶۹ (±۰.۴۶)	۱۶.۴۵ (±۰.۷۹)	۱۰.۵۶ (±۰.۵۸)	۵۴۲ (±۰.۳۰)	σ	
±۹۲۰/۲۱	±۶۸۷/۰۲	±۱۸۲۷/۳۶	±۹۸۸/۲۷	±۱۳۶۱/۸۹	±۱۲۵۷/۷۷	±۱۰۳۶/۴۰	AIC	
۴۶.۹۵ (±۰.۲۶)	۲.۷۸ (±۰.۱۳)	۹۸.۵۲ (±۰.۷۳)	۲۱/۳۵ (±۰.۳۲)	۱۴۶.۰ (±۱/۱۷)	۹۶.۷۳ (±۰.۸۴)	۴۶/۵۳ (±۰.۲۱)	μ	تی
۲.۵۷ (±۰.۲۷)	۱/۱۹ (±۰.۱۳)	۷/۴۰ (±۰.۸۰)	۳.۴۳ (±۰.۳۱)	۱۲.۹۷ (±۰.۱۵)	۱۰.۰۶ (±۰.۸۳)	۴.۸۱ (±۰.۴۲)	σ	
۲.۹۶ (±۰.۷۹)	۷/۷۲ (±۰.۷۱)	۷.۷۵ (±۰.۸۲)	۴.۲۱ (±۱.۳۲)	۹.۷۴ (±۰.۴۳)	۲۱.۱۸ (±۲۷.۹۷)	۹.۴۲ (±۰.۹۸)	df	
۹۰۹/۵۸	۶۶۸/۲۶	۱۲۶۰/۹۹	۹۶۸/۷۹	۱۳۶۷/۷۶	۱۲۵۷/۳۲	۱۰۳۶/۸۸	AIC	
۱۲.۲۲ (±۰.۲۵)	۰/۹۹ (±۰.۰۰۱۴)	۸۷.۷۰ (±۰.۷۲)	۵۰.۶۹ (±۰.۳۵)	۱۱۹.۸۷ (±۱/۰۶)	۸۷.۶۵ (±۰.۷۹)	۵۰.۴۳ (±۰.۴۲)	μ	نرمال چوله
۶۰۰/۱ (±۰.۰۸)	۲۷/۰۲ (±۰.۰۰۱۱)	۱۷/۱۲ (±۰.۰۵)	۶.۴۲ (±۰.۲۵)	۲۰.۶۷ (±۰.۱۷)	۱۶.۷۱ (±۰.۵۷)	۶.۷۲ (±۰.۴۰)	σ	
۷.۴۹ (±۰.۰۸)	۱۵۲۳/۷۷ (±۰.۰۰۰۰۱)	۷/۵۵ (±۰.۰۸)	-۱.۶۷ (±۰.۱۷)	۷.۶۰ (±۰.۱۴)	۱/۷۹ (±۰.۱۷)	-۱.۱۰ (±۰.۱۵)	λ	
۹۰۶/۸۱	۵۷۷۷/۱۷	۱۲۵۳/۲۱	۹۷۹/۵۰	۱۳۵۶/۰۸	۱۲۵۵/۴۱	۱۰۳۷/۰۵	AIC	
۲۶.۱۷ (±۰.۵۰)	۰/۹۹ (±۰.۰۰۱)	۹۰/۲۶ (±۱.۶۲)	۱۸/۸۶ (±۰.۴۰)	۱۳۶.۵۷ (±۱۰.۳۷)	۸۷.۶۶ (±۱/۸۱)	۴۸.۲۶ (±۱.۱۲)	μ	تی چوله
۱۶.۸۴ (±۰.۱۶)	۲۷.۳۹ (±۰.۱۱)	۱۳۵/۰۹ (±۰.۲۰)	۴.۳۲ (±۰.۲۲)	۲۰۰/۳۹ (±۰.۴۵)	۲۱۶/۱۴ (±۰.۱۱)	۲۵.۸۲ (±۰.۲۰)	σ	
۷.۰۹ (±۰.۴۹)	۳۱۸۹۳۲۷۸ (±۱۸۵۹۸)	۲/۰۹ (±۰.۴۹)	۱/۵۳ (±۰.۴۶)	۰/۷۰ (±۰.۹۳)	۱/۷۸ (±۰.۴۴)	-۰.۴۴ (±۰.۶۷)	λ	
۳.۹۱ (±۱.۳۷)	۷.۳۰ (±۱.۰۱)	۴.۵۹ (±۱.۹۹)	۱/۱۵ (±۰.۲۵)	۸/۱۶ (±۰.۷۹)	۲۵۱۹/۲۱ (±۵۹۷۲۸)	۹.۶۶ (±۶.۶۲)	df	
۸۹۶.۴۵	۵۵۸/۲۷	۱۲۴۷/۰۷	۹۵۷/۱۶	۱۳۶۱/۲۸	۱۲۵۷/۲۱	۱۰۳۶/۵۰	AIC	

*: فشار خون دیاستولیک، **: فشار خون سیستولیک،

+: دور کمر، +: شاخص توده بدنی،

جدول ۶:تابع برخی از توزیع‌های چوله معروف دیگر

توزيع انتخابی	کشیدگی	چولگی	نرمال
نرمال	همواره عدم اختلاف معنی‌دار	همواره عدم اختلاف معنی‌دار	عدم اختلاف معنی‌دار از همتوزیعی با توزیع نرمال
	از کشیدگی توزیع نرمال	از چولگی توزیع نرمال	
تی	اختلاف معنی‌دار	عدم اختلاف معنی‌دار	اختلاف معنی‌دار از همتوزیعی با توزیع نرمال
	از کشیدگی توزیع نرمال	از چولگی توزیع نرمال	
نرمال چوله	عدم اختلاف معنی‌دار	عدم اختلاف معنی‌دار	اختلاف معنی‌دار از همتوزیعی با توزیع نرمال
	از کشیدگی توزیع نرمال	از چولگی توزیع نرمال	
نرمال چوله	عدم اختلاف معنی‌دار	اختلاف معنی‌دار	اختلاف معنی‌دار از همتوزیعی با توزیع نرمال
	از کشیدگی توزیع نرمال	از چولگی توزیع نرمال	
تی چوله	اختلاف معنی‌دار	اختلاف معنی‌دار	اختلاف معنی‌دار از همتوزیعی با توزیع نرمال
	از کشیدگی توزیع نرمال	از چولگی توزیع نرمال	

مراجع

- [1] Abtahi, A., M. Towhidi and J. Behboodian (2009). An appropriate empirical version of skew-normal density. *Stat Papers*.
- [2] Anscombe, F. J. and W. J. Glynn (1983). Distribution of kurtosis statistic for normal statistics. *Biometrika*, **70(1)**, 227-234.
- [3] Arellano-Valle, R., H. Go'mez and F. Quintana (2004). A new class of skew-normal distributions. *Communications in Statistics: Theory and Methods*, **33(7)**, 1465–1480.
- [4] Arellano-Valle, R. B., H. Bolfarine and V. H. Lachos (2005). Skew-normal linear mixed models. *Journal of Data Science*, **3**, 415–438.
- [5] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statist. Theory and Applications*, **12(2)**, 171-178.
- [6] Azzalini, A. (2005). The skew-normal distribution and related multivariate families. *Scand. J. Statist. Theory and Applications*, **32(2)**, 159–200.
- [7] Azzalini, A. and A. Capitanio (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society*, **61**(Series B), 579-602.
- [8] Azzalini, A. and A. Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, **65**, 367-389.
- [9] Azzalini, A. and A. DallaValle (1996). The multivariate skew-normal distribution. *Biometrika*, **83(4)**, 715-726.
- [10] Azzalini, A., Genton, M., (2008). Robust likelihood methods based on the skew-t and related distributions. *International Statistical Review*, **76**, 106–129.
- [11] Balakrishnan, N. (2002). Discussion on Skew Multivariate Models Related to Hidden Truncation and/or Selective Reporting by B. C. Arnold and R. J. Beaver. *Test*, **11**, 37-39.
- [12] Bandyopadhyay, D., V. H. Lachos, C. A. Abanto-Valle and P. Ghosh (2010). Linear mixed models for skew-normal/independent bivariate responses with an application to periodontal disease. *Statist. Med*, **29**, 2643–2655.
- [13] D'Agostino, R. B. (1970). Transformation to Normality of the Null Distribution of G1. *Biometrika*, **57(3)**, 679-681.

- [14] Dey, D. (2010). *Estimation of the Parameters of Skew Normal Distribution by Approximating the Ratio of the Normal Density and Distribution Functions*. *Applied Statistics.*, CALIFORNIA, UNIVERSITY OF CALIFORNIA RIVERSIDE. PHD.
- [15] Elal-Olivero, D., H. W. Gomez and F. A. Quintana (2009). Bayesian modeling using a class of bimodal skew-elliptical distributions. *Journal of Statistical Planning and Inference*, **139**(4), 1484-1492.
- [16] Eugene, N., C. Lee and F. Famoye (2002). Beta-normal distribution and its applications. *Commun Stat Theory Methods*, **31**, 497–512.
- [17] FERNANDEZ, C. and J. STEEL MF (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, **93**, 359-371.
- [18] Ferreira, J. T. A. S. and J. STEEL MF (2004). Discussion of Mc Jones, Families of distributions arising from distributions of order statistics. *Test*, **13**, 30–33.
- [19] Ferreira, J. T. A. S. and J. STEEL MF (2006). A constructive representation of univariate skewed distributions. *J Am Stat Assoc*, **101**, 823–829.
- [20] Genton, M. G. (2004). *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*, USA, Chapman & Hall/CRC.
- [21] Gómez, H. W., O. Venegas and H. Bolfarine (2007). Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics*, **18**(4), 395-407.
- [22] Gupta, A. K. (2003). Multivariate skew t-distributions. *Statistics*, **37**, 359–363.
- [23] Gupta, A. K., F. C. Chang and W. J. Huang (2002). Some skew-symmetric models. *Random Operators and Stochastic Equations*, **10**(2), 133-140.
- [24] Gupta, A. K. and S. Nadarajah (2004). On the moments of the beta-normal distribution. *Commun Stat Theory Methods*, **33**(1), 1-13.
- [25] Hasanalipour, P. and M. Sharafi (2010). A new generalized Balakrishnan skew-normal distribution. *Stat Papers*.
- [26] Jamalizadeh, A., A. R. Arabpour and N. Balakrishnan (2009). A generalized skew two-piece skew-normal distribution. *Stat Papers*.
- [27] Jara, A., F. Quintana and E. San Martin (2008). Linear mixed models with skew-elliptical distributions: a Bayesian approach. *Computational Statistics and Data Analysis*, **52**, 5033–5045.

- [28] Kheradmandi, A., M. Mohammadzadeh and N. Sanjari Farsipour (2010). Generalizations of the Skew t-Normal Distribution and their Properties. *Journal of Sciences, Islamic Republic of Iran*, **21(4)**, 343-351.
- [29] KIM, H. J. (2005). On a class of two-piece skew-normal distributions. *Statistics*, **39**, 537–553.
- [30] Lachos, V. H., D. K. Dey and V. G. Canchoc (2009). Robust linear mixed models with skew-normal independent distributions from a Bayesian perspective. *Journal of Statistical Planning and Inference*, **139**, 4098 - 4110.
- [31] Lee, J. C., T. I. Lin, K. J. Lee and Y. L. Hsu (2005). Bayesian analysis of Box-Cox transformed linear mixed models with ARMA(p,q) dependence. *Journal of Statistical Planning and Inference*, **133(2)**, 435-451.
- [32] Lin, T. I. and J. C. Lee (2008). Estimation and prediction in linear mixed models with skew-normal random effects for longitudinal data. *Statistics in Medicine*, **27**, 1490–1507.
- [33] Ma, Y., M. G. Genton and M. Davidian (2004). *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality. Linear mixed models with flexible generalized skew-elliptical random effect*, Chapman & Hall, CRC Press, Boca Raton, FL,.
- [34] Montazeri, A., A. Harirchi, M. Shariati, G. Garmaroudi, M. Ebadi and A. Fateh (2003). The 12 item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version. *Health and Quality of life Outcome*, **1**, 66-70.
- [35] Mudholkar, G. S. and A. D. Hutson (2000). The epsilon-skew-normal distribution for analyzing near-normal data. *Journal of Statistical Planning and Inference*, **83(2)**, 291-309
- [36] Nadarajah, S. (2008). A new model for symmetric and skewed data. *Probab Eng Inf Sci*, **22**, 261–271.
- [37] Nadarajah, S. and S. Kotz (2003). Skewed distributions generated by the normal kernel. *Statist. Probab. Lett.*, **65(3)**, 269-277.
- [38] Nekoukhoo, V. and M. H. Alamatsaz (2011). A family of skew-symmetric-Laplace distributions. *Stat Papers*.
- [39] Patrick, R. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Applied Statistics*, **31**, 115–124.
- [40] Pewsey, A. (2000). Problems of inference for Azzalini's skew-normal distribution. *J. Appl. Statist*, **27(7)**, 859- 870.
- [41] Sahu, S. K., D. K. Dey and M. D. Branco (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics*, **31**, 129–150.

- [42] SHARAFI, M. and J. BEHBOUDIAN (2006). "New skew normal density." *Journal of Statistical Research Of Iran (JSRI)*, **3(1)**, 47-61.
- [43] Spinosa, J. S. (2008). *Inference for the Multiparameter Skew Normal Distribution*, Applied statistic. California, University of California Riverside. PHD.
- [44] Tsung, I. L. and J. C. Lee (2003). On modelling data from degradation sample paths over time. *Australian and New Zealand Journal of statistics*, **45(3)**, 257-270.
- [45] Wang, J. and M. Genton (2006). The multivariate skew-slash distribution. *Journal of Statistical Planning and Inference*, **136**, 209–220.
- [46] Yadegari, I., A. GERAMI and M. J. Khaledi (2008). A generalization of the Balakrishnan skew-normal distribution. *Stat Probab Lett*, **78**, 1165–1167.

ضمیمه

فرمول‌های تکمیلی

$$f_{\lambda_1, \lambda_2, \rho}(x) = \frac{\pi}{\cos^{-1}\left(\frac{-(\rho + \lambda_1 \lambda_2)}{\sqrt{1+\lambda_1^2} \sqrt{1+\lambda_2^2}}\right) + \cos^{-1}\left(\frac{-(\rho - \lambda_1 \lambda_2)}{\sqrt{1+\lambda_1^2} \sqrt{1+\lambda_2^2}}\right) + 2\tan^{-1}(\lambda_2)} \\ \times \varphi(x) \Phi(\lambda_1 x, \lambda_1 |x|, \rho), \quad -\infty < x, \lambda_1, \lambda_2 < +\infty \quad (23)$$

$$f_{\lambda_1, \lambda_2}(x) = \frac{1}{\pi + 2\tan^{-1}(\lambda_2)} \varphi(x) \Phi(\lambda_1 x) \Phi(\lambda_2 |x|), \quad -\infty < x, \lambda_1, \lambda_2 < +\infty. \quad (24)$$

فرم داده‌ها در نرم افزار R