



## توزیع چوله نرمال چند مدی

فروغ نقیبی، سید محمد رضا علوی<sup>\*</sup>، رحیم چینی پرداز

گروه آمار، دانشکده علوم ریاضی و کامپیوتر، دانشگاه شهید چمران اهواز

تاریخ پذیرش: ۱۴۰۰/۲/۲۵

تاریخ دریافت: ۱۳۹۹/۸/۱۶

دبیر مسئول: غلامعلی پرهام

چکیده: در این مقاله ابتدا توزیع نرمال چند مدی که تعمیمی از توزیع نرمال است و توزیع‌های متقارن یک مدی، دو مدی و سه مدی را پوشش می‌دهد، معرفی می‌شود و سپس توزیع چوله نرمال چند مدی به عنوان یک توزیع چوله-متقارن که بوسیله تابع توزیع نرمال استاندارد تولید می‌شود، معرفی و خواصی از این توزیع بررسی می‌شود. روش‌های تولید داده از توزیع پیشنهادی مطرح شده و حالت مکانی و مقیاسی توزیع معرفی می‌شود. برآورد پارامترها به روش ماکسیمم درستنمایی با توجه خاصیت مجانبی آنها محاسبه و با استفاده از شبیه‌سازی برآوردها ارزیابی می‌شوند. مجموعه داده‌های معروف آفشان به این توزیع برازش داده و با استفاده از معیارهای آکائیک، آکائیک تعمیم یافته و معیار بیزی شوارتر با چند توزیع مرسوم مقایسه می‌گردند. نتیجه‌گیری نتایج حاصل از این مقاله آخرین بخش آن خواهد بود.

واژه‌های کلیدی: چوله نرمال، چند مدی، داده‌های آفشان، برآورد.

رده‌بندی ریاضی: 62E05; 62E10

### ۱ مقدمه

آرالینی [۵] برای مدل‌بندی داده‌های نامتقارن یک مدی توزیع چوله نرمال را که دارای یک پارامتر چولگی  $\lambda$  است و با نماد  $SN(\lambda)$  نمایش داده می‌شود به صورت

$$f(x; \lambda) = 2\phi(x)\Phi(\lambda x), \quad \lambda \in R \quad (1.1)$$

تعریف کرد که  $\phi(\cdot)$  و  $\Phi(\cdot)$  به ترتیب تابع چگالی و تابع توزیع احتمال نرمال استاندارد هستند. با جایگذاری یک تابع چگالی متقارن به جای  $\phi(x)$  در  $SN(\lambda)$  توزیع‌های چوله نرمال دیگری به عنوان توزیع‌های چوله-متقارن بر اساس لم آرالینی [۶] معرفی گردید. با استفاده از این ایده گامز و همکاران [۸] توزیع چوله-t نرمال و نکوخو و علامت ساز [۱۴] توزیع چوله لاپلاس-نرمال را معرفی کردند. شرفی و بهبودیان [۱۷] و یادگاری و همکاران [۱۹] تعمیم‌هایی از توزیع چوله نرمال ارائه دادند. توزیع‌های چوله زیادی توسط محققین مختلفی از جمله ماملی و موزیو، [۱۳] راسخی و همکاران [۱۶]، علوی و طرهانی [۳] و مالکی و نعمت‌اللهی [۱۲] مطالعه شده است.

علوی [۲] با استفاده از ایده توزیع‌های وزنی، توزیع نرمال دو مدی مرتبه  $\beta$  را معرفی و آن را با  $BN(\beta)$  نمایش داد که تابع چگالی احتمال آن به صورت:

$$g(x; \beta) = \frac{|x|^\beta}{k} \phi(x) \quad (۲.۱)$$

تعریف می‌شود. در اینجا  $\beta \geq 0$ ،  $k = \frac{\Gamma(\frac{\beta+1}{2}) 2^{\frac{\beta}{2}}}{\Gamma(\frac{1}{2})}$  پارامتر مد و  $\Gamma(\cdot)$  تابع گاما است. این توزیع حول صفر متقارن و با  $\beta = 0$  توزیع نرمال استاندارد حاصل می‌شود.

در این مقاله توزیع جدیدی از نرمال معرفی می‌گردد که می‌تواند حالت چندمدی داشته و توزیع‌های نرمال استاندارد و نرمال دو مدی ([۲]) را به عنوان حالت خاص پوشش دهد. با توجه به انعطاف پذیری توزیع پیشنهادی، امکان معرفی گونه آن برای برازش داده‌های غیر متقارن چند مدی نیز فراهم می‌گردد. خواص توزیع، برآورد پارامترها و مقایسه آن با بعضی از توزیع‌های پیشنهادی از طریق شبیه‌سازی و داده‌های واقعی در این مقاله پیگیری می‌شود. مقاله در شش بخش به صورت زیر تنظیم شده است.

در بخش دوم جهت انعطاف پذیری بیشتر توزیع‌های متقارن با استفاده از توزیع نرمال به معرفی توزیع نرمال متقارن چند مدی پرداخته، ضمن بررسی خواص آن، نشان داده می‌شود که این توزیع شامل توزیع‌های یک مدی و چند مدی است. در بخش سوم و در ادامه گسترش این توزیع، گونه چوله آن با نام توزیع چوله نرمال چند مدی معرفی می‌شود. انعطاف پذیری و خواصی از آن را بیان کرده و در بخش چهارم چند روش برای تولید داده از آن بیان می‌شود. در بخش پنجم توزیع مکانی-مقیاسی چوله نرمال چندمدی بررسی و برآورد پارامترها به روش ماکسیمم درستنمایی با روش عددی محاسبه و توسط یک مطالعه شبیه سازی ارزیابی می‌شوند. در بخش ششم توزیع پیشنهادی به مجموعه معروف داده‌های آبفشان برازش داده می‌شود و برتری این برازش با چند توزیع رقیب نشان داده می‌شود.

## ۲ توزیع نرمال چند مدی

فرض کنید  $g(x; \beta)$  و  $G(x; \beta)$  به ترتیب تابع چگالی احتمال و تابع توزیع تجمعی توزیع نرمال دو مدی مرتبه  $\beta$  باشد، توزیع نرمال چند مدی با تابع چگالی زیر را معرفی می‌کنیم:

$$g(x; \alpha, \beta) = \frac{(\alpha + |x|^\beta)}{(\alpha + k)} \phi(x) \quad (۱.۲)$$

که در اینجا نیز  $\phi(x)$  توزیع نرمال استاندارد و  $\alpha \geq 0$  غیر تصادفی است. توزیع نرمال چند مدی با نماد  $BN(\alpha, \beta)$  و تابع توزیع آن را با  $G(x; \alpha, \beta)$  نشان داده و برابر است با:

$$g(x; \alpha, \beta) = \frac{\alpha \Phi(x) + k G(x, \beta)}{\alpha + k} \quad (۲.۲)$$

میانگین و واریانس توزیع به ترتیب صفر و

$$\frac{1}{\alpha + k} \left( \alpha + \frac{k}{\Gamma(\frac{1}{2})} \Gamma\left(\frac{\beta+1}{2} + 1\right) \left(\frac{1}{2}\right)^{\left(\frac{\beta+1}{2} + 1\right)} \right)$$

است.

گشتاورهای فرد، صفر و گشتاورهای زوج از رابطه زیر به دست می‌آیند:

$$E(X^{2r}) = \frac{1}{\alpha + k} \left( \frac{\alpha}{\Gamma(\frac{1}{2})} \Gamma\left(r + \frac{1}{2}\right) (2)^r + \frac{k}{\Gamma(\frac{k}{2})} \Gamma\left(\frac{\beta+1}{2} + r\right) \left(\frac{1}{2}\right)^{\left(\frac{\beta+1}{2} + r\right)} \right)$$

تابع مولد گشتاور آن یک ترکیب خطی از تابع مولد گشتاور نرمال  $(M(t))$  و تابع گشتاور نرمال دو مدی  $(M(t, \beta))$  است یعنی:

$$M(t, \alpha, \beta) = \frac{\alpha}{\alpha + k} M(t) + \frac{k}{k + \alpha} M(t, \beta)$$

قضیه ۱.۲. اگر  $X \sim BN(\alpha, \beta)$  در آن صورت:  
(الف)

$$X^2 \sim \frac{\alpha}{(k + \alpha)} \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{k}{(k + \alpha)} \Gamma\left(\frac{\beta + 1}{2}, \frac{1}{2}\right)$$

(ب)  $Y = |X| \sim RHBN(\alpha, \beta)$  که نماد  $RHBN(\alpha, \beta)$  بیانگر توزیع نیم نرمال چند مدی راست با تابع چگالی احتمال زیر است:

$$f(y; \beta) = \frac{\Upsilon(\alpha + y^\beta)}{\Gamma\left(\frac{\beta+1}{2}\right) \Upsilon^{\frac{\beta+1}{2}}} e^{-\frac{1}{2}y^2}, \quad y \geq 0 \quad (3.2)$$

(ج)  $Y = -|X| \sim LHBN(\alpha, \beta)$  که نماد  $LHBN(\alpha, \beta)$  بیانگر توزیع نیم نرمال چند مدی چپ با تابع چگالی احتمال زیر است:

$$f(y; \beta) = \frac{\Upsilon(\alpha + y^\beta)}{\Gamma\left(\frac{\beta+1}{2}\right) \Upsilon^{\frac{\beta+1}{2}}} e^{-\frac{1}{2}y^2}, \quad y \leq 0 \quad (4.2)$$

برای اثبات خواننده توجه دارد که تابع چگالی در اینجا آمیخته‌ای از توزیع نرمال و توزیع  $BN(\beta)$  است.

### مد در توزیع $BN(\alpha, \beta)$

برای به دست آوردن مدهای توزیع، کافی است مشتق تابع چگالی را برابر صفر قرار داده و پس از تعیین علامت مدها را به دست آورد. البته چون توزیع متقارن است و یکی از ریشه‌های معادله نیز همیشه صفر است واضح است که مدها نسبت به صفر متقارن و در مواردی که تعداد ریشه‌های معادله فرد باشد یکی از مدها صفر است. در حالت کلی جواب‌های غیر صفر معادله از حل معادله زیر بدست می‌آید:

$$\beta|x|^{\beta-2} - \alpha - |x|^\beta = 0$$

برای مثال وقتی  $\beta = 2$  و  $\alpha < 2$  دو مد عبارتند از:  $-\sqrt{2-\alpha}$  و  $\sqrt{2-\alpha}$  و برای  $\alpha \geq 2$  توزیع یک مد دارد که صفر است. برای  $\beta = 4$  و  $\alpha = 0$  توزیع دو مد در  $-2$  و  $2$  دارد و برای  $\alpha < 4$  سه مد به ترتیب عبارتند از:  $0$  و  $-\sqrt{\frac{4-\sqrt{16-4\alpha}}{2}}$  و  $\sqrt{\frac{4-\sqrt{16-4\alpha}}{2}}$  هستند و برای  $\alpha > 4$  یک مد دارد که صفر است. و

### روش تولید داده از توزیع $BN(\alpha, \beta)$

به دلیل مشکل بودن روش معکوس برای تولید داده از الگوریتم پذیرش و رد استفاده می‌کنیم. مراحل این روش عبارتند از:

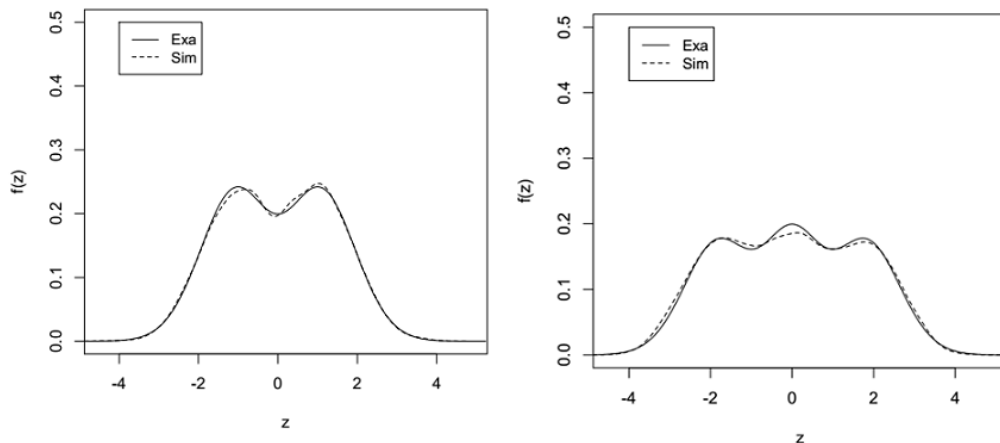
- یک عدد به تصادف از توزیع یکنواخت روی بازه صفر و یک انتخاب و آن را با  $U$  نشان می‌دهیم.
- یک عدد به تصادف از توزیع نرمال استاندارد انتخاب و آن را  $y$  در نظر می‌گیریم. با استفاده از یک عدد مثبت مانند  $c$  اگر  $u$  کمتر از نسبت  $\frac{g(x; \alpha, \beta)}{c\phi(x)}$  باشد  $y$  را به عنوان یک عدد از  $BN(\alpha, \beta)$  می‌پذیریم و در غیر این صورت مراحل ۱ و ۲ را آن قدر تکرار می‌کنیم تا داده‌ای انتخاب شود.

در شکل (۱) نمودارهای داده‌های شبیه‌سازی و واقعی برای توزیع‌های  $BN(1, 2)$  و  $BN(3, 4)$  نشان داده شده است.

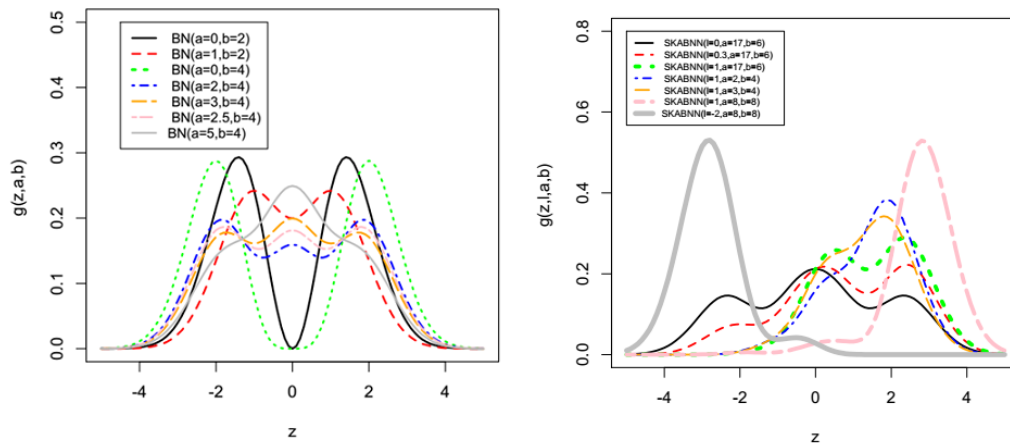
## ۳ توزیع چوله نرمال چند مدی

فرض کنید  $g(x; \alpha, \beta)$  تابع چگالی احتمال نرمال چند مدی  $BN(\alpha, \beta)$  باشد، در آن صورت بر اساس لم آزالینی [۶] تابع چگالی احتمال توزیع چوله نرمال چند مدی به صورت زیر معرفی می‌گردد

$$f(x; \alpha, \lambda, \beta) = \Upsilon g(x; \alpha, \beta) \Phi(\lambda x), \quad x \in R, \lambda \in R \quad (1.3)$$



شکل ۱: نمودار داده شبیه‌سازی و واقعی برای توزیع (سمت چپ  $BN(1, 2)$  و سمت راست  $BN(3, 4)$ )



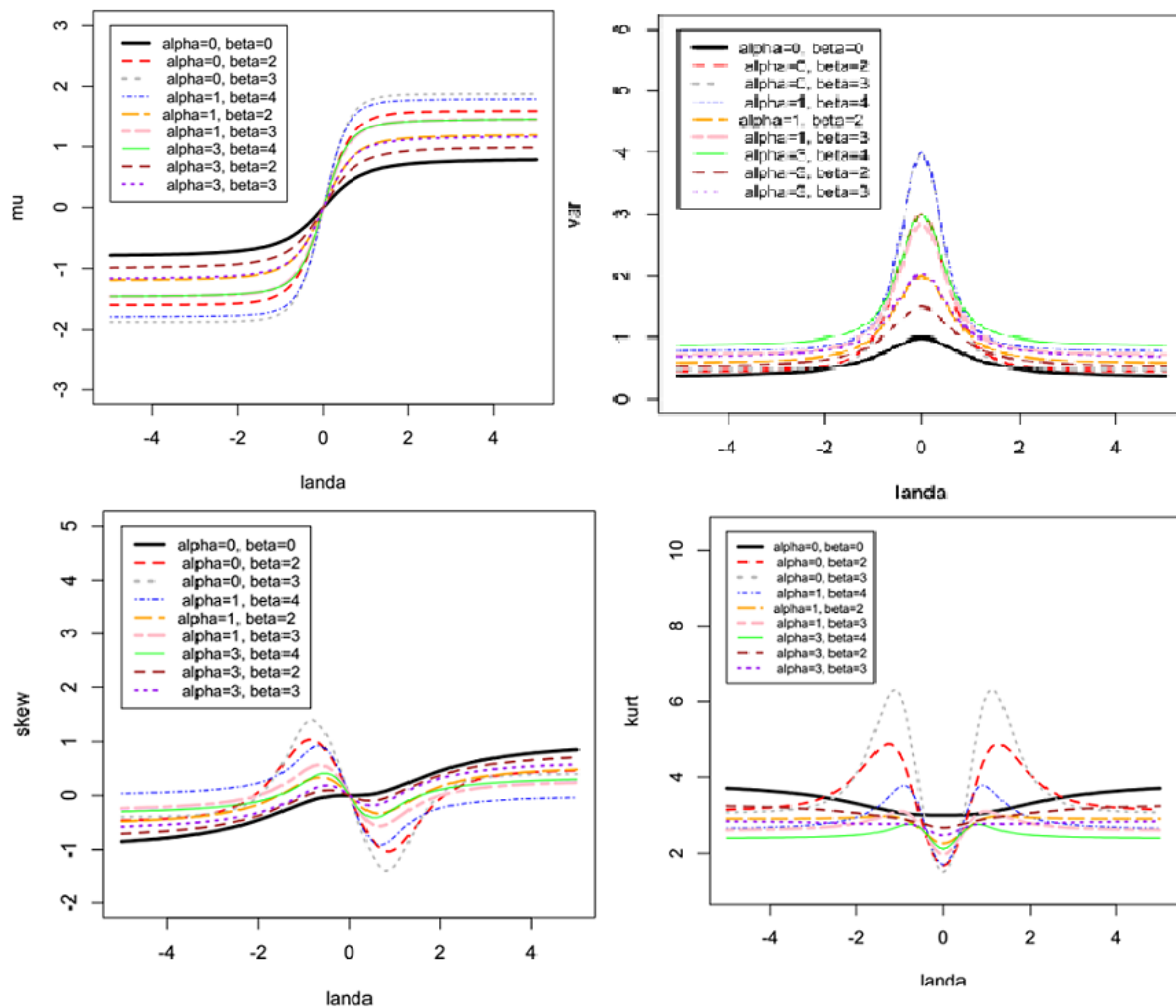
شکل ۲: نمودار تابع چگالی توزیع نرمال چند مدی (سمت چپ) و چوله نرمال چند مدی (سمت راست) برای مقادیر مختلف  $\lambda$  و  $\beta$

که در آن  $\lambda$  پارامتر چولگی است. این توزیع را با نماد  $SKABNN(\lambda, \alpha, \beta)$  نمایش می‌دهیم. با توجه به متقارن بودن  $g(x; \alpha, \beta)$  و رابطه  $\Phi(-\lambda x) = 1 - \Phi(\lambda x)$  داریم:

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x; \alpha, \lambda, \beta) dx &= \int_{-\infty}^{+\infty} \Psi g(x; \alpha, \beta) \Phi(\lambda x) dx \\ &= \int_{-\infty}^0 \Psi g(x; \alpha, \beta) \Phi(\lambda x) dx + \int_0^{+\infty} \Psi g(x; \alpha, \beta) \Phi(\lambda x) dx \\ &= \int_0^{+\infty} \Psi g(x; \alpha, \beta) dx = 1 \end{aligned}$$

بنابراین چگالی بودن  $f(x; \alpha, \lambda, \beta)$  تایید می‌شود. شکل (۲) تابع چگالی احتمال چند توزیع نرمال چند مدی و توزیع چوله نرمال چند مدی برای مقادیر مختلف  $\lambda$  و  $\beta$  نمایش داده شده است. دامنه تغییرات میانگین، واریانس، ضریب چولگی و ضریب کشیدگی نرمال چند مدی و توزیع چوله نرمال چند مدی نسبت به  $\lambda$  و به ازای چند مقدار  $\alpha$  و  $\beta$  در شکل (۳) نشان داده شده است.

همان طور که در شکل (۳) دیده می‌شود وقتی  $\alpha$  ثابت باشد دامنه ضریب چولگی با افزایش  $\beta$  نیز افزایش می‌یابد. اما برای  $\beta$  ثابت با افزایش  $\alpha$  کاهش پیدا می‌کند. قابل توجه است که نمودار ممتد سیاه تغییرات ضریب چولگی توزیع چوله نرمال آزالینی را نشان می‌دهد و قابلیت انعطاف توزیع نسبت به توزیع آزالینی به خوبی مشاهده می‌شود. این توزیع انعطاف پذیر می‌تواند داده‌های یک مدی و چند مدی، متقارن و چوله را مدل بندی کند.



شکل ۳: نمودار تغییرات شاخص‌های آماری در توزیع چوله نرمال چند مدی نسبت به  $\lambda$  به ازای چند مقدار  $\alpha$  و  $\beta$  (چپ بالا: میانگین)، واریانس (راست بالا: واریانس)، (چپ پایین: چولگی) و (راست پایین: کشیدگی)

توزیع‌های زیر حالات خاص  $SKABNN(\lambda, \alpha, \beta)$  هستند:

(الف)  $SKABNN(0, 0, 0)$  توزیع نرمال استاندارد است.

(ب)  $SKABNN(\lambda, 0, 0)$  توزیع چوله نرمال آزالینی است.

(ج)  $SKABNN(0, 0, \beta)$  توزیع نرمال دو مدی مرتبه  $\beta$  علوی ([۲]) است.

(د) اگر  $X \sim SKABNN(\lambda, 0, \beta)$  و  $\lambda \rightarrow \infty$ ، در آن صورت  $X \sim RHBN(\beta)$ .

(ه) اگر  $X \sim SKABNN(\lambda, 0, \beta)$  و  $\lambda \rightarrow -\infty$  در آن صورت  $X \sim LHBN(\beta)$ .

(و)  $X \sim SKABNN(\lambda, 0, 2)$  حالت خاصی است که علوی و طرهانی [۳] معرفی کردند.

قضیه ۱.۳. اگر  $G(x, \alpha, \beta)$  و  $F(x, \lambda, \alpha, \beta)$  به ترتیب تابع توزیع‌های  $BN(\alpha, \beta)$  و  $SKABNN(\lambda, \alpha, \beta)$  باشند، در آن صورت برای هر  $x$  و  $\lambda$  حقیقی مقدار داریم:

$$F(x; \lambda, \alpha, \beta) = G(x; \alpha, \beta) - H(x; \lambda, \alpha, \beta) \quad (۲.۳)$$

که

$$H(x; \lambda, \alpha, \beta) = \frac{1}{\sqrt{\pi}\Gamma(\frac{\beta+1}{2})} \sum_{i=0}^{\infty} \frac{(-1)^i \lambda^{i+1} \Gamma(\frac{\beta+2}{2} + i, \frac{1}{2}x^2)}{i!(1+2i)} \quad (۳.۳)$$

اثبات: بدون از دست دادن کلیت مسئله فرض می‌کنیم در (۲.۳)  $\lambda \geq 0$  و  $x \geq 0$  داریم:

$$F(x; \lambda, \alpha, \beta) = \int_{-\infty}^x \Psi g(t; \alpha, \beta) \Phi(\lambda t) dt = \int_{-\infty}^x \int_{-\infty}^{\lambda t} \Psi g(t; \alpha, \beta) \phi(u) du dt$$

و

$$\begin{aligned} \Psi G(x; \alpha, \beta) &= \int_{-\infty}^x \Psi g(t; \alpha, \beta) dt = \int_{-\infty}^x \Psi g(t; \alpha, \beta) \left[ \int_{-\infty}^{\lambda t} \phi(u) du + \int_{\lambda t}^{+\infty} \phi(u) du \right] dt \\ &= F(x; \lambda, \alpha, \beta) + \int_{-\infty}^x \int_{\lambda t}^{\infty} \Psi g(t; \alpha, \beta) \phi(u) du dt + \int_{-\infty}^x \int_0^{\infty} \Psi g(t; \alpha, \beta) \phi(u) du dt \\ &= F(x; \lambda, \alpha, \beta) + \int_{-\infty}^x \int_{\lambda t}^{\infty} \Psi g(t; \alpha, \beta) \phi(u) du dt + G(x; \alpha, \beta) \end{aligned}$$

آن‌گاه

$$\begin{aligned} F(x; \lambda, \alpha, \beta) &= G(x; \alpha, \beta) - \int_{-\infty}^x \int_{\lambda t}^{\infty} \Psi g(t; \alpha, \beta) \phi(u) du dt \\ &= G(x; \alpha, \beta) + \int_{-\infty}^x \int_0^{\lambda t} \Psi g(t; \alpha, \beta) \phi(u) du dt \\ &= G(x; \alpha, \beta) + \left[ \int_{-\infty}^{\infty} \int_0^{\lambda t} \Psi g(t; \alpha, \beta) \phi(u) du dt - \int_x^{\infty} \int_0^{\lambda t} \Psi g(t; \alpha, \beta) \phi(u) du dt \right] \\ &= G(x; \alpha, \beta) + \int_{-\infty}^{\infty} \Psi g(t; \alpha, \beta) [\Phi(\lambda t) - \Phi(x)] dt - H(x; \lambda, \alpha, \beta) \\ &= G(x; \alpha, \beta) - H(x; \lambda, \alpha, \beta) \end{aligned}$$

که

$$\begin{aligned} H(x; \lambda, \alpha, \beta) &= \int_x^{\infty} \int_0^{\lambda t} \Psi g(t) \phi(u) du dt = \int_x^{\infty} \frac{(\alpha + |t|)^{\beta} e^{-\frac{1}{2}t^2} \Gamma(\frac{1}{2}, \frac{1}{2}\lambda^2 t^2)}{\alpha + \Gamma(\frac{\beta+1}{2}) 2^{\frac{\beta+1}{2}} \sqrt{\pi}} dt \\ &= \int_x^{\infty} \frac{t^{\beta} e^{-\frac{1}{2}t^2}}{\alpha + \Gamma(\frac{\beta+1}{2}) 2^{\frac{\beta+1}{2}} \sqrt{\pi}} \sum_{i=0}^{\infty} \frac{(-1)^i (\frac{1}{2}\lambda^2 t^2)^{\frac{1}{2}+i}}{i!(\frac{1}{2}+i)} dt \end{aligned}$$

اگر تبدیل  $y = \frac{1}{\gamma} t^\gamma$  به کار رود، آنگاه با استفاده از  $\Gamma(\gamma, x) = \int_x^\infty e^{-t} t^{(\gamma-1)} dt$  روابط (۱) ۳.۳۸۱ و (۱) ۸.۳۵۴ در [۹]، رابطه (۹) به دست می آید. خواص زیر به سادگی برای تابع  $H(x; \lambda, \alpha, \beta)$  حاصل می شوند:

$$H(-x; \lambda, \alpha, \beta) = H(x; \lambda, \alpha, \beta) \quad (۴.۳)$$

$$H(x; -\lambda, \alpha, \beta) = -H(x; \lambda, \alpha, \beta) \quad (۵.۳)$$

برای مقادیر منفی  $x$  و  $\lambda$  با استفاده از  $\Phi(-\lambda x) = 1 - \Phi(\lambda x)$  و روابط (۳.۴) و (۳.۵) به راحتی از رابطه (۳.۲) تایید و در نتیجه اثبات قضیه ۳.۱ کامل می شود.

قضیه ۲.۳. فرض کنید  $X \sim SKABNN(\lambda, \alpha, \beta)$  و  $F(x; -\lambda, \alpha, \beta)$  تابع توزیع آن باشد، در آن صورت:

الف)  $-X \sim SKABNN(-\lambda, \alpha, \beta)$

ب)  $F(x; -\lambda, \alpha, \beta) = 1 - F(-x; \lambda, \alpha, \beta)$

ج)  $X^\gamma \sim \frac{\alpha}{k+\alpha} \Gamma(\frac{1}{\gamma}, \frac{1}{\gamma}) + \frac{k}{k+\alpha} \Gamma(\frac{\beta+1}{\gamma}, \frac{1}{\gamma})$

د)  $|X| \sim RHBNN(\alpha, \beta)$

ه)  $-|X| \sim LHBNN(\alpha, \beta)$

و) اگر  $X_1 \sim BN(\alpha, \beta)$  و  $h(\cdot)$  تابع زوجی باشد، در آن صورت  $h(X)$  و  $h(X_1)$  هم توزیع هستند

اثبات: اثبات موارد "الف" و "ب" ساده است و برای "ج" داریم:

$$\begin{aligned} P(X^\gamma \leq y) &= \int_{-y}^y f(x; \alpha, \lambda, \beta) dx = \int_{-y}^y \gamma g(x; \alpha, \beta) \Phi(\lambda x) dx \\ &= \int_{-y}^0 \gamma g(x; \alpha, \beta) \Phi(\lambda x) dx + \int_0^y \gamma g(x; \alpha, \beta) \Phi(\lambda x) dx \\ &= \int_0^y \gamma g(x; \alpha, \beta) dx \end{aligned}$$

با مشتق گیری نسبت به  $y$  اثبات "ج" حاصل می شود. بقیه موارد نیز شبیه "ج" است.

قضیه ۳.۳. میانگین، واریانس، ضریب چولگی و ضریب کشیدگی  $SKBNN(\lambda, \alpha, \beta)$  نسبت به  $\lambda$  به ترتیب فرد، زوج، فرد و زوج هستند.

اثبات: اگر  $X \sim SKBNN(\lambda, \alpha, \beta)$  میانگین و واریانس  $X$  عبارتند از:

$$E(X) = \mu_{SKBNN}(\lambda, \alpha, \beta) = \int_{-\infty}^{\infty} x^\gamma g(x; \alpha, \beta) \Phi(\lambda x) dx$$

$$\begin{aligned} E(X - \mu_{SKBNN}(\lambda, \alpha, \beta))^\gamma &= \sigma_{SKBNN}^\gamma(\lambda, \alpha, \beta) \\ &= \int_{-\infty}^{\infty} \gamma (x - \mu_{SKBNN}(\lambda, \alpha, \beta))^\gamma g(x; \alpha, \beta) \Phi(\lambda x) dx \end{aligned}$$

بنابراین با توجه به متقارن بودن  $g(\cdot; \alpha, \beta)$  و تغییر متغیر  $y = -x$

$$\begin{aligned} \mu_{SKBNN}(-\lambda, \alpha, \beta) &= \int_{-\infty}^{\infty} x^\gamma g(x; \alpha, \beta) \Phi(-\lambda x) dx \\ &= \int_{-\infty}^{\infty} -y^\gamma g(-y; \alpha, \beta) \Phi(\lambda y) dy = -\mu_{SKBNN}(\lambda, \alpha, \beta) \end{aligned}$$

با استدلال مشابه داریم:

$$\sigma_{SKBNN}^\gamma(-\lambda, \alpha, \beta) = \sigma_{SKBNN}^\gamma(\lambda, \alpha, \beta)$$

با توجه به تعریف ضریب چولگی داریم:

$$\begin{aligned}\alpha_{\text{rSBNN}}(\lambda, \alpha, \beta) &= \int_{-\infty}^{\infty} \left( \frac{x - \mu_{\text{SBNN}}(\lambda, \alpha, \beta)}{\sigma_{\text{SBNN}}(\lambda, \alpha, \beta)} \right)^{\text{r}} g(x; \alpha, \beta) \Phi(\lambda x) dx \\ \alpha_{\text{rSBNN}}(-\lambda, \alpha, \beta) &= \int_{-\infty}^{\infty} \left( \frac{x - \mu_{\text{SBNN}}(-\lambda, \alpha, \beta)}{\sigma_{\text{SBNN}}(-\lambda, \alpha, \beta)} \right)^{\text{r}} g(x; \alpha, \beta) \Phi(-\lambda x) dx \\ &= \int_{-\infty}^{\infty} \left( \frac{-y + \mu_{\text{SBNN}}(\lambda, \alpha, \beta)}{\sigma_{\text{SBNN}}(\lambda, \alpha, \beta)} \right)^{\text{r}} g(y; \alpha, \beta) \Phi(y) dy \\ &\quad - \alpha_{\text{rSBNN}}(\lambda, \alpha, \beta)\end{aligned}$$

برای ضریب کشیدگی داریم:

$$\alpha_{\text{fSBNN}}(\lambda, \alpha, \beta) = \int_{-\infty}^{\infty} \left( \frac{x - \mu_{\text{SBNN}}(\lambda, \alpha, \beta)}{\sigma_{\text{SBNN}}(\lambda, \alpha, \beta)} \right)^{\text{f}} g(x; \alpha, \beta) \Phi(\lambda x) dx$$

با استدلال مشابه داریم:

$$\alpha_{\text{fSBNN}}(-\lambda, \alpha, \beta) = \alpha_{\text{fSBNN}}(\lambda, \alpha, \beta)$$

و اثبات قضیه ۳.۳ کامل می‌شود.

#### ۴ روش‌های تولید داده از توزیع چوله نرمال چند مدی

در این بخش بر اساس روش‌های بیان شده توسط آزالینی [۶] و وانگ و همکاران [۱۸] روش‌هایی برای تولید نمونه از توزیع چوله نرمال چند مدی ارائه می‌شود.

قضیه ۱.۴: فرض کنید  $Z \sim N(0, 1)$  و  $Y \sim BN(\alpha, \beta)$  متغیرهای تصادفی مستقل باشند، در این صورت:  
 $X(Y | \lambda Y > Z) \sim SKABNN(\lambda, \alpha, \beta)$

اثبات:

$$P(Y \leq x | \lambda Y > Z) = \frac{P(Y \leq x, Z < \lambda Y)}{P(Z < \lambda Y)}$$

صورت کسر برابر است با:

$$\begin{aligned}P(Y \leq x, Z < \lambda Y) &= \int_{-\infty}^x P(Z < \lambda Y | Y = y) g(y; \alpha, \beta) dy \\ &= \int_{-\infty}^x \Phi(\lambda y) g(y; \alpha, \beta) dy.\end{aligned}$$

و مخرج کسر برابر است با:

$$P(Z < \lambda y) = \int_{-\infty}^{\infty} P(Z < \lambda Y | Y = y) g(y; \alpha, \beta) dy = \int_{-\infty}^{\infty} \Phi(\lambda y) g(y; \alpha, \beta) dy = \frac{1}{\text{r}}.$$

بنابراین:

$$F_X(x) = \text{r} \int_{-\infty}^x \Phi(\lambda y) g(y; \alpha, \beta) dy = \int_{-\infty}^x \text{r} g(y; \alpha, \beta) \Phi(\lambda y) dy$$

که تابع توزیع تجمعی  $SKABNN(\lambda, \alpha, \beta)$  است. بنابراین  $X \sim SKABNN(\lambda, \alpha, \beta)$  و اثبات کامل می‌شود.

روش اول تولید داده:

با استفاده از قضیه ۴.۱ می‌توان از توزیع  $SKABNN(\lambda, \beta)$  با اجرای مراحل زیر داده تولید کرد.



۱. تولید یک مقدار تصادفی مانند  $Z$  از توزیع نرمال استاندارد

۲. تولید یک مقدار تصادفی مانند  $Y$  از  $BN(\alpha, \beta)$

۳. اگر  $X = Y$ ,  $\lambda Y > Z$  در غیر این صورت مراحل ۱ و ۲ را تا حصول داده  $X$  تکرار می کنیم.

اگر سه مرحله فوق  $n$  بار به طور مستقل تکرار شوند یک نمونه تصادفی به حجم  $n$  از  $SKABNN(\lambda, \alpha, \beta)$  تولید می شود.

قضیه ۲.۴. فرض کنید  $Z \sim N(0, 1)$  و  $Y \sim BN(\alpha, \beta)$  متغیرهای تصادفی مستقل باشند، اگر متغیر تصادفی  $X$  را به صورت زیر تعریف کنیم:

$$X = \begin{cases} Y, & \text{اگر } Z \leq \lambda Y \\ -Y, & \text{اگر } Z > \lambda Y \end{cases}$$

آن گاه  $X \sim SKABNN(\lambda, \alpha, \beta)$

اثبات:

تابع توزیع  $X$  به صورت زیر است:

$$\begin{aligned} & \int_{-\infty}^x P(Z \leq \lambda Y | Y = y)g(y; \alpha, \beta)dy + \int_{-\infty}^x P(Z > \lambda Y | -Y = y)g(-y; \alpha, \beta)dy \\ &= \int_{-\infty}^x \Phi(\lambda y)g(y; \alpha, \beta)dy + \int_{-\infty}^x [1 - \Phi(-\lambda y)]g(y; \alpha, \beta)dy = \int_{-\infty}^x 2g(y; \alpha, \beta)\Phi(\lambda y)dy. \end{aligned}$$

اثبات با مشتق گرفتن نسبت به  $x$  تکمیل می شود.

روش دوم تولید داده

با استفاده از قضیه ۴.۲ و اجرای مراحل زیر می توان از توزیع  $SKABNN(\lambda, \alpha, \beta)$  یک داده تولید کرد.

۱. تولید یک مقدار تصادفی مانند  $Z$  از توزیع نرمال استاندارد

۲. تولید یک مقدار تصادفی مانند  $Y$  از  $BN(\alpha, \beta)$

۳. اگر  $X = Y$ ,  $Z \leq \lambda Y$  در غیر این صورت  $X = -Y$  اگر سه مرحله فوق  $n$  بار به طور مستقل تکرار شوند یک نمونه تصادفی به حجم  $n$  از  $SKABNN(\lambda, \alpha, \beta)$  تولید می شود.

دو قضیه زیر براساس ایده وانگ و همکاران ([۱۸]) که توزیع یکنواخت را جایگزین توزیع نرمال کردند نتیجه می شود.

قضیه ۳.۴. فرض کنید  $Z \sim U(0, 1)$  و  $Y \sim BN(\alpha, \beta)$  متغیرهای تصادفی مستقل باشند، در این صورت:

$$X(Y|Z \leq \Phi(\lambda Y)) \sim SKABNN(\lambda, \alpha, \beta)$$

اثبات: داریم

$$\begin{aligned} F_X(x) &= P(Y \leq x | Z \leq \Phi(\lambda Y)) = \frac{P(Y \leq x, Z \leq \Phi(\lambda Y))}{P(Z \leq \Phi(\lambda Y))} \\ &= \frac{\int_{-\infty}^x P(Z \leq \Phi(\lambda Y) | Y = y)g(y; \alpha, \beta)dy}{\int_{-\infty}^{\infty} P(Z \leq \Phi(\lambda Y) | Y = y)g(y; \alpha, \beta)dy} = \int_{-\infty}^x 2g(y; \alpha, \beta)\Phi(\lambda y)dy \end{aligned}$$

اثبات با مشتق گرفتن نسبت به  $x$  کامل می شود.

روش سوم تولید داده

با استفاده از قضیه ۴.۳ می توان از توزیع  $SKABNN(\lambda, \alpha, \beta)$  با اجرای مراحل زیر داده تولید کرد.

۱. تولید یک مقدار تصادفی مانند  $Z$  از توزیع یکنواخت روی بازه  $(0, 1)$

۲. تولید یک مقدار تصادفی مانند  $Y$  از  $BN(\alpha, \beta)$

۳. اگر  $X = Y, Z \leq \Phi(Y)$  در غیر این صورت مراحل ۱ و ۲ را تا حصول داده  $X$  تکرار می‌کنیم. اگر سه مرحله فوق  $n$  بار به طور مستقل تکرار شوند، یک نمونه تصادفی به حجم  $n$  از  $SKABNN(\lambda, \alpha, \beta)$  تولید می‌شود.

قضیه ۴.۴. فرض کنید  $Z \sim U(0, 1)$  و  $Y \sim BN(\alpha, \beta)$  متغیرهای تصادفی مستقل باشند، اگر متغیر تصادفی  $X$  را به صورت زیر تعریف کنیم:

$$X = \begin{cases} Y, & Z \leq \Phi(\lambda Y) \\ -Y, & Z > \Phi(\lambda Y) \end{cases}$$

آن گاه  $X \sim SKABNN(\lambda, \alpha, \beta)$

اثبات:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x P(Z \leq \phi(\lambda Y) | Y = y) g(y; \alpha, \beta) dy \\ &+ \int_{-\infty}^x P(Z \geq \Phi(\lambda Y) | Y = y) g(-y; \alpha, \beta) dy \\ &= \int_{-\infty}^x \Phi(\lambda y) g(y; \alpha, \beta) dy + \int_{-\infty}^x (1 - \Phi(-\lambda y)) g(y; \alpha, \beta) dy \end{aligned}$$

اثبات توسط مشتق گرفتن  $F_X(x)$  نسبت به  $x$  کامل می‌شود.

روش چهارم تولید داده

با استفاده از قضیه ۴.۴ و اجرای مراحل زیر می‌توان از توزیع  $SKABNN(\lambda, \alpha, \beta)$  یک داده تولید کرد.

۱. تولید یک مقدار تصادفی مانند  $Z$  از توزیع یکنواخت روی بازه  $(0, 1)$

۲. تولید یک مقدار تصادفی مانند  $Y$  از  $BN(\alpha, \beta)$

۳. اگر  $X = Y, Z \leq \Phi(\lambda Y)$  در غیر این صورت  $X = -Y$ .

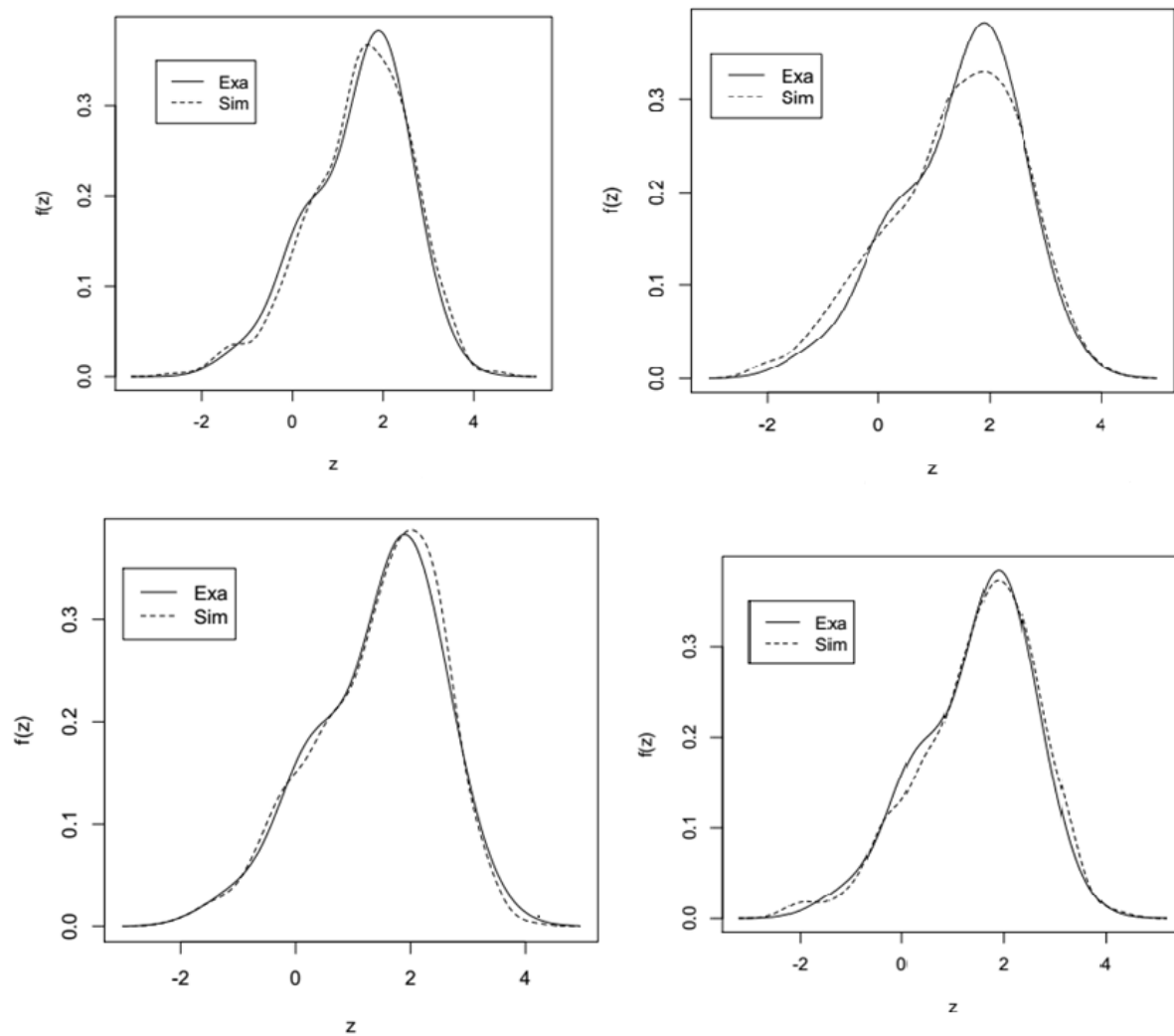
اگر سه مرحله  $n$  بار به طور مستقل تکرار شوند، یک نمونه تصادفی به حجم  $n$  از  $SKABNN(\lambda, \alpha, \beta)$  تولید می‌شود. شکل (۴) نمودارهای چگالی واقعی و شبیه‌سازی شده با چهار روش مذکور از توزیع  $SKABNN(1, 2, 4)$  را نشان می‌دهد.

## ۵ خانواده مکانی - مقیاسی توزیع چوله نرمال چند مدی

اگر  $Z \sim SKABNN(\lambda, \alpha, \beta)$  و  $\sigma$  یک عدد حقیقی مثبت باشد با استفاده از تبدیل  $X = \mu + \sigma Z$  متغیر تصادفی  $X$  دارای توزیع چوله نرمال چند مدی با پارامتر مکان  $\mu$  و پارامتر مقیاس  $\sigma$  است که با نماد  $SKABNN(\lambda, \alpha, \beta, \mu, \sigma)$  نمایش داده می‌شود. تابع چگالی احتمال  $X$  به صورت

$$\begin{aligned} f(x; \lambda, \alpha, \beta, \mu, \sigma) &= \frac{2}{\sigma} g\left(\frac{x - \mu}{\sigma}; \alpha, \beta\right) \Phi\left(\lambda \frac{x - \mu}{\sigma}\right) \\ &= \frac{2(\alpha + |x - \mu|^\beta) e^{-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2} \Phi\left(\lambda \frac{x - \mu}{\sigma}\right)}{\alpha + \Gamma\left(\frac{\beta+1}{2}\right) 2^{\frac{\beta+1}{2}} \sigma^{\beta+1}} \end{aligned}$$

می‌باشد. این خانواده مکانی - مقیاسی توزیع چوله نرمال چند مدی، انعطاف پذیر می‌باشد و انواع داده‌های یک، دو و سه مدی متقارن و چوله را پوشش می‌دهد.



شکل ۴: نمودار تابع چگالی احتمالی واقعی و شبیه‌سازی  $SKABNN(1, 2, 4)$  با چهار روش (چپ بالا: روش اول)، (راست بالا: روش دوم)، (چپ پایین: روش سوم) و (راست پایین: روش چهارم)

اگر  $x_1, x_2, \dots, x_n$  مقادیر محقق شده یک نمونه تصادفی از  $SKABNN(\lambda, \alpha, \beta, \mu, \sigma)$  باشند در آن صورت لگاریتم تابع درستنمایی به صورت زیر می‌باشد:

$$\log L(\mu, \sigma, \lambda, \beta) = n \log \gamma - n \log \sigma + n \Gamma\left(\frac{1}{\gamma}\right) - n \log \left( \alpha + \Gamma\left(\frac{\beta + 1}{\gamma}\right) \gamma^{\frac{\beta}{\gamma}} \right) \\ + \sum_{i=1}^n \log \left( \alpha + \left| \frac{x_i - \mu}{\sigma} \right|^{\beta} \right) - \frac{1}{\gamma \sigma^{\gamma}} \sum_{i=1}^n (x_i - \mu)^{\gamma} + \sum_{i=1}^n \log \Phi\left(\lambda \frac{x_i - \mu}{\sigma}\right)$$

برآورد ماکسیمم درستنمایی از حل معادلاتی که از مشق‌های جزئی  $\log L(\mu, \sigma, \lambda, \beta)$  نسبت به تک تک پارامترها و معادل صفر قرار دادن توأم آنها حاصل می‌شود، به دست می‌آیند. با توجه به غیر خطی بودن معادلات از روش‌های عددی برآوردها قابل محاسبه هستند و لذا از ذکر معادلات پیچیده آن‌ها خوداری شده است. با استفاده از دستور `optim` یا `nlnmb` در بسته نرم افزار `R` می‌توان مقادیر برآورد پارامترها را برای یک مجموعه داده عددی معلوم محاسبه نمود.

وجود برآورد ماکسیمم درستنمایی  $\lambda$  در توزیع نرمال چوله آزالینی توسط لژیو [۱۰] و پیوسی [۱۵] بررسی شده است. آنها بیان کرده‌اند که اگر حداقل دو داده در مجموعه داده‌های نمونه علامت متفاوت داشته باشند، برآورد  $\lambda$  وجود دارد. مارتینز و همکاران [۱۱] نشان دادند که این شرط برای هر کلاس از توزیع‌ها با تابع چگالی احتمال به صورت زیر نیز صادق است:

$$f(x) = \gamma h(x) \Phi(\lambda x)$$

که  $h(x)$  تابع چگالی احتمال متقارنی حول صفر است. بنابراین چون تابع چگالی چوله نرمال چند مدی متعلق به این خانواده است پس اگر حداقل دو داده علامت مختلف داشته باشند، برآورد پارامتر  $\lambda$  وجود خواهد داشت.

## ۶ شبیه سازی و کاربرد در داده‌های واقعی

در این بخش با استفاده از یک مطالعه شبیه‌سازی با در نظر گرفتن مراحل زیر برآوردهای ماکسیمم درستنمایی پارامترهای توزیع چوله نرمال چند مدی ارزیابی می‌شوند:

۱. نمونه‌هایی به حجم  $n$  (۲۰۰ و ۵۰۰) از توزیع  $SKABNN(\mu, \sigma, \lambda, \alpha, \beta)$  با پارامترهای معلوم تولید می‌شود.

۲. برای هر نمونه در مرحله ۱ برآوردهای ماکسیمم درستنمایی پارامترهای توزیع با استفاده از دستور `optim` به روش `L-BFGS-B` در نرم افزار `R` محاسبه می‌شوند. نتایج شبیه‌سازی برای مقادیر مختلفی از پارامترها و حجم نمونه در جداول ۱ و ۲ نشان داده شده‌اند. با توجه به این دو جدول با افزایش حجم نمونه برآوردها ناریب و خطای معیار ناچیزی دارند. برآوردهای مکان و مقیاس در مقایسه با دیگر پارامترها دقیق‌تر هستند. لازم به ذکر است که برای دیگر تعداد نمونه شبیه‌سازی انجام شد و نتایج مشابه به دست آمد.

### کاربرد در داده‌های واقعی

در این بخش توزیع چوله نرمال چند مدی را به مجموعه داده‌های معروف آفشان (چشمه‌های آب گرم) برازش داده و با چند توزیع رقیب مقایسه می‌شوند. این داده‌ها توسط محققین زیادی از جمله آزالینی و باومن [۴]، علوی و طرهانی [۳]، راسخی و همکاران [۱۶] و غیره استفاده شده‌اند. این مجموعه داده در "MASS" در بسته نرم افزار `R` به نام "faithful" موجود است. داده‌ها بیانگر ۲۹۹ جفت اندازه‌گیری دو متغیر فاصله زمانی بین شروع فوران‌های پی در پی (زمان انتظار) و مدت زمان فوران بعدی (مدت زمان) آفشان‌های قدیمی پارک ملی سنگ زرد در ایالات متحده هستند. در این مقاله، متغیر مدت زمان مطالعه شده است. در جدول ۳ توصیف مجموعه داده‌ها آمده است. بافت‌نگار داده‌ها در شکل (۴) نشان می‌دهد که داده‌ها دو مدی هستند. جدول ۴ برآورد پارامترهای مدل پیشنهادی (`SKABNN`) و پنج مدل رقیب شامل مدل (`ABFN`) بولفرینیت و همکاران [۱۷]، مدل (`SBN`) علوی و طرهانی [۳]، مدل (`SSCN`) راسخی و همکاران [۱۶]، مدل (`SN`) چوله نرمال آزالینی [۵] و مدل آمیخته دو نرمال (`Mixed-N`) را نشان می‌دهد. شکل (۵) نمودار تابع چگالی احتمال توزیع پیشنهادی و توزیع‌های رقیب به همراه بافت‌نگار داده‌ها را نشان می‌دهد. شکل (۵) بیانگر این است که مدل پیشنهادی به خوبی به داده‌ها برازش شده است. در جدول ۵ نیز با سه معیار `AIC`، `BIC` و `AICC` مدل پیشنهادی با مدل‌های رقیب مقایسه شده‌اند که نشان می‌دهد مدل پیشنهادی بهترین مدل برازش شده در بین مدل‌های رقیب است.

## ۷ نتیجه‌گیری

در مقاله حاضر توزیع چوله نرمال چند مدی معرفی و خواص آن مورد بررسی قرار گرفته است. این توزیع تعمیمی مناسب از توزیع نرمال برای توزیع‌های متقارنی است که ممکن است بیش از یک مد داشته باشند. روش‌هایی برای تولید داده از این توزیع، پیشنهاد شده و برآورد

جدول ۱: میانگین و خطای معیار شبیه سازی برآورد پارامترها برای حجم نمونه ۲۰۰

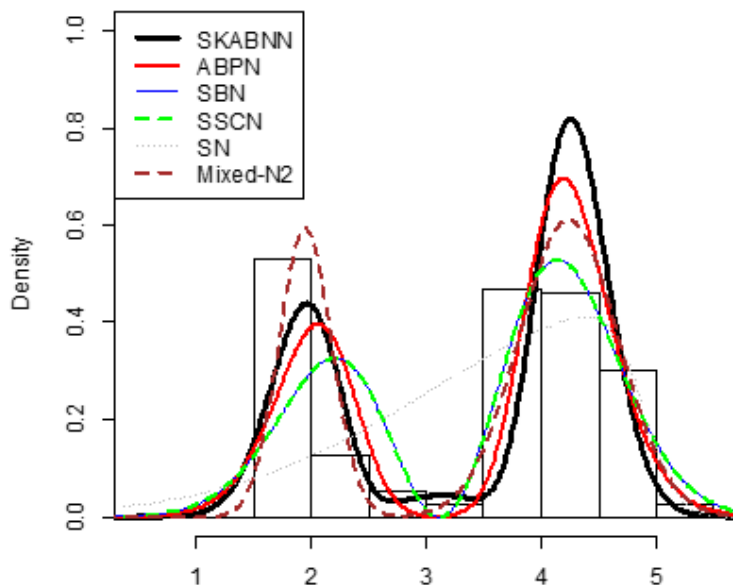
$SD[\hat{\sigma}]$	$E[\hat{\sigma}]$	$SD[\hat{\mu}]$	$E[\hat{\mu}]$	$SD[\hat{\lambda}]$	$E[\hat{\lambda}]$	$SD[\hat{\beta}]$	$E[\hat{\beta}]$	$SD[\hat{\alpha}]$	$E[\hat{\alpha}]$	$\sigma$	$\mu$	$\lambda$	$\beta$	$\alpha$
./۰۷۸۸	./۹۷۴۷	./۱۲۳۵	./۰۰۰۸	./۲۰۱۶	./۰۰۲۰	./۴۴۷۰	./۱۱۵۵	./۵۳۵۱	./۰۷۴۷	۱	.	.	.	.
./۰۸۳۱	./۹۷۱۶	./۱۲۹۰	./۰۰۰۴	./۱۹۹۶	./۰۰۰۶	./۵۵۳۲	./۱۵۹۱	./۶۹۷۳	./۷۴۲۲	۱	.	.	.	۱
./۰۵۵۶	./۹۸۶۰	./۰۴۷۳	./۰۰۰۴	./۰۵۵۱	./۰۰۰۰	./۳۲۶۳	۲/۱۲۱۲	./۰۲۲۲	./۰۰۷۵	۱	.	.	۲	.
./۱۳۷۶	./۹۹۹۹	./۲۱۰۵	./۰۰۳۰	./۱۴۸۰	./۰۰۱۹	۱/۲۲۴۴	۲/۱۳۷۹	۱/۲۸۵۴	۱/۲۲۴۱	۱	.	.	۲	۱
./۱۱۰۴	./۹۹۲۹	./۱۸۲۸	./۰۳۵۲	./۴۶۰۲	۱/۱۱۷۷	./۳۶۰۸	./۱۱۲۵	./۳۹۲۸	./۰۴۹۱	۱	.	۱	.	.
./۱۱۹۶	./۹۸۷۳	./۲۰۰۰	./۰۲۸۳	./۴۸۵۸	۱/۰۹۹۷	./۴۸۷۵	./۱۵۵۳	./۶۸۴۰	./۵۷۰۶	۱	.	۱	.	۱
./۰۵۶۹	./۹۸۶۹	./۰۶۳۸	./۰۰۴۲	./۱۲۸۱	۱/۰۰۳۷	./۳۴۵۸	۲/۱۲۴۴	./۰۲۲۳	./۰۰۶۷	۱	.	۱	۲	.
./۱۵۹۲	۱/۰۰۳۸	./۳۵۹۲	./۰۰۰۹	./۴۴۳۱	۱/۰۲۷۱	۱/۲۳۶۲	۱/۹۵۳۷	۱/۲۷۷۲	./۹۹۳۴	۱	.	۱	۲	۱
./۰۷۸۰	./۹۷۴۷	./۱۱۹۵	./۹۹۹۸	./۱۹۴۶	./۰۰۱۰	./۴۲۷۲	./۱۱۳۴	./۴۷۹۰	./۰۷۰۶	۱	۱	.	.	.
./۰۸۲۳	./۹۷۰۷	./۱۲۶۷	۱/۰۰۲۲	./۱۹۸۰	./۰۰۲۳	./۵۴۰۸	./۱۵۶۳	./۷۱۲۶	./۷۲۸۴	۱	۱	.	.	۱
./۰۵۵۸	./۹۸۵۱	./۰۴۷۶	./۹۹۹۹	./۰۵۴۷	./۰۰۰۳	./۳۲۴۱	۲/۱۲۶۸	./۰۲۲۳	./۰۰۷۵	۱	۱	.	۲	.
./۱۳۶۶	۱/۰۰۱۱	./۲۱۱۷	./۹۹۹۲	./۱۴۹۸	./۰۰۰۸	۱/۲۳۰۴	۲/۱۳۴۰	۱/۲۸۵۵	۱/۲۲۵۶	۱	۱	.	۲	۱
./۱۱۱۹	./۹۹۳۲	./۱۸۲۲	./۹۶۴۳	./۴۶۸۸	۱/۱۲۲۰	./۳۴۰۶	./۱۰۷۶	./۳۸۷۵	./۰۴۵۹	۱	۱	۱	.	.
./۱۱۶۸	./۹۸۷۱	./۱۹۱۹	./۹۷۰۵	./۴۸۴۱	۱/۱۰۴۵	./۴۶۲۴	./۱۴۹۲	./۶۵۳۱	./۵۶۵۹	۱	۱	۱	.	۱
./۰۵۷۲	./۹۸۶۸	./۰۶۵۷	./۹۹۳۳	./۱۲۹۰	۱/۰۰۳۷	./۳۵۲۰	۲/۱۳۵۵	./۰۲۲۴	./۰۰۶۸	۱	۱	۱	۲	.
./۱۵۹۱	۱/۰۰۵۸	./۳۶۶۰	۱/۰۰۱	./۴۷۶۴	۱/۰۲۹۴	۱/۲۵۰۵	۱/۹۴۹۲	۱/۳۱۰۰	۱/۰۱۰۹	۱	۱	۱	۲	۱
./۱۷۸۷	۱/۹۵۷۷	./۳۳۲۷	./۰۰۳۵	./۲۸۳۰	./۰۰۲۱	./۴۸۸۲	./۱۳۳۵	./۵۵۱۷	./۰۹۵۳	۲	.	.	.	.
./۱۹۲۴	۱/۹۵۸۶	./۳۷۴۸	./۰۰۰۶	./۳۱۹۰	./۰۰۱۳	./۵۷۹۶	./۱۷۵۴	./۷۴۵۳	./۷۰۷۵	۲	.	.	.	۱
./۱۱۲۰	۱/۹۷۳۸	./۰۹۵۱	./۰۰۰۰	./۰۵۴۵	./۰۰۰۳	./۳۲۳۵	۲/۱۱۷۷	./۰۲۱۳	./۰۰۷۴	۲	.	.	۲	.
./۲۷۳۰	۱/۹۹۵۴	./۴۲۳۱	./۰۰۴۵	./۱۴۷۶	./۰۰۱۰	۱/۲۳۳۹	۲/۱۶۲۷	۱/۲۹۶۷	۱/۲۴۱۶	۲	.	.	۲	۱
./۲۲۷۶	۱/۹۸۴۱	./۳۶۳۷	./۰۷۷۱	./۴۶۷۵	۱/۱۲۰۹	./۴۶۰۸	./۱۳۸۴	./۵۳۴۷	./۰۸۱۵	۲	.	۱	.	.
./۲۳۷۱	۱/۹۷۶۵	./۳۹۸۰	./۰۶۶۱	./۴۹۳۵	۱/۱۱۱۶	./۵۲۴۲	./۱۶۴۶	./۷۱۱۸	./۵۴۲۲	۲	.	۱	.	۱
./۱۱۶۰	۱/۹۷۲۵	./۱۲۹۶	./۰۰۸۵	./۱۲۸۷	۱/۰۰۱۱	./۳۶۱۳	۲/۱۳۰۳	./۰۴۷۳	./۰۰۷۵	۲	.	۱	۲	.
./۳۱۶۷	۲/۰۰۶۰	./۶۸۹۹	./۰۰۸۶	./۴۳۲۱	۱/۰۲۰۱	۱/۲۵۲۴	۱/۹۶۶۹	۱/۳۳۳۹	۱/۰۴۴۶	۲	.	۱	۲	۱
./۱۷۹۶	۱/۹۵۸۳	./۳۳۵۸	./۹۹۹۵	./۲۹۱۴	./۰۰۲۵	./۵۰۸۹	./۱۳۹۷	./۵۸۳۲	./۱۰۰۸	۲	۱	.	.	.
./۱۹۶۵	۱/۹۵۹۶	./۳۷۲۶	۱/۰۰۱۶	./۳۲۶۳	./۰۰۰۲	./۵۶۰۹	./۱۶۷۴	./۷۵۷۶	./۷۰۶۷	۲	۱	.	.	۱
./۱۱۲۷	۱/۹۷۱۹	./۰۹۵۵	۱/۰۰۱۰	./۰۵۴۳	./۰۰۰۲	./۳۲۹۱	۲/۱۲۱۷	./۰۲۱۶	./۰۰۷۳	۲	۱	.	۲	.
./۲۷۴۱	۲/۰۰۲۱	./۴۱۸۷	./۹۹۳۴	./۱۴۹۴	./۰۰۲۳	۱/۲۱۶۴	۲/۱۲۷۱	۱/۲۹۳۳	۱/۲۲۵۷	۲	۱	.	۲	۱
./۲۲۲۶	۱/۹۸۵۷	./۳۶۶۱	./۹۲۳۶	./۴۷۳۵	۱/۱۲۰۷	۴۲۹۴/	۰./۱۲۷۱	./۴۷۲۷	./۰۶۷۷	۲	۱	۱	.	.
./۲۳۹۹	۱/۹۷۶۲	./۳۸۹۶	./۹۳۴۶	./۴۹۲۰	۱/۱۱۴۶	./۵۳۱۱	./۱۶۵۹	./۷۰۹۱	./۵۳۲۰	۲	۱	۱	.	۱
./۱۱۴۴	۱/۹۷۴۶	./۱۲۸۰	./۹۹۲۵	./۱۳۰۶	۱/۰۰۳۶	./۳۴۵۷	۲/۱۲۳۳	./۰۲۱۱	./۰۰۶۹	۲	۱	۱	۲	.
./۳۲۰۶	۲/۰۰۶۶	./۷۰۲۷	./۹۹۷۰	./۴۴۸۴	۱/۰۲۴۵	۱/۲۶۴۰	۱/۹۸۵۹	۱/۳۷۸۹	۱/۰۵۷۸	۲	۱	۱	۲	۱

جدول ۲: میانگین و خطای معیار شبیه‌سازی برآورد پارامترها برای حجم نمونه ۵۰۰

$SD[\hat{\sigma}]$	$E[\hat{\sigma}]$	$SD[\hat{\mu}]$	$E[\hat{\mu}]$	$SD[\hat{\lambda}]$	$E[\hat{\lambda}]$	$SD[\hat{\beta}]$	$E[\hat{\beta}]$	$SD[\hat{\alpha}]$	$E[\hat{\alpha}]$	$\sigma$	$\mu$	$\lambda$	$\beta$	$\alpha$
۰/۰۴۱۳	۰/۹۸۴۳	۰/۰۴۸۰	۰/۰۰۰۴	۰/۰۵۵۸	۰/۰۰۰۰	۰/۲۲۲۴	۰/۰۴۷۸	۰/۲۹۸۳	۰/۰۳۰۳	۱	.	.	.	.
۰/۰۴۳۹	۰/۹۸۳۱	۰/۰۶۱۸	۰/۰۰۰۰	۰/۰۷۴۰	۰/۰۰۱۰	۰/۲۹۳۵	۰/۰۷۵۷	۰/۴۸۵۵	۰/۸۳۸۰	۱	.	.	.	۱
۰/۰۳۴۶	۰/۹۹۳۰	۰/۰۲۸۷	۰/۰۰۰۰	۰/۰۳۴۷	۰/۰۰۰۱	۰/۱۸۸۷	۲/۰۵۴۶	۰/۰۰۸۶	۰/۰۰۳۳	۱	.	.	۲	.
۰/۰۸۶۸	۰/۹۹۹۸	۰/۱۱۸۰	۰/۰۰۱۲	۰/۰۸۴۰	۰/۰۰۰۳	۰/۷۲۱۸	۲/۰۴۹۳	۰/۵۶۳۰	۱/۰۶۹۷	۱	.	.	۲	۱
۰/۰۷۳۳	۰/۹۹۹۷	۰/۱۲۲۵	۰/۰۲۵۸	۰/۲۸۷۵	۱/۰۶۹۱	۰/۱۸۵۰	۰/۰۵۲۱	۰/۲۶۴۳	۰/۰۲۵۰	۱	.	۱	.	.
۰/۰۷۸۵	۰/۹۹۶۵	۰/۱۳۹۲	۰/۰۲۰۳	۰/۳۱۴۹	۱/۰۵۶۶	۰/۲۷۵۷	۰/۰۷۹۲	۰/۵۳۶۸	۰/۷۵۳۵	۱	.	۱	.	۱
۰/۰۳۵۵	۰/۹۹۲۹	۰/۰۳۷۰	۰/۰۰۱۶	۰/۰۷۸۲	۰/۹۹۸۵	۰/۱۹۹۳	۲/۰۶۱۷	۰/۰۰۸۶	۰/۰۰۳۳	۱	.	۱	۲	.
۰/۱۰۲۷	۱/۰۰۶۲	۰/۲۸۲۰	۰/۰۰۹۴	۰/۲۹۶۴	۱/۰۱۴۷	۰/۸۷۷۴	۱/۹۵۶۵	۰/۷۲۲۹	۰/۹۶۹۲	۱	.	۱	۲	۱
۰/۰۴۱۰	۰/۹۸۴۰	۰/۰۵۰۶	۰/۹۹۹۷	۰/۰۵۷۷	۰/۰۰۰۱	۰/۲۳۲۳	۰/۰۴۹۷	۰/۳۱۲۰	۰/۰۳۲۲	۱	۱	.	.	.
۰/۰۴۴۵	۰/۹۸۲۸	۰/۰۵۶۰	۱/۰۰۰۶	۰/۰۶۴۹	۰/۰۰۰۳	۰/۲۹۳۵	۰/۰۷۷۴	۰/۴۸۸۸	۰/۸۴۲۲	۱	۱	.	.	۱
۰/۰۳۴۷	۰/۹۹۳۰	۰/۰۲۸۹	۱/۰۰۰۱	۰/۰۳۴۶	۰/۰۰۰۲	۰/۱۸۸۹	۲/۰۵۷۵	۰/۰۰۸۷	۰/۰۰۳۵	۱	۱	.	۲	.
۰/۰۸۷۱	۱/۰۰۰۴	۰/۱۱۷۰	۰/۹۹۹۰	۰/۰۸۳۷	۰/۰۰۰۰	۰/۷۲۰۸	۲/۰۴۳۰	۰/۵۵۷۸	۱/۰۶۵۴	۱	۱	.	۲	۱
۰/۰۷۳۹	۰/۹۹۹۴	۰/۱۲۵۳	۰/۹۷۵۰	۰/۲۹۰۰	۱/۰۶۸۴	۰/۲۱۰۳	۰/۰۵۳۲	۰/۳۳۰۵	۰/۰۲۹۴	۱	۱	۱	.	.
۰/۰۷۸۱	۰/۹۹۵۱	۰/۱۳۵۴	۰/۹۷۹۲	۰/۳۰۴۹	۱/۰۵۷۲	۰/۲۸۷۱	۰/۰۸۲۷	۰/۵۵۶۷	۰/۷۶۲۷	۱	۱	۱	.	۱
۰/۰۳۵۸	۰/۹۹۳۲	۰/۰۳۶۹	۰/۹۹۷۹	۰/۰۷۸۴	۰/۹۹۹۰	۰/۱۹۷۸	۲/۰۵۹۷	۰/۰۰۸۶	۰/۰۰۳۳	۱	۱	۱	۲	.
۰/۱۰۱۵	۱/۰۰۷۰	۰/۲۷۶۸	۰/۹۹۴۹	۰/۲۹۰۵	۱/۰۱۳۶	۰/۸۶۷۳	۱/۹۴۲۹	۰/۶۶۷۱	۰/۹۵۷۹	۱	۱	۱	۲	۱
۰/۱۰۰۸	۱/۹۷۱۸	۰/۱۷۵۱	۰/۰۰۱۰	۰/۱۴۰۸	۰/۰۰۱۷	۰/۳۰۴۸	۰/۰۶۷۳	۰/۴۲۳۰	۰/۰۶۰۲	۲	.	.	.	.
۰/۱۰۸۵	۱/۹۶۹۴	۰/۱۹۶۸	۰/۰۰۳۳	۰/۱۵۵۳	۰/۰۰۱۴	۰/۳۷۰۲	۰/۰۹۳۶	۰/۵۸۸۴	۰/۸۳۳۸	۲	.	.	.	۱
۰/۰۶۹۸	۱/۹۸۷۶	۰/۰۵۷۸	۰/۰۰۰۰	۰/۰۳۴۱	۰/۰۰۰۵	۰/۱۸۸۸	۲/۰۵۲۹	۰/۰۰۸۶	۰/۰۰۳۴	۲	.	.	۲	.
۰/۱۷۴۸	۲/۰۰۲۵	۰/۲۳۱۲	۰/۰۰۰۷	۰/۰۸۴۴	۰/۰۰۰۵	۰/۷۲۰۵	۲/۰۳۵۸	۰/۵۴۶۷	۱/۰۶۲۵	۲	.	.	۲	۱
۰/۱۴۸۹	۱/۹۹۶۸	۰/۲۵۲۵	۰/۰۵۰۵	۰/۲۹۶۵	۱/۰۶۷۲	۰/۲۵۵۰	۰/۰۶۵۶	۰/۳۹۳۸	۰/۰۴۸۰	۲	.	۱	.	.
۰/۱۵۵۵	۱/۹۹۰۸	۰/۲۷۷۵	۰/۰۴۵۳	۰/۳۱۸۹	۱/۰۶۱۰	۰/۳۴۲۷	۰/۰۹۵۰	۰/۶۳۱۹	۰/۷۳۱۳	۲	.	۱	.	۱
۰/۰۷۱۹	۱/۹۸۷۵	۰/۰۷۴۸	۰/۰۰۵۳	۰/۰۷۹۴	۱/۰۰۱۳	۰/۲۰۱۳	۲/۰۵۸۸	۰/۰۰۸۹	۰/۰۰۳۴	۲	.	۱	۲	.
۰/۲۰۶۶	۲/۰۰۹۳	۰/۵۵۶۴	۰/۰۱۶۶	۰/۲۹۳۸	۱/۰۱۴۵	۰/۸۹۰۸	۱/۹۷۱۴	۰/۷۵۹۲	۰/۹۸۸۴	۲	.	۱	۲	۱
۰/۰۹۶۵	۱/۹۷۱۳	۰/۱۵۴۶	۰/۹۹۹۸	۰/۱۲۲۷	۰/۰۰۰۵	۰/۲۹۳۸	۰/۰۶۲۶	۰/۴۱۷۶	۰/۰۵۵۴	۲	۱	.	.	.
۰/۱۰۶۷	۱/۹۷۱۰	۰/۱۹۳۴	۱/۰۰۱۶	۰/۱۴۹۹	۰/۰۰۱۶	۰/۳۶۷۸	۰/۰۹۱۸	۰/۵۷۳۳	۰/۸۳۷۸	۲	۱	.	.	۱
۰/۰۶۹۴	۱/۹۸۵۳	۰/۰۵۷۷	۱/۰۰۰۴	۰/۰۳۴۴	۰/۰۰۰۸	۰/۱۹۰۰	۲/۰۵۶۷	۰/۰۰۸۹	۰/۰۰۳۴	۲	۱	.	۲	.
۰/۱۷۴۶	۲/۰۰۲۹	۰/۲۳۸۰	۰/۹۹۵۷	۰/۰۸۵۵	۰/۰۰۱۹	۰/۷۲۴۲	۲/۰۳۲۵	۰/۵۴۶۷	۱/۰۵۸۰	۲	۱	.	۲	۱
۰/۱۴۷۹	۱/۹۹۶۹	۰/۲۵۵۹	۰/۹۴۶۱	۰/۲۹۷۰	۱/۰۶۹۱	۰/۲۷۷۴	۰/۰۶۹۸	۰/۳۶۷۰	۰/۰۴۹۳	۲	۱	۱	.	.
۰/۱۵۸۰	۱/۹۸۸۴	۰/۲۸۰۰	۰/۹۶۱۲	۰/۳۱۴۶	۱/۰۵۵۳	۰/۳۳۳۸	۰/۰۹۳۱	۰/۶۰۹۳	۰/۷۱۹۶	۲	۱	۱	.	۱
۰/۰۷۱۵	۱/۹۸۹۱	۰/۰۷۴۳	۰/۹۹۴۹	۰/۰۷۸۷	۱/۰۰۰۱	۰/۲۰۱۶	۲/۰۵۴۷	۰/۰۰۸۶	۰/۰۰۳۳	۲	۱	۱	۲	.
۰/۲۰۴۷	۲/۰۰۹۸	۰/۵۵۲۷	۰/۹۹۰۹	۰/۲۹۴۱	۱/۰۱۰۹	۰/۸۷۸۸	۱/۹۵۴۸	۰/۷۰۹۶	۰/۹۷۴۹	۲	۱	۱	۲	۱

جدول ۳: آماره های توصیفی مجموعه داده‌ها

مجموعه داده	تعداد داده‌ها	میانگین	انحراف معیار	ضریب چولگی	ضریب کشیدگی
data Geyser	۲۹۹	۳/۴۶	۱/۱۴	-۰/۴۵	۱/۵۷



شکل ۵: نمودار بافت‌نگار داده‌های مدت زمان فوران و تابع چگالی احتمال توزیع برازش‌شده

جدول ۴: برآورد پارامترهای مدل پیشنهادی و مدل‌های رقیب

					Model
$\hat{\alpha} = ۲,۰۹$	$\hat{\beta} = ۷,۰۸$	$\hat{\lambda} = ۰,۱۴$	$\hat{\sigma} = ۰,۴۳$	$\hat{\mu} = ۳,۰۸$	SKABNN
	$\hat{\alpha} = ۱۳,۶۸$	$\hat{\beta} = ۰,۲۲$	$\hat{\sigma} = ۰,۶۸$	$\hat{\mu} = ۳,۰۹$	ABPN
		$\hat{\lambda} = ۰,۲۱$	$\hat{\sigma} = ۰,۶۹$	$\hat{\mu} = ۳,۱۱$	SBN
	$\hat{\alpha} = ۲۵۸۱$	$\hat{\lambda} = ۰,۲۱$	$\hat{\sigma} = ۰,۶۹$	$\hat{\mu} = ۳,۱۱$	SSCN
		$\hat{\lambda} = -۷,۹۷$	$\hat{\sigma} = ۱,۸۳$	$\hat{\mu} = ۴,۸۹$	SN
$\hat{\rho} = ۰,۳۳$	$\hat{\sigma}_2 = ۰,۴۳$	$\hat{\mu}_2 = ۴,۲۳$	$\hat{\sigma}_1 = ۰,۲۲$	$\hat{\mu}_1 = ۱,۹۵$	Mixed-N

جدول ۵: جدول ۵ مقدار معیارهای نیکویی برازش مدل پیشنهادی در مقایسه با مدل‌های رقیب

BIC	AICC	AIC	Model
۶۱۸/۷۹	۵۹۰/۵۰	۶۰۰/۲۹	SKABNN
۶۴۸/۹۹	۶۲۶/۳۳	۶۳۴/۱۹	ABPN
۶۹۳/۲۵	۶۷۶/۲۳	۶۸۲/۱۵	SBN
۶۹۸/۹۵	۶۷۶/۲۸	۶۸۴/۱۵	SSCN
۸۶۸/۵۷	۸۵۱/۵۵	۸۵۷/۴۷	SN
۶۲۴/۷۹	۵۹۶/۴۹	۶۰۶/۷۹	Mixed-N

پارامترهای توزیع با روش ماکسیمم درستنمایی به دست آمده است. شبیه‌سازی از این توزیع و سپس توانایی روش به دست آوردن برآوردها به صورت عددی نشان داده شده است. در ادامه توزیع پیشنهادی با اضافه شدن دو پارامتر مکانی و مقیاس گسترش یافته است. با به کارگیری این توزیع در برازش داده‌های چشمه‌های آب گرم و مقایسه نتایج با چند توزیع مرسوم، نشان داده شده است که این توزیع می‌تواند برای داده‌های چند مدی مورد استفاده محققین قرار گیرد. بدیهی است گسترش به توزیع‌هایی چند مدی که بتوانند حالت‌های نامتقارن را هم پوشش دهند، می‌تواند تحقیق بعدی این مطالعه باشد.

## فهرست منابع

- [۱] س.م.ر. علوی، ر. چینی پرداز، استنباط در توزیع نرمال بر اساس نمونه‌گیری وزنی، مجله علوم آماری، ۱۳۸۴، ۱ (۱)، ۷۳-۸۸.
- [2] Alavi, S. M. R., On a new bimodal normal family, *Journal of Statistical Research of Iran*, (2011), **8** (2), 163-175.
- [3] Alavi, S. M. R. and Tarhani, M., On a Skew Bimodal Normal-Normal distribution fitted to the Old-Faithful geyser data, *Communications in Statistics - Theory and Methods*, (2016), **46(15)**, 7301-7312.
- [4] Azzalini, A., Bowman, A. W., A Look at Some Data on the Old Faithful Geyser, *Journal of Applied Statistics*, (1990). **39**, 357-365.
- [5] Azzalini, A., A Class of Distributions which Includes the Normal Ones, *Scandinavian Journal of Statistics*, (1985). **12**, 171-178.
- [6] Azzalini, A., The Skew-normal Distribution and Relative Multivariate Families, *Scandinavian Journal of Statistics*, (2005). **32**, 159-188.
- [7] Bolfarine H., Gómez, H.W., Rivas, L. The log-bimodal-skew-normal model. A geochemical application, *Journal of Chemometrics*, (2011). **25(6)**, 329-332.
- [8] Gómez, H. W., Venegas, O. and Bolfarine, H., Skew-Symmetric Distributions Generated by the Distribution Function of the Normal Distribution, *Environmetrics*, (2007). **18**, 395-407
- [9] Gradshteyn, I.S. and Ryzhik, I.M., *Tables of Integrals, Seires and Products*, Academic press, (1965).
- [10] Liseo, B., The Skew-Normal Class of Densities: Aspects of Inference From the Bayesian Point of View, *Statistica*, (1990), **50(1)**, 71-82.
- [11] Martinez, E. H., Varela, H., Gomez, H. W. and Bolfarine, H., A Note on the Likelihood and Moments of the Skew-Normal Distribution, *Statistical Operation Research Transaction*, (2008), **32(1)**, 57-66.
- [12] Maleki, M. and Nematollahi, A. R., Bayesian Approach to Epsilon-Skew-Normal Family, *Communications in Statistics-Theory and Methods*, (2016), **46 (15)**, 7546-7561.
- [13] Mamei, V. and Musio, M., A Generalization of the Skew-Normal Distribution: the Beta-Skew-Normal Distribution, *Communications in Statistics: Theory and Methods*, (2013), **42**, 2229-2242.
- [14] Nekoukhou, V. and Alamatsaz, M. H., A family of skew-symmetric-Laplace distributions, *Stistical papers*, (2012), **53**, 685-696.
- [15] Pewsey, A., Problems of Inference for Azzalini's Skew-Normal Distribution, *Journal of Applied Statistics*, (2000), **27(7)**, 859-870.



- [16] Rasekhi, M., Chinipardaz, R. and Alavi, S. M. R., A Flexible Generalization of the Skew Normal Distribution Based on a Weighted Normal Distribution, *Statistical Methods and Application*, (2015), **25 (3)**, 375-394.
- [17] Sharafi, M. and Behboodian, J., The Balakrishnan Skew-Normal Density, *Statistical Papers*, (2007), **49**, 769-778.
- [18] Wang, J., Boyer, J. and Genton, M. G., A skew-symmetric representation of multivariate distributions, *Statist. Sinica*, (2004), **14**, 1259 – 1270.
- [19] Yadegari, I., Gerami, A. and Khaledi, M. J., A Generalized of the Balakrishnan Skew-normal Distribution, *Statistics and Probability Letters*, (2008), **78**, 1165- 1167.



## Skew multimodal normal distribution

F. Naghibi, S.M.R. Alavi <sup>†</sup>, R. Chinipardaz

Department of Statistics, Faculty of Mathematics and Computer Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

Received: 2020/11/16

Accepted: 2021/5/15

Communicated by: Gholam Ali Parham

**Abstract:** In this paper, a multimodal normal distribution is introduced. This distribution is a generalization of the standard normal distribution and covers the symmetric distributions with up to three modes. Then the skew multimodal normal distribution is introduced as a skewed-symmetric distribution generated by the standard normal distribution. Some properties of this distribution are investigated. Data generation methods from this distribution are proposed and, its location and scale distribution is introduced. The maximum likelihood estimates of the parameters are studied and evaluated using simulations. The famous geysers dataset is fitted to this distribution and compared with several competing distributions.

**Keywords:** skew normal, multimodal, geysers data, estimation.

**Mathematics Subject Classification (2010):** 62E05, 62E10.



©2021 Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (<http://creativecommons.org/licenses/by-nc/4.0/>).

<sup>†</sup> Corresponding author: [alavi\\_m@scu.ac.ir](mailto:alavi_m@scu.ac.ir)