

توزیع نیم-نرمال-تعمیم‌یافته: مدل و ویژگی‌ها

عاطفه پورچیت‌ساز، عیسی محمودی*؛
دانشگاه یزد، دانشکده ریاضی، گروه آمار

دریافت ۹۵/۳/۲۵ پذیرش ۹۵/۰۹/۱۵

چکیده

در این مقاله، تعمیم جدیدی از توزیع نیم-نرمال با نام توزیع نیم-نرمال-تعمیم‌یافته معرفی می‌شود. این توزیع در حالت خاص شامل توزیع نیم-نرمال است. ویژگی‌های این توزیع بررسی می‌شود. همچنین تابع چگالی احتمال، تابع توزیع، گشتاور مرتبه‌ی r ام، ضرایب چولگی و کشیدگی و تابع مولد گشتاور بررسی می‌شود. در باره برخی از جنبه‌های استنباطی مربوط به برآورد ماکسیمم درست‌نمایی بحث می‌شود. در پایان، انعطاف‌پذیری این توزیع با کاربرد مجموعه داده‌های واقعی نشان داده شده است.

واژه‌های کلیدی: برآوردگر گشتاوری، برآوردگر ماکسیمم درست‌نمایی، تابع نرخ خطر، توزیع نیم-نرمال.

مقدمه

یک توزیع مناسب برای برآزش به داده‌های مثبت، توزیع نیم-نرمال است. اگر Z یک متغیر تصادفی نرمال استاندارد باشد، $Z \sim N(0, 1)$ آن‌گاه $X = |Z|$ از توزیع نیم-نرمال استاندارد (نرمال بریده شده به چپ^۱ در نقطه صفر) و $X = -|Z|$ از توزیع نیم-نرمال منفی (نرمال بریده شده به راست^۲ در نقطه صفر) پیروی می‌کند. توزیع نیم-نرمال حالت خاصی از توزیع‌های نرمال تاخورد^۳ و نرمال بریده شده به چپ در نقطه صفر است.

به‌عنوان یک توسیع، $Y = \mu + \sigma X$ یک متغیر تصادفی نیم-نرمال کلی با تابع چگالی به‌صورت:

$$f(y; \mu, \sigma) = \frac{2}{\sigma} \varphi\left(\frac{y-\mu}{\sigma}\right), y > \mu, -\infty < \mu < \infty, \sigma > 0,$$

است که در آن φ تابع چگالی نرمال استاندارد است. از نماد $Y \sim HN(\mu, \sigma)$ برای نمایش این توزیع استفاده می‌شود. متغیر تصادفی $Y^* = \mu - \sigma X$ از توزیع نیم-نرمال کلی منفی پیروی می‌کند.

توزیع نیم-نرمال کلی، یک توزیع حدی مهم از کلاس توزیع‌های چوله-نرمال سه پارامتری آزالینی معرفی شده در [۲] است. اغلب برای داده‌های با چولگی زیاد، برآوردهای ماکسیمم درست‌نمایی برآزش شده از یک توزیع چوله-نرمال، مطابق یک توزیع نیم-نرمال کلی است [۲]، [۳]، [۹].

توزیع نیم-نرمال کلی با $\mu = 0$ که اولین بار دانیل [۶] در سال ۱۹۵۹ معرفی کرد، در بسیاری از زمینه‌ها استفاده شده است. این توزیع با نماد $HN(\sigma)$ نشان داده می‌شود.

در ادامه به تعمیم‌های دیگر توزیع نیم-نرمال، از جمله توزیع نرمال تاخورد و توزیع نرمال بریده شده اشاره می‌شود.

توزیع نرمال تاخوردده یک توزیع پیوسته در نظریه احتمال و آمار است. فرض کنید $X \sim N(\mu, \sigma^2)$ آن گاه متغیر تصادفی $Y = |X|$ دارای توزیع نرمال تاخوردده است که در آن زمانی که $\mu = 0$ ، توزیع نیم-نرمال با پارامتر σ به دست می‌آید.

در آمار و احتمال توزیع نرمال بریده، متغیر تصادفی است که به طور نرمال توزیع شده است که از پایین یا بالا (یا هر دو) محدود شده است. توزیع نرمال بریده دارای کاربردهای گسترده‌ای در آمار و اقتصادسنجی است. به عنوان مثال، برای مدل سازی احتمال برآمدهای دوحالتی در مدل پروبیت و برای مدل داده‌های سانسور شده در مدل توبیت استفاده می‌شود.

فرض کنید $X \sim N(\mu, \sigma^2)$ ، $X \in (a, b)$ و $-\infty \leq a < b \leq \infty$. آن گاه متغیر تصادفی X با شرط $a < X < b$ دارای توزیع نرمال بریده است. زمانی که $\mu = 0$ و $a = 0$ و $b = \infty$ ، توزیع نرمال بریده از چپ در نقطه صفر به دست می‌آید که همان توزیع نیم-نرمال با پارامتر σ است [۷].

تعمیمی از توزیع نیم-نرمال را کورای و آناندا [۵] در ۲۰۰۸ بررسی کردند. در این مقاله ابتدا به معرفی توزیع جدید می‌پردازیم و توابع چگالی، توزیع، قابلیت اعتماد و نرخ خطر و هم‌چنین گشتاورها و تابع چگالی آماره‌های ترتیبی برای آن به دست می‌آید. سپس در بخش استنباط، به بررسی برآوردگرهای گشتاوری و برآوردگرهای ماکسیمم درست‌نمایی می‌پردازیم. سپس شبیه‌سازی و بعد از آن کاربرد دو مجموعه داده واقعی بررسی می‌شود و در پایان به بحث و نتیجه‌گیری می‌پردازیم.

معرفی توزیع نیم-نرمال تعمیم یافته

در این بخش، توابع چگالی، توزیع، قابلیت اعتماد، نرخ خطر، گشتاورها و برخی دیگر از ویژگی‌های توزیع نیم-نرمال تعمیم یافته بررسی شده است. هم‌چنین تعدادی از نمودارهای توابع چگالی، توزیع، قابلیت اعتماد، نرخ خطر، ضرایب چولگی و کشیدگی مربوط به توزیع نیم-نرمال تعمیم یافته برای برخی مقادیر پارامترها رسم شده است.

تابع چگالی

با توجه به این که توزیع مذکور در این مقاله می‌تواند در حالت خاص توزیع‌های نیم-نرمال و نمایی را در برگیرد، بنا براین برای داده‌هایی که به یکی از توزیع‌های نیم-نرمال یا نمایی خوب برازش می‌شوند این مدل رقیب مناسبی برای این دو توزیع است. هم‌چنین برای آن دسته از داده‌هایی که توزیع نیم-نرمال و نمایی به خوبی برازش نمی‌دهند، مدل جدید با یک پارامتر اضافه می‌تواند بهتر از زیرمدل‌هایش به داده‌ها برازش شود.

توزیع نرمال تعمیم یافته با پارامترهای μ ، σ و S را اولین بار ناداراجا [۸] در سال ۲۰۰۵ معرفی کرد. این توزیع زنگی شکل و تک‌مدی است که مد آن در نقطه $X = \mu$ رخ می‌دهد. توزیع نرمال تعمیم یافته در حالت خاص، توزیع‌های نرمال و لاپلاس را شامل می‌شود. در این مقاله هدف معرفی تعمیم جدیدی از توزیع نیم-نرمال با استفاده از توزیع نرمال تعمیم یافته است.

فرض کنید متغیر تصادفی X دارای توزیع نرمال تعمیم یافته با پارامترهای $\mu = 0$ ، σ و S باشد اگر $Y = |X|$ ، در این صورت متغیر تصادفی Y دارای توزیع نیم-نرمال تعمیم یافته با پارامترهای σ و S است.

متغیر تصادفی Y دارای توزیع نیم-نرمال تعمیم‌یافته با پارامترهای σ و s است هرگاه دارای این تابع چگالی احتمال باشد:

$$f_Y(y; \sigma, s) = \frac{1}{\sigma \Gamma(\frac{1}{s} + 1)} e^{-\frac{y^s}{\sigma^s}}, \quad y > 0, \sigma > 0, s > 0.$$

از نماد $Y \sim HGN^r(\sigma, s)$ برای نمایش این توزیع استفاده می‌شود.

نکته ۱. اگر $Y \sim HGN(\sigma', s)$ و $\sigma'^2 = 2\sigma^2$ و $s = 2$ آن‌گاه توزیع نیم-نرمال با پارامتر σ به دست می‌آید.

هم‌چنین اگر $\sigma' = \sigma$ و $s = 1$ ، آن‌گاه توزیع نمایی با پارامتر σ به دست می‌آید.

قضیه ۱. تابع چگالی توزیع نیم-نرمال تعمیم‌یافته به‌ازای تمام مقادیر $\sigma, s > 0$ همواره نزولی است.

اثبات. از سویی

$$\lim_{y \rightarrow 0} f_Y(y; \sigma, s) = \frac{1}{\sigma \Gamma(\frac{1}{s} + 1)}, \quad \lim_{y \rightarrow \infty} f_Y(y; \sigma, s) = 0,$$

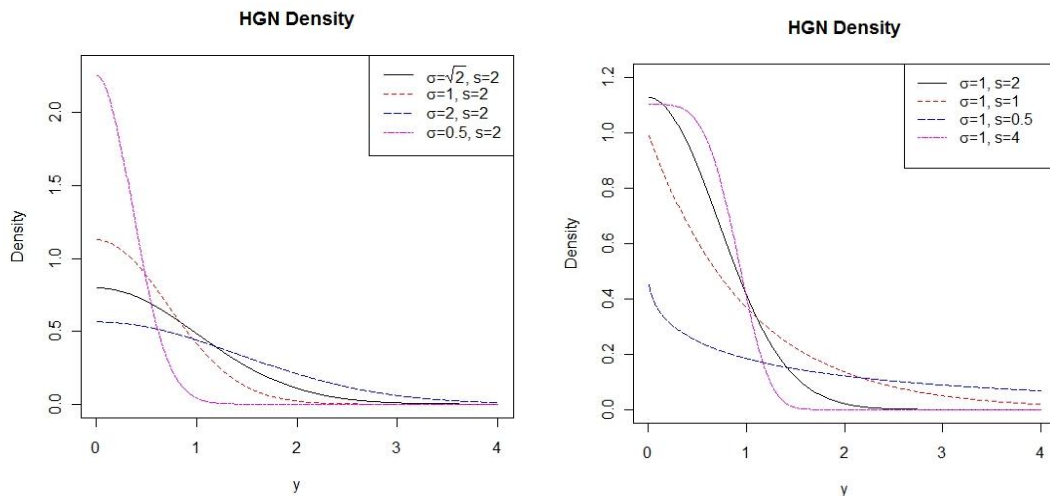
از سویی دیگر، مشتق مرتبه اول تابع چگالی توزیع نیم-نرمال تعمیم‌یافته بدین صورت است:

$$\frac{d}{dy} f_Y(y; \sigma, s) = \frac{1}{\sigma \Gamma(\frac{1}{s} + 1)} e^{-\frac{y^s}{\sigma^s}} \left(-\frac{sy^{s-1}}{\sigma^s} \right),$$

که همواره منفی است و بنا براین تابع چگالی توزیع نیم-نرمال تعمیم‌یافته برای تمام مقادیر $\sigma, s > 0$ همواره نزولی است.

در شکل ۱ نمودار تابع چگالی احتمال توزیع نیم-نرمال تعمیم‌یافته برای مقادیر متفاوت پارامترهای σ و s رسم

شده است.



شکل ۱. تابع چگالی احتمال توزیع نیم-نرمال تعمیم‌یافته برای مقادیر مختلف پارامترهای σ و s

قضیه ۲. فرض کنید $X \sim \text{Gamma}(\frac{1}{s}, \frac{1}{\sigma^s})$ با تابع چگالی $f_X(x; s, \sigma) = \frac{x^{\frac{1}{s}-1} e^{-\frac{x}{\sigma^s}}}{\sigma \Gamma(\frac{1}{s})}$ ، آن‌گاه $Y = X^{\frac{1}{s}}$ دارای

توزیع نیم-نرمال تعمیم‌یافته است یعنی $Y \sim HGN(\sigma, s)$

اثبات. با استفاده از تبدیل ژاکوبی تابع چگالی متغیر تصادفی Y بدین صورت به دست می‌آید:

$$f_Y(y) = f_X(y^s) \left| \frac{d}{dy} y^s \right| = \frac{s}{\sigma \Gamma(\frac{1}{s})} \exp \left\{ -\frac{y^s}{\sigma^s} \right\},$$

و بنا براین قضیه اثبات می‌شود.

تابع توزیع

قضیه ۳. اگر متغیر تصادفی $Y \sim HGN(\sigma, s)$ باشد، آن‌گاه تابع توزیع متغیر تصادفی Y بدین صورت است:

$$F_Y(y; \sigma, s) = G(y^s; \frac{1}{s}, \frac{1}{\sigma^s}),$$

که در آن $G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})$ تابع توزیع تجمعی گاما با پارامتر شکل $\frac{1}{s}$ و پارامتر مقیاس $\frac{1}{\sigma^s}$ در نقطه y^s است.

اثبات. با توجه به قضیه ۲ می‌توان تابع توزیع نیم-نرمال تعمیم‌یافته را بدین صورت به دست آورد:

$$F_Y(y; \sigma, s) = P(Y \leq y) = P(X^{\frac{1}{s}} \leq y) = P(X \leq y^s) = F_X(y^s) = G(y^s; \frac{1}{s}, \frac{1}{\sigma^s}),$$

و بنا براین قضیه اثبات شد.

توابع قابلیت اعتماد و نرخ خطر

تابع قابلیت اعتماد متغیر تصادفی Y ، زمانی که $Y \sim HGN(\sigma, s)$ باشد از این رابطه پیروی می‌کند:

$$R(y; \sigma, s) = 1 - G(y^s; \frac{1}{s}, \frac{1}{\sigma^s}).$$

به این ترتیب، تابع نرخ خطر برای توزیع نیم-نرمال تعمیم‌یافته بدین صورت به دست می‌آید:

$$h(y; \sigma, s) = \frac{\frac{1}{\sigma^{\frac{1}{s}+1}} e^{-\frac{y^s}{\sigma^s}}}{1 - G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})}. \quad (1)$$

گزاره ۱. برای تابع نرخ خطر به دست آمده در رابطه (۱) داریم:

$$\lim_{y \rightarrow 0} h(y; \sigma, s) = \frac{s}{\sigma \Gamma(\frac{1}{s})},$$

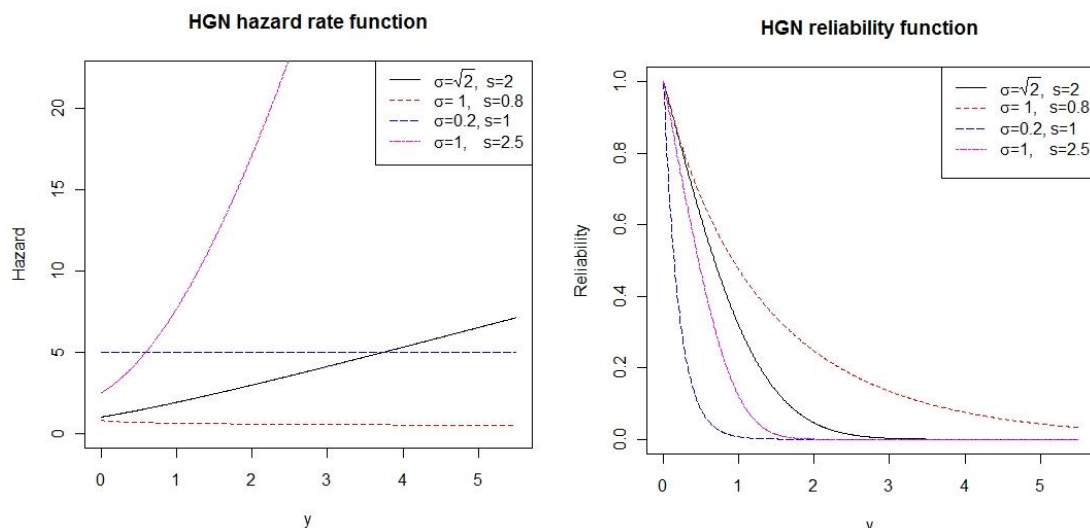
$$\lim_{y \rightarrow \infty} h(y; \sigma, s) = \begin{cases} 0, & 0 < s < 1 \\ \frac{1}{\sigma}, & s = 1 \\ \infty, & s > 1 \end{cases}.$$

در شکل ۲ نمودار توابع قابلیت اعتماد و نرخ خطر برای توزیع نیم-نرمال تعمیم‌یافته، به‌ازای مقادیر متفاوت

پارامترهای σ و s رسم شده است.

گشتاورها

در این قسمت گشتاور مرتبه r ام و تابع مولد گشتاور توزیع نیم-نرمال تعمیم‌یافته به دست آمده است.



شکل ۲. نمودار توابع قابلیت اعتماد و نرخ خطر برای توزیع نیم-نرمال تعمیم یافته، برای مقادیر مختلف پارامترهای σ و s

قضیه ۴. فرض کنید $Y \sim HGN(\sigma, s)$ ، آن گاه برای $r = 1, 2, \dots$ گشتاور مرتبه r ام توزیع عبارت است از:

$$\mu_r = E[Y^r] = \frac{\sigma^r \Gamma\left(\frac{r+1}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}. \quad (2)$$

اثبات. گشتاور مرتبه r ام توزیع نیم-نرمال تعمیم یافته با استفاده از قضیه ۲ از رابطه (۳) به دست می آید:

$$E[Y^r] = E\left[X^{\frac{r}{s}}\right] = \int_0^\infty \frac{x^{\frac{r+1}{s}-1} e^{-\frac{x}{\sigma^s}}}{\sigma \Gamma\left(\frac{1}{s}\right)} dx = \frac{\sigma^r \Gamma\left(\frac{r+1}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}, \quad (3)$$

و بدین ترتیب اثبات قضیه کامل می شود.

قضیه ۵. فرض کنید $Y \sim HGN(\sigma, s)$ باشد. آن گاه تابع مولد گشتاور متغیر تصادفی Y بدین صورت است:

$$\begin{aligned} M_Y(t) &= \frac{1}{\sigma \Gamma\left(\frac{1}{s} + 1\right)} \int_0^\infty \exp\left\{-\left(\frac{y^s}{\sigma^s} - ty\right)\right\} dy \\ &= \frac{1}{\Gamma\left(\frac{1}{s}\right)} \sum_{n=0}^\infty \frac{(t\sigma)^n}{n!} \Gamma\left(\frac{n+1}{s}\right). \end{aligned}$$

نتیجه ۱. فرض کنید $Y \sim HGN(\sigma, s)$ باشد، آن گاه امید ریاضی و واریانس متغیر تصادفی Y از این رابطه به دست می آید:

$$\begin{aligned} E[Y] &= \frac{\sigma \Gamma\left(\frac{2}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}, \\ \text{Var}(Y) &= \frac{\sigma^2 \left(\Gamma\left(\frac{1}{s}\right) \Gamma\left(\frac{3}{s}\right) - \Gamma^2\left(\frac{2}{s}\right) \right)}{\Gamma^2\left(\frac{2}{s}\right)}. \end{aligned}$$

فرض کنید $\mu = E[Y]$ باشد، آن گاه برای $r = 1, 2, \dots$ گشتاور مرکزی مرتبه r ام (گشتاور حول میانگین) توزیع نیم-نرمال تعمیم یافته بدین صورت به دست می آید:

$$E[(Y - \mu)^r] = \sum_{k=0}^r \binom{r}{k} E[Y^k] (-\mu)^{r-k}. \quad (4)$$

با جای‌گذاری (۲) در رابطه (۴) داریم:

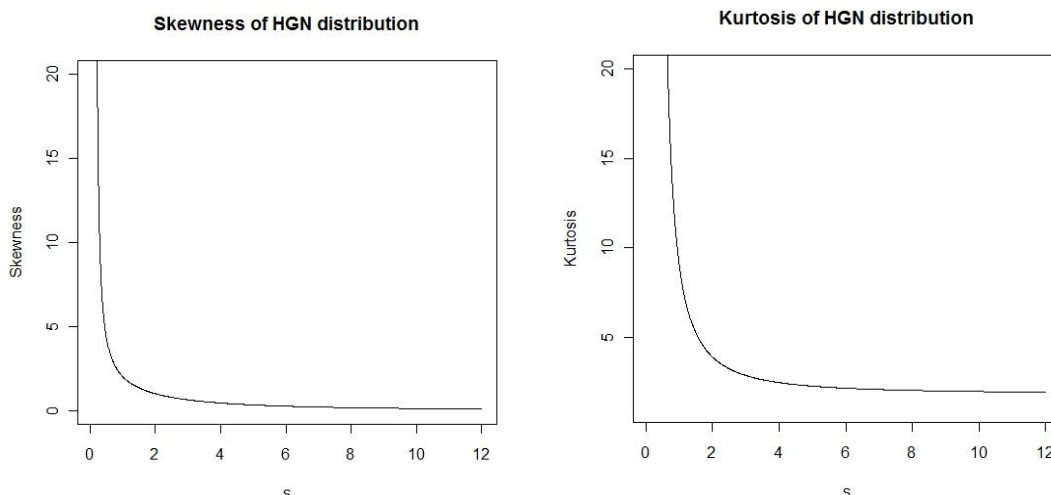
$$E[(Y - \mu)^r] = \frac{\sigma^r \Gamma^r(\frac{2}{s})}{\Gamma^r(\frac{1}{s})} \sum_{k=0}^r \binom{r}{k} (-1)^{r-k} \frac{\Gamma^{k-1}(1/s) \Gamma((k+1)/s)}{\Gamma^k(\frac{2}{s})}.$$

نتیجه ۲. فرض کنید $Y \sim HGN(\sigma, s)$ باشد، آن‌گاه ضرایب چولگی و کشیدگی آن به ترتیب بدین صورت به دست می‌آید:

$$\sqrt{\beta_1} = \frac{\Gamma^2(1/s)\Gamma(4/s) - 3\Gamma(1/s)\Gamma(2/s)\Gamma(3/s) + 2\Gamma^3(2/s)}{\left(\Gamma\left(\frac{1}{s}\right)\Gamma\left(\frac{3}{s}\right) - \Gamma^2\left(\frac{2}{s}\right)\right)^{3/2}},$$

$$\beta_2 = \frac{\Gamma^3\left(\frac{1}{s}\right)\Gamma\left(\frac{5}{s}\right) - 4\Gamma^2\left(\frac{1}{s}\right)\Gamma\left(\frac{2}{s}\right)\Gamma\left(\frac{4}{s}\right) + 2\Gamma^3\left(\frac{2}{s}\right)}{\left(\Gamma\left(\frac{1}{s}\right)\Gamma\left(\frac{3}{s}\right) - \Gamma^2\left(\frac{2}{s}\right)\right)^2} + \frac{6\Gamma(1/s)\Gamma^2(2/s)\Gamma(3/s) - 3\Gamma^4(2/s)}{\left(\Gamma\left(\frac{1}{s}\right)\Gamma\left(\frac{3}{s}\right) - \Gamma^2\left(\frac{2}{s}\right)\right)^2}.$$

در شکل ۳ نمودار ضرایب چولگی و کشیدگی توزیع نیم-نرمال تعمیم‌یافته برای برخی از مقادیر پارامتر s رسم شده است.



شکل ۳. نمودارهای مربوط به ضرایب چولگی و کشیدگی توزیع نیم-نرمال تعمیم‌یافته برای برخی مقادیر s

تابع چگالی آماره‌های ترتیبی توزیع نیم-نرمال تعمیم‌یافته

فرض کنید Y_1, \dots, Y_n نمونه‌ای تصادفی از توزیع نیم-نرمال تعمیم‌یافته باشد، آن‌گاه تابع چگالی r امین آماره

ترتیبی آن بدین صورت به دست می‌آید:

$$f_{Y_{r:n}}(y; \sigma, s) = \frac{n!}{(r-1)!(n-r)! \sigma \Gamma\left(\frac{1}{s} + 1\right)} \exp\left\{-\frac{y^s}{\sigma^s}\right\} \times \left[G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{r-1} \left[1 - G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{n-r}. \quad (5)$$

با استفاده از رابطه (۵)، می‌توان توابع چگالی احتمال کوچک‌ترین و بزرگ‌ترین آماره ترتیبی را به دست آورد. در حالت خاص با قرار دادن $r = 1$ و $r = n$ داریم:

$$f_{Y_{1:n}}(y; \sigma, s) = \frac{n}{\sigma \Gamma\left(\frac{1}{s} + 1\right)} \exp\left\{-\frac{y^s}{\sigma^s}\right\} \left[1 - G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{n-1},$$

9

$$f_{Y_{n:n}}(y; \sigma, s) = \frac{n}{\sigma \Gamma\left(\frac{1}{s} + 1\right)} \exp\left\{-\frac{y^s}{\sigma^s}\right\} \left[G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{n-1}.$$

تابع توزیع r امین آماره ترتیبی مربوط به توزیع نیم-نرمال تعمیم یافته بدین صورت به دست می‌آید:

$$F_{Y_{r:n}}(y; \sigma, s) = \sum_{k=r}^n \binom{n}{k} \left[G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^k \left[1 - G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{n-k}.$$

تابع چگالی احتمال توأم i امین و j امین آماره ترتیبی، $(Y_{i:n}, Y_{j:n})$ از توزیع نیم-نرمال تعمیم یافته که $1 \leq i < j \leq n$ از این رابطه پیروی می‌کند:

$$f_{Y_{i:n}, Y_{j:n}}(x, y) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \left[G(x^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{i-1} \\ \times \left[G(y^s; \frac{1}{s}, \frac{1}{\sigma^s}) - G(x^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{j-i-1} \left[1 - G(y^s; \frac{1}{s}, \frac{1}{\sigma^s})\right]^{n-j} \\ \times \frac{1}{\sigma^2 \Gamma^2\left(\frac{1}{s} + 1\right)} \exp\left\{-\frac{x^s + y^s}{\sigma^s}\right\}, \quad x < y.$$

استنباط

در این بخش، در باره برآوردهای گشتاوری و ماکسیمم درست‌نمایی برای پارامترهای σ و s از مدل نیم-نرمال تعمیم یافته بحث می‌شوند.

برآوردهای گشتاوری

فرض کنید Y_1, \dots, Y_n نمونه تصادفی از متغیر تصادفی $Y \sim HGN(\sigma, s)$ است، آن‌گاه با استفاده از رابطه (۲) داریم:

$$E[Y] = \frac{\sigma \Gamma(2/s)}{\Gamma(1/s)}, \quad (6)$$

$$E[Y^2] = \frac{\sigma^2 \Gamma(3/s)}{\Gamma(1/s)} \quad (7)$$

حال فرض کنید، \bar{y} میانگین نمونه‌ای (اولین گشتاور نمونه) و \bar{y}^2 دومین گشتاور نمونه بر اساس نمونه مشاهده شده y_1, \dots, y_n باشد. با مساوی قرار دادن روابط (۶) و (۷) به ترتیب با \bar{y} و \bar{y}^2 می‌توان معادله (۸) دست پیدا کرد:

$$\frac{\Gamma(1/s)\Gamma(3/s)}{\Gamma^2(2/s)} = \frac{\bar{y}^2}{\bar{y}^2}, \quad (8)$$

که با حل رابطه (۸) نسبت به s ، برآورد گشتاوری پارامتر s به دست می‌آید، سپس با جای گذاری این برآورد در رابطه

$$\sigma = \frac{\Gamma(1/s)}{\Gamma(2/s)} \bar{y},$$

برآورد گشتاوری پارامتر σ نیز به دست می‌آید.

برآوردگرهای ماکسیمم درست‌نمایی

در این بخش برآوردگرهای ماکسیمم درست‌نمایی برای توزیع نیم-نرمال تعمیم‌یافته بحث می‌شود. فرض کنید Y_1, \dots, Y_n نمونه‌ای تصادفی به حجم n از توزیع نیم-نرمال تعمیم‌یافته با پارامترهای σ و s و مقادیر مشاهده شده y_1, \dots, y_n باشند، در این صورت تابع درست‌نمایی بدین صورت است:

$$L(\sigma, s) = \frac{1}{\sigma^n \Gamma^n\left(\frac{1}{s} + 1\right)} \exp\left\{-\frac{\sum_{i=1}^n y_i^s}{\sigma^s}\right\},$$

و تابع لگاریتم درست‌نمایی آن از این رابطه پیروی می‌کند:

$$l(\sigma, s) = -n \log \sigma - n \log \Gamma\left(\frac{1}{s} + 1\right) - \frac{\sum_{i=1}^n y_i^s}{\sigma^s}.$$

با مشتق‌گیری از تابع لگاریتم درست‌نمایی نسبت به پارامترها داریم:

$$\frac{\partial}{\partial \sigma} l(\sigma, s) = -\frac{n}{\sigma} + \frac{s \sum_{i=1}^n y_i^s}{\sigma^{s+1}},$$

$$\frac{\partial}{\partial s} l(\sigma, s) = \frac{n}{s^2} \psi\left(\frac{1}{s} + 1\right) - \frac{\sum_{i=1}^n (y_i^s \log y_i) - \log \sigma \sum_{i=1}^n y_i^s}{\sigma^s},$$

که در آن ψ تابع دو گاما $(\psi(\alpha) = \frac{d}{d\alpha} \log \Gamma(\alpha))$ است.

بنا بر این معادلات ماکسیمم درست‌نمایی بدین صورت به دست می‌آید:

$$-\frac{n}{\sigma} + \frac{s \sum_{i=1}^n y_i^s}{\sigma^{s+1}} = 0, \quad (9)$$

$$\frac{n}{s^2} \psi\left(\frac{1}{s} + 1\right) - \frac{\sum_{i=1}^n (y_i^s \log y_i) - \log \sigma \sum_{i=1}^n y_i^s}{\sigma^s} = 0, \quad (10)$$

حال می‌توان با حل معادله (۹) نسبت به σ به رابطه (۱۱) دست یافت:

$$\sigma = \left(\frac{s}{n} \sum_{i=1}^n y_i^s\right)^{1/s}, \quad (11)$$

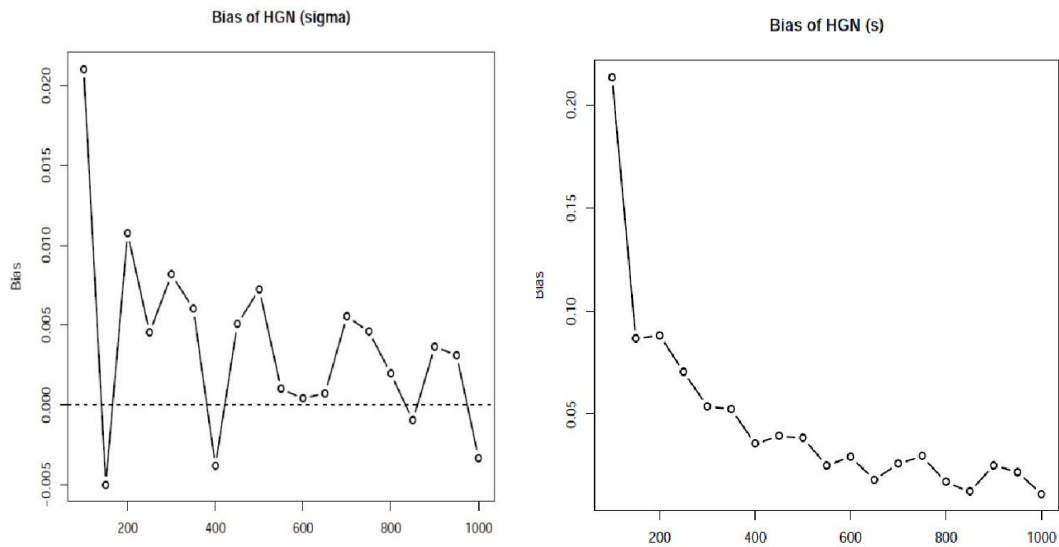
و با جای‌گذاری (۱۱) در رابطه (۱۰) و حل آن نسبت به s ، می‌توان برآورد ماکسیمم درست‌نمایی s و سپس σ را به دست آورد.

شبیه‌سازی

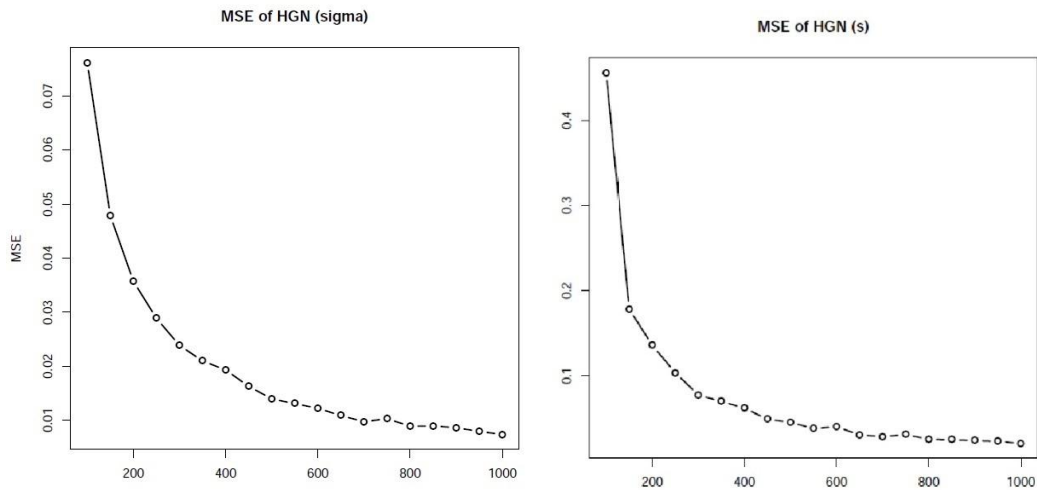
در این بخش، شبیه‌سازی به روش نیوتن رافسون انجام می‌شود. نتایج حاصل از شبیه‌سازی به منظور بررسی سازگاری برآوردهای ماکسیمم درست‌نمایی در این بخش داده شده است. نتایج شبیه‌سازی برآوردها، انحراف استاندارد، اریبی و جذر میانگین توان دوم خطای پارامترهای توزیع نیم-نرمال تعمیم‌یافته با استفاده از روش ماکسیمم درست‌نمایی به‌ازای مقادیر مختلف n و با تکرار 1000 ($rep = 1000$) در جدول ۱ آمده است. چنان‌که از مقادیر جدول مشاهده می‌شود با افزایش n برآوردها به مقدار واقعی پارامترها نزدیک‌تر می‌شوند که باعث کاهش اریبی و جذر میانگین توان دوم خطای پارامترها می‌شود.

در ادامه برای مثال، نمودارهای مربوط به شبیه‌سازی توزیع نیم-نرمال تعمیم‌یافته با پارامترهای $\sigma = 2$ و $s = 2$ آورده شده است.

در شکل‌های ۴ و ۵ به ترتیب نمودارهای مربوط به اریبی و میانگین توان دوم خطای پارامترهای شبیه‌سازی شده توزیع $HGN(2, 2)$ رسم شده است. چنان‌که در شکل دیده می‌شود، با افزایش n مقادیر اریبی و میانگین توان دوم پارامترهای σ و s به صفر نزدیک می‌شود.



شکل ۴. نمودارهای مربوط به اریبی پارامترهای شبیه‌سازی شده توزیع $HGN(2, 2)$



شکل ۵. نمودارهای مربوط به میانگین توان دوم خطای پارامترهای شبیه‌سازی شده توزیع $HGN(2, 2)$

کاربردها در مورد مجموعه داده‌های واقعی

در این بخش، به منظور نشان دادن میزان انعطاف‌پذیری توزیع نیم-نرمال تعمیم‌یافته نسبت به توزیع نیم-نرمال از دو سری داده واقعی برای برآزش استفاده می‌شود. در نهایت مقایسه‌ای بین این توزیع با توزیع نیم-نرمال، با توجه به مقادیر AIC ، $AICc$ ، BIC ، آماره آزمون‌های اندرسون-دارلینگ، کرام-فون میزس، کولموگروف اسمیرنوف و p -مقدار متناظر صورت می‌گیرد.

کاربرد این دو سری داده واقعی برای برازش به دو مدل نیم-نرمال و نیم-نرمال تعمیم‌یافته، نشان از بهتر برازش شدن داده‌های این پژوهش به مدل نیم-نرمال تعمیم‌یافته نسبت به نیم-نرمال دارد.

جدول ۱. میانگین برآورد‌ها، انحراف استاندارد، اربیبی، جذر میانگین توان دوم خطا و کوواریانس برآوردگرهای $\hat{\sigma}$ و $\hat{\delta}$ در توزیع نیم-نرمال تعمیم‌یافته با استفاده از روش ماکسیمم درست‌نمایی

n	(σ, δ)	AE		Sd		Bias		\sqrt{MSE}		cov($\hat{\sigma}, \hat{\delta}$)
		$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	$\hat{\sigma}$	$\hat{\delta}$	
۱۰۰	(۰/۵, ۰/۵)	۰/۶۱۷	۰/۵۲۰	۰/۳۸۵	۰/۰۸۹	۰/۱۱۷	۰/۰۲۰	۰/۴۱۳	۰/۰۹۵	۰/۰۳۵
	(۰/۵, ۱)	۰/۵۲۸	۱/۰۶۸	۰/۱۴۱	۰/۲۱۸	۰/۰۲۸	۰/۰۶۸	۰/۱۵۳	۰/۲۴۶	۰/۰۳۲
	(۰/۵, ۲)	۰/۵۰۵	۲/۱۸۸	۰/۰۶۸	۰/۵۸۹	۰/۰۰۵	۰/۱۸۸	۰/۰۶۶	۰/۶۴۲	۰/۰۳۰
	(۱, ۰/۵)	۰/۲۱۴	۰/۵۲۰	۰/۷۵۸	۰/۰۸۹	۰/۲۱۴	۰/۰۲۰	۰/۷۶۷	۰/۰۹۱	۰/۰۶۲
	(۱, ۱)	۰/۰۳۲	۰/۰۵۶	۰/۲۷۹	۰/۲۱۳	۰/۰۳۲	۰/۰۵۶	۰/۲۸۴	۰/۲۳۰	۰/۰۵۶
	(۱, ۲)	۱/۰۰۷	۲/۱۷۸	۰/۱۳۶	۰/۵۷۵	۰/۰۰۷	۰/۱۷۸	۰/۱۳۵	۰/۶۰۳	۰/۰۶۱
	(۲, ۰/۵)	۲/۴۶۸	۰/۵۲۰	۱/۵۳۵	۰/۰۸۹	۰/۴۶۸	۰/۰۲۰	۱/۶۳۱	۰/۰۹۳	۰/۱۳۵
	(۲, ۱)	۲/۰۶۲	۱/۰۵۵	۰/۵۵۸	۰/۲۱۳	۰/۰۶۲	۰/۰۵۵	۰/۵۶۹	۰/۲۲۹	۰/۱۱۳
(۲, ۲)	۲/۰۲۱	۲/۲۱۳	۰/۲۶۸	۰/۶۴۷	۰/۰۲۱	۰/۲۱۳	۰/۲۶۷	۰/۶۷۵	۱/۳۰	
۳۰۰	(۰/۵, ۰/۵)	۰/۵۲۸	۰/۵۰۴	۰/۱۹۷	۰/۰۴۹	۰/۰۲۸	۰/۰۰۴	۰/۲۰۰	۰/۰۴۹	۰/۰۰۹
	(۰/۵, ۱)	۰/۵۰۸	۱/۰۱۷	۰/۰۸۲	۰/۱۱۲	۰/۰۰۸	۰/۰۱۷	۰/۰۸۱	۰/۱۱۱	۰/۰۰۸
	(۰/۵, ۲)	۰/۵۰۳	۲/۰۶۳	۰/۰۳۹	۰/۲۸۰	۰/۰۰۳	۰/۰۶۳	۰/۰۴۰	۰/۲۹۲	۰/۰۰۹
	(۱, ۰/۵)	۱/۰۷۴	۰/۵۰۶	۰/۳۹۸	۰/۰۴۹	۰/۰۷۴	۰/۰۰۶	۰/۴۱۰	۰/۰۵۰	۰/۰۱۹
	(۱, ۱)	۱/۰۰۴	۱/۰۱۲	۰/۱۶۲	۰/۱۱۱	۰/۰۰۴	۰/۰۱۲	۰/۱۶۶	۰/۱۱۲	۰/۰۱۷
	(۱, ۲)	۱/۰۰۴	۲/۰۶۰	۰/۰۷۸	۰/۲۷۷	۰/۰۰۴	۰/۰۶۰	۰/۰۸۰	۰/۲۹۸	۰/۰۲۰
	(۲, ۰/۵)	۲/۴۶۸	۰/۵۲۰	۱/۵۳۵	۰/۰۸۹	۰/۴۶۸	۰/۰۲۰	۱/۶۳۱	۰/۰۹۳	۰/۱۳۵
	(۲, ۱)	۲/۰۰۷	۱/۰۱۱	۰/۳۲۴	۰/۱۱۱	۰/۰۰۷	۰/۰۱۱	۰/۳۳۲	۰/۱۱۲	۰/۰۳۴
(۲, ۲)	۲/۰۰۸	۲/۰۵۴	۰/۱۵۷	۰/۲۷۷	۰/۰۰۸	۰/۰۵۴	۰/۱۵۴	۰/۲۷۸	۰/۰۳۵	
۵۰۰	(۰/۵, ۰/۵)	۰/۵۲۶	۰/۵۰۵	۰/۱۵۱	۰/۰۳۷	۰/۰۲۶	۰/۰۰۵	۰/۱۴۹	۰/۰۳۷	۰/۰۰۵
	(۰/۵, ۱)	۰/۵۰۲	۱/۰۰۷	۰/۰۶۳	۰/۰۸۵	۰/۰۰۲	۰/۰۰۷	۰/۰۶۱	۰/۰۸۲	۰/۰۰۴
	(۰/۵, ۲)	۰/۵۰۱	۲/۰۳۳	۰/۰۳۰	۰/۲۰۹	۰/۰۰۱	۰/۰۳۳	۰/۰۳۰	۰/۲۱۲	۰/۰۰۵
	(۱, ۰/۵)	۱/۰۳۵	۰/۵۰۲	۰/۲۹۹	۰/۰۳۷	۰/۰۳۵	۰/۰۰۲	۰/۲۹۹	۰/۰۳۷	۰/۰۱۱
	(۱, ۱)	۱/۰۱۴	۱/۰۱۵	۰/۱۲۷	۰/۰۸۶	۰/۰۱۴	۰/۰۱۵	۰/۱۳۸	۰/۰۸۷	۰/۰۱۰
	(۱, ۲)	۱/۰۰۴	۲/۰۴۴	۰/۰۶۱	۰/۲۱۰	۰/۰۰۴	۰/۰۶۳	۰/۰۶۱	۰/۲۲۶	۰/۰۱۲
	(۲, ۰/۵)	۲/۰۹۸	۰/۵۰۵	۰/۶۰۴	۰/۰۳۷	۰/۰۹۸	۰/۰۰۵	۰/۵۹۳	۰/۰۳۷	۰/۰۲۱
	(۲, ۱)	۲/۰۲۹	۱/۰۱۴	۰/۲۵۳	۰/۰۸۶	۰/۰۲۹	۰/۰۱۴	۰/۲۵۹	۰/۰۸۸	۰/۰۲۱
(۲, ۲)	۲/۰۰۷	۲/۰۳۹	۰/۱۲۱	۰/۲۰۹	۰/۰۰۷	۰/۰۳۹	۰/۱۱۸	۰/۲۱۳	۰/۰۲۰	

مجموعه داده اول

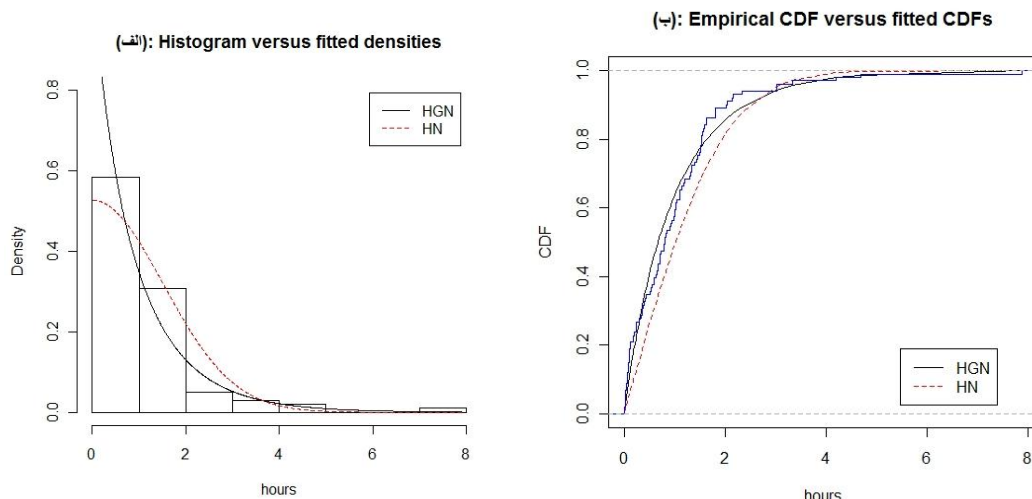
اولین مجموعه‌ی داده‌ها نشان‌دهنده زمان بین شکست خستگی آلیاژ کولار ۴۹ که در معرض فشار ثابت در سطح تنش ۹۰ درصد قرار می‌گیرد تا زمانی که همه آن شکسته می‌شود، است. برای مشاهده پژوهش‌های قبلی با این مجموعه داده، اندروز و هرزبرگ [۱] در سال ۱۹۸۵ و بارلو و همکاران [۴] در سال ۱۹۸۴ را ببینید.

جدول ۲ برآورد پارامترها را برای توزیع‌های نیم-نرمال و نیم-نرمال-تعمیم‌یافته با استفاده از روش ماکسیمم درست‌نمایی (MLE) به همراه مقادیر AIC ، $AICc$ ، BIC ، آماره‌ی آزمون‌های اندرسون-دارلینگ، کرامر-فون میزس، کولموگروف اسمیرنوف و p -مقدار متناظر نشان می‌دهد.

برای این مجموعه داده، AIC ، $AICc$ و BIC برازش را بهتر را برای مدل نیم-نرمال-تعمیم‌یافته نشان می‌دهد. آماره‌های اندرسون-دارلینگ، کرامر-فون میزس و کولموگروف اسمیرنوف نیز بر برازش بهتر داده‌ها به توزیع نیم-نرمال-تعمیم‌یافته نسبت به توزیع نیم-نرمال تأکید دارند. p -مقدار مربوط به آماره‌ی کولموگروف اسمیرنوف حاکی از آن است که مدل نیم-نرمال-تعمیم‌یافته برای این مجموعه از داده‌ها نیست.

در شکل ۶ الف توابع چگالی مدل‌های برازش شده به این مجموعه داده در کنار نمودار هیستوگرام داده‌ها رسم شده است که حاکی از آن است که توزیع نیم-نرمال-تعمیم‌یافته برازش بهتری را برای این مجموعه داده فراهم می‌کند. در شکل ۶ ب تابع توزیع مدل‌های برازش شده در برابر تابع توزیع تجربی مربوط به داده‌ها رسم شده که نتیجه قبلی را تأیید می‌کند.

QQ-plot های رسم شده در شکل ۷ نیز برازش بهتر توزیع نیم-نرمال-تعمیم‌یافته نسبت به توزیع نیم-نرمال را نشان می‌دهد.



شکل ۶. نمودار تابع چگالی و تابع توزیع مدل‌های برازش داده شده به مجموعه داده تنش-گسیختگی

مجموعه داده‌ی دوم

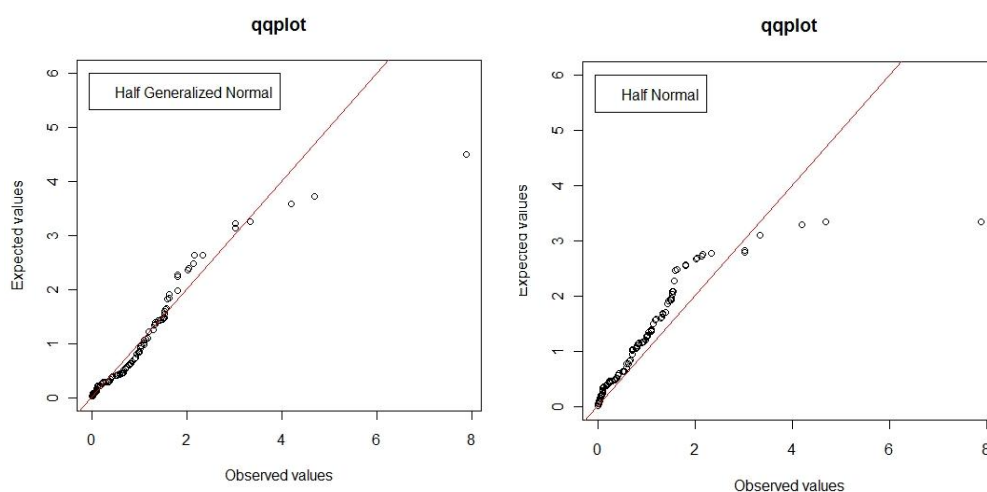
دومین مجموعه از داده‌ها مربوط به زمان بین خرابی‌های پیاپی در حوادث پمپ‌های کمکی رآکتور است که این مجموعه داده را سوپراواردانا و همکاران [۱۰] در سال ۱۹۹۹ بررسی کردند.

برآورد پارامترها را برای توزیع‌های نیم-نرمال و نیم-نرمال-تعمیم‌یافته با استفاده از روش ماکسیمم درست‌نمایی به همراه مقادیر AIC ، $AICc$ ، BIC ، آماره‌ی آزمون‌های اندرسون-دارلینگ، کرامر-فون میزس، کولموگروف اسمیرنوف و p -مقدار متناظر در جدول ۲ به دست آمده است.

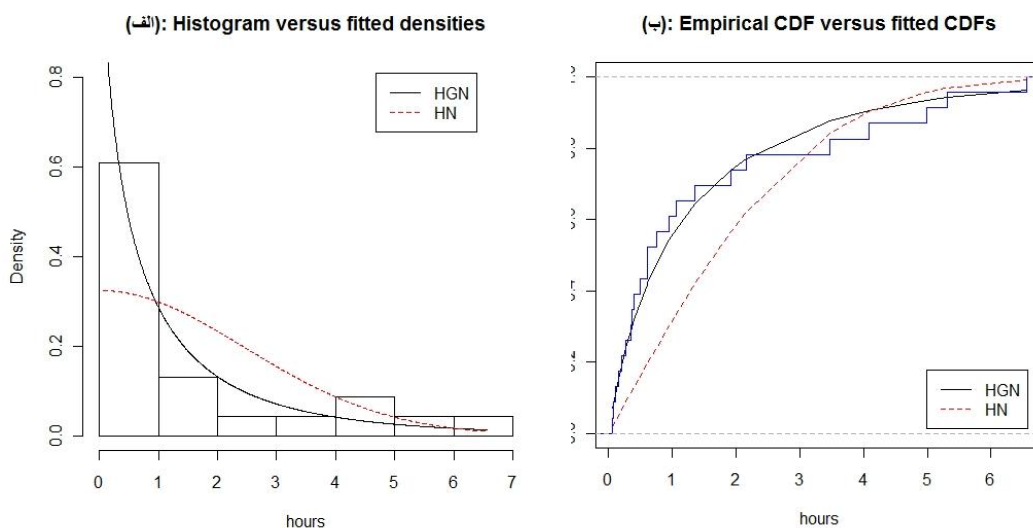
برای این مجموعه داده، مقادیر AIC ، $AICc$ ، BIC برازش را بهتر را برای مدل نیم-نرمال-تعمیم‌یافته نشان می‌دهند. آماره‌های اندرسون-دارلینگ، کرامر-فون میزس و کولموگروف اسمیرنوف نیز بر برازش بهتر داده‌ها به توزیع

نیم-نرمال تعمیم‌یافته نسبت به توزیع نیم-نرمال تأکید دارند. p -مقدار مربوط به آماره‌ی کولموگروف اسمیرنوف حکایت از آن دارد که در سطح معنی‌داری ۵ درصد فرض مناسب بودن توزیع نیم-نرمال برای برازش شدن به این مجموعه از داده‌ها پذیرفته نیست.

در شکل ۸ الف توابع چگالی مدل‌های برازش شده به این مجموعه داده در کنار نمودار هیستوگرام داده‌ها رسم شده است که حاکی از آن است که توزیع نیم-نرمال تعمیم‌یافته برازش بهتری را برای این مجموعه داده فراهم می‌کند. در شکل ۸ ب تابع توزیع مدل‌های برازش شده در برابر تابع توزیع تجربی مربوط به داده‌ها رسم شده که نتیجه قبلی را تأیید می‌کند.

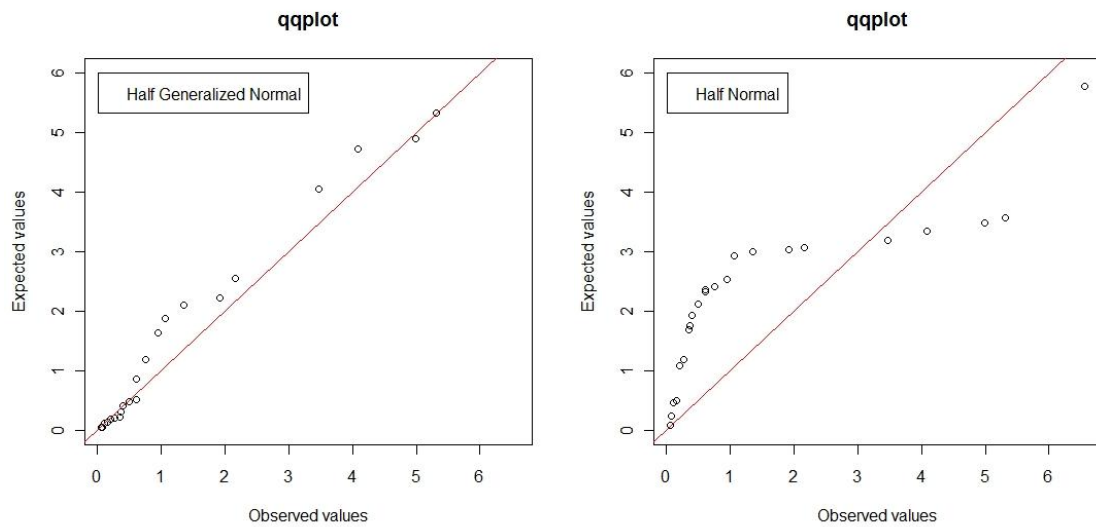


شکل ۷. نمودار QQ-plot مجموعه داده تنش-گسیختگی در برابر توزیع‌های نیم-نرمال و نیم-نرمال تعمیم‌یافته



شکل ۸. نمودار توابع چگالی و توزیع مدل‌های برازش شده به مجموعه داده زمان بین خرابی‌های پمپ‌های کمکی رآکتور

QQ-plot های رسم شده در شکل ۹ نیز برازش بهتر توزیع نیم-نرمال تعمیم‌یافته نسبت به توزیع نیم-نرمال را نشان می‌دهد.



شکل ۹. نمودار QQ-plot مجموعه داده‌ی زمان بین خرابی‌های پمپ‌های کمکی رآکتور در برابر توزیع‌های نیم-نرمال و نیم-نرمال تعمیم‌یافته

جدول ۲. برآورد پارامترها، انحراف استاندارد برآوردها، مقادیر AIC ، $AICc$ ، BIC ، آماره‌های اندرسون-دارلینگ،

کرامر-فون میزس، کولموگوروف-اسمیرنوف و p -value برای توزیع‌های نیم-نرمال و نیم-نرمال تعمیم‌یافته

	مجموعه داده‌ی اول		مجموعه داده‌ی دوم	
	HGN	HN	HGN	HN
$\hat{\sigma}(Sd)$	۰/۸۴۰ (۰/۲۸۷)	۱/۵۱۴ (۰/۱۰۶)	۰/۴۴۵ (۰/۴۹۹)	۲/۴۶۱ (۰/۳۶۳)
$\hat{s}(Sd)$	۰/۸۸۲ (۰/۱۶۸)	—	۰/۵۷۴ (۰/۱۹۷)	—
AIC	۲۱۰/۵۰۷	۲۳۲/۳۳۳	۶۸/۴۳۰	۷۶/۸۰۷
AICc	۲۱۰/۶۳۰	۲۳۲/۳۷۴	۶۹/۰۳۰	۷۶/۹۹۸
BIC	۲۰۹/۲۸۵	۲۳۳/۱۱۱	۶۵/۷۲۷	۷۶/۱۰۵
AD	۱/۲۲۷	۶/۲۹۰	۰/۳۰۸	۴/۵۸۷
CM	۰/۳۸۰	۰/۹۶۶	۰/۱۲۲	۰/۸۷۹
KS	۰/۰۸۷	۰/۱۴۱	۰/۱۰۱	۰/۳۲۷
p-value	۰/۴۳۶	۰/۰۳۵	۰/۹۵۶	۰/۰۱۱

بحث و نتیجه‌گیری

در این مقاله تعمیم جدیدی از توزیع نیم-نرمال، تحت عنوان توزیع نیم-نرمال تعمیم‌یافته بیان شد. این توزیع در حالت خاص توزیع نیم-نرمال را در بر می‌گیرد.

برای این توزیع برخی ویژگی‌های ریاضی بررسی شد. هم‌چنین تابع چگالی احتمال، تابع توزیع، گشتاور مرتبه r ام، ضرایب چولگی و کشیدگی و تابع مولد گشتاور استخراج شد. برخی از جنبه‌های استنباطی مربوط به برآوردهای

ماکسیمم درست‌نمایی بحث شد و در پایان انعطاف‌پذیری این توزیع با به کار گرفتن آن در برازش کردن به مجموعه داده‌های واقعی بررسی شد. با توجه به نتایج حاصل از مقایسه در بخش کاربردها می‌توان نتیجه گرفت که این تعمیم توزیع نیم-نرمال توانمندتر از توزیع نیم-نرمال است.

تشکر و قدردانی

نهایت تشکر و سپاس را از سردبیر مجله و اعضای هیات تحریریه آن دارند. هم‌چنین از داوران مقاله، که با نظرات ارزشمند خود بر غنای علمی مقاله افزودند، تشکر و قدردانی ویژه می‌نمایند. از دانشگاه یزد نیز بخاطر حمایت و پشتیبانی از این مقاله قدردانی می‌گردد.

منابع

1. Andrews D. F., Herzberg A. M., "Data: A Collection of Problems from Many Fields for the Student and Researcher Work", Springer Series in Statistics, New York (1985).
2. Azzalini A., "A class of distributions which includes the normal ones", *Scandinavian Journal of Statistics*, 12 (1985) 171-178.
3. Azzalini A., Capitanio A., "Statistical applications of the multivariate skew-normal distribution", *Journal of the Royal Statistical Society, Series B*, 61 (1999) 579-602.
4. Barlow R. E., Toland E., Freeman T., "A Bayesian Analysis of Stress-Rupture Life of Kevlar 49/Epoxy Spherical Pressure Vessels", In: *Proceedings of the Canadian Conference in Applied Statistics*. Marcel Dekker, New York (1984).
5. Cooray K., Ananda M., "A generalization of the half-normal distribution with applications to lifetime data", *Communications in Statistics-Theory and Methods*, 37 (2008) 1323-1337.
6. Daniel C., "Use of half normal plots in interpreting factorial two-level experiments", *Technometrics*, 1 (1959) 311-341.
7. Johnson N. L., Kotz S., Balakrishnan N., "Continuous Univariate Distributions", Vol. 1, 2nd Ed.; Wiley: New York (1994).
8. Nadarajah S., "A generalized normal distribution", *Journal of Applied Statistics*, 32 (5) (2005) 685-694.
9. Pewsey A., "Problems of inference for Azzalini's skew-normal distribution", *Journal of Applied Statistics*, 27 (2000) 859-870.
10. Suprawhardana M. S., Prayoto S. M., Sangadji, "Total time on test plot analysis for mechanical components of the RSG-GAS reactor". *Atom Indones*, 25 (1999) 81-93.