

توزیع نرمال-یکنواخت

حمزه ترابی، نرگس منتظری همدش، فاطمه قاسمیان دیانی
گروه آمار، دانشگاه یزد

تاریخ دریافت: ۱۳۹۲/۶/۱۸ تاریخ آخرین بازنگری: ۱۳۹۲/۱۲/۲۲

چکیده: در این مقاله، نخست خانواده‌های مختلفی که با هسته‌های توزیع‌های بتا تعمیم یافته، بتا، کومار، گاما تعمیم یافته، گاما، وایبول، لگ-گاما و لوژستیک ساخته شده است، مرور می‌شود. سپس یک خانواده کلی از توزیع‌ها با هسته توزیع نرمال پیشنهاد می‌شود. حالت خاصی از این خانواده، توزیع نرمال-یکنواخت، تعریف و ویژگی‌های مختلف آن استخراج شده است. پارامترهای این توزیع با روش ماکسیمم درست‌نمایی و مینیمم فاصله برآورد شده است. در پایان، این توزیع جدید به صورت کارا برای تحلیل یک مجموعه داده بقای واقعی به کار برده می‌شود.

واژه‌های کلیدی: برآوردگر ماکسیمم درست‌نمایی، برآوردگر مینیمم فاصله، توزیع نرمال-یکنواخت، نرخ شکست.

۱ مقدمه

توزیع‌های جدید معرفی شده در سال‌های اخیر، قابلیت انعطاف‌پذیری بیشتری برای برازش به داده‌ها را دارند و باعث می‌شوند بتوانیم مدل‌های بهتری به داده‌های موجود برازش دهیم؛ بنابراین تلاش این مقاله، ساخت خانواده جدیدی از توزیع‌ها

آدرس الکترونیک مسئول مقاله: حمزه ترابی، htorabi@yazd.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): ۶۲N۰۳، ۶۲F۱۰، ۶۲E۱۵

است که بتواند برای برخی از داده‌ها، مدل بهتری ارائه دهد.

فرض کنید G و g به ترتیب تابع توزیع انباشته و تابع چگالی متغیر تصادفی پیوسته Y با دامنه $S_Y = (a, b)$ باشند، که در آن $a < b$ و $a, b \in [-\infty, \infty]$. همچنین فرض کنید X یک متغیر تصادفی پیوسته با تابع توزیع انباشته F و تابع چگالی f باشد. آلزاتره و همکاران (۲۰۱۳) خانواده G_F را به صورت

$$G_F(x) = \int_a^{W(F(x))} g(t) dt = G(W(F(x))) \quad (1)$$

معرفی کردند، که در آن $R \rightarrow [0, 1] : W$ تابعی مشتق‌پذیر و اکیداً صعودی است به طوری که $\lim_{x \rightarrow 0^+} W(x) = a$ و $\lim_{x \rightarrow 1^-} W(x) = b$. توزیع جدید G_F با ترکیب خاصی از دو توزیع F و G ساخته می‌شود، که دامنه آن با توزیع F برابر است اما با انتخاب W خاصی، توزیع جدید G_F تعمیمی از توزیع F می‌شود. توزیع G را هسته و توزیع F را جد^۱ توزیع G_F می‌گویند.

ایوجین و همکاران (۲۰۰۲) و جونز (۲۰۰۴) با در نظر گرفتن g به عنوان تابع چگالی بتا و قرار دادن $W(F(x)) = F(x)$ خانواده توزیع‌های ساخته شده با هسته بتا را معرفی کردند. زوگرافوس و بالاکریشنان (۲۰۰۹) و ترابی و منتظری (۲۰۱۲a) با در نظر گرفتن g به عنوان تابع چگالی گاما و قرار دادن $W(F(x)) = \frac{F(x)}{\sqrt{1-F(x)}}$ و $W(F(x)) = -\ln(1-F(x))$ تولید شده از هسته گاما را معرفی کردند. ترابی و منتظری (۲۰۱۲b) با در نظر گرفتن g به عنوان تابع چگالی لوژستیک و قرار دادن $W(F(x)) = \ln(\frac{F(x)}{1-F(x)})$ خانواده تولید شده از هسته لوژستیک را معرفی کردند. امینی و همکاران (۲۰۱۲) با در نظر گرفتن g به عنوان تابع چگالی لگ-گاما و قرار دادن $W(F(x)) = F(x)$ خانواده توزیع‌های ساخته شده با هسته لگ-گاما را معرفی کرد. بنابراین می‌توان گفت آلزاتره و همکاران (۲۰۱۳) با ارائه رابطه (۱) کلاس جامع‌تری شامل این خانواده‌ها را معرفی کردند. در حالت کلی، انتخاب تابع W مناسب، وابسته به دامنه تکیه‌گاه متغیر تصادفی Y است.

در این مقاله، ابتدا توزیع‌های ساخته شده از هسته‌های مختلف بتا تعمیم یافته،

^۱ Parent

گاما تعمیم یافته، لگ-گاما و لوژستیک مرور می‌شوند. سپس با انتخاب هسته جدید نرمال و استفاده از $W(F(x)) = \Phi^{-1}(F(x))$ خانواده تعمیم یافته‌ای از توزیع‌ها بر اساس هسته نرمال معرفی می‌شود. با انتخاب مناسب این تابع W ، خانواده G_F معرفی می‌شود که علاوه بر اینکه تعمیمی از توزیع F است، دارای قابلیت انعطاف‌پذیری بالایی (متقارن، گودالی شکل، چوله به راست، چوله به چپ، 'J' وارون و 'J' شکل) نیز هست. سپس در زیر بخش ۱.۲ با انتخاب توزیع یکنواخت برای F یک توزیع ویژه از این خانواده، به نام نرمال-یکنواخت ارائه شده است. در پایان، سایر ویژگی‌های این توزیع جدید در بخش‌های ۳ و ۴ بررسی خواهد شد.

۱.۱ توزیع‌های تولید شده از هسته بتا تعمیم یافته

فرض کنید متغیر تصادفی X دارای توزیع بتا تعمیم یافته با تابع توزیع انباشته

$$G(x) = \frac{\gamma}{B(\alpha, \beta)} \int_0^x t^{\alpha\gamma-1} (1-t)^\beta dt, \quad 0 < x < 1$$

باشد، که در آن $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$ تابع بتا است. توزیع بتا و توزیع کومار به ترتیب به‌ازای $\gamma = 1$ و $\alpha = 1$ حالت خاصی از توزیع بتا تعمیم یافته هستند.

الکساندر و همکاران (۲۰۱۲) با شروع از یک تابع توزیع انباشته F و قرار دادن آن به‌عنوان کران بالای G کلاس توزیع‌های تولید شده از هسته بتا تعمیم یافته را به‌صورت

$$G_F(x) = \frac{\gamma}{B(\alpha, \beta)} \int_0^{F(x)} t^{\alpha\gamma-1} (1-t)^\beta dt, \quad \alpha, \beta, \gamma > 0 \quad (2)$$

معرفی کردند، که تابع چگالی (۲) به‌صورت

$$g_F(x) = \frac{1}{B(\alpha, \beta)} f(x) (F(x))^{\alpha-1} (\bar{F}(x))^{\beta-1}$$

است، که در آن $\bar{F}(x) = 1 - F(x)$. این یک خانواده کلی تولید شده با هسته بتا نامیده می‌شود. نکته قابل توجه برای این خانواده آن است که دامنه توزیع‌های G_F و

۱۷۲ توزیع نرمال-یکنواخت

F برابر هستند. توزیع جدید ساخته شده تعمیمی از توزیع F است؛ زیرا اگر
 $g_F(x) = f(x)$ آنگاه $\alpha = \beta = \gamma = 1$

برای شبیه‌سازی از این خانواده اگر متغیر تصادفی Y دارای توزیع بتا تعمیم یافته باشد، آنگاه متغیر تصادفی $X = F^{-1}(Y)$ دارای تابع توزیع G_F است.

• برای $\gamma = 1$ ، جونز (۲۰۰۴) کلاس توزیع‌های تولید شده از هسته بتا را در (۲) معرفی کرد و آن را خانواده تولید شده از بتا نامید.

• برای $\alpha = 1$ ، کوردریو و کاستر (۲۰۱۱) با شروع از یک تابع توزیع انباشته F و قرار دادن آن به‌عنوان کران بالای G کلاس توزیع‌های تولید شده از هسته کومار را به‌صورت

$$G_F(x) = 1 - (1 - (F(x))^\gamma)^\beta, \quad \gamma > 0, \beta > 0$$

معرفی کردند.

۲.۱ توزیع‌های تولید شده از هسته گاما تعمیم یافته

فرض کنید متغیر تصادفی X دارای توزیع گامای تعمیم یافته با تابع توزیع

$$G(x) = \frac{\gamma}{\Gamma(\alpha)\beta^{\alpha\gamma}} \int_0^x e^{-(\frac{t}{\beta})^\gamma} t^{\alpha\gamma-1} dt, \quad x > 0$$

باشد، که در آن $\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$ تابع گاما است. توزیع گاما و توزیع وایبول به‌ترتیب به‌ازای $\gamma = 1$ و $\alpha = 1$ حالت خاصی از توزیع گاما تعمیم یافته هستند.

تابع توزیع دو کلاس از توزیع‌های تعمیم یافته از متغیر تصادفی گاما به‌صورت زیر هستند:

$$G_F(x) = \frac{\gamma}{\Gamma(\alpha)\beta^{\alpha\gamma}} \int_0^{-\ln(\bar{F}(x))} e^{-(\frac{t}{\beta})^\gamma} t^{\alpha\gamma-1} dt, \quad (3)$$

$$G_{\bar{F}}(x) = 1 - \frac{\gamma}{\Gamma(\alpha)\beta^{\alpha\gamma}} \int_0^{-\ln(F(x))} e^{-(\frac{t}{\beta})^\gamma} t^{\alpha\gamma-1} dt. \quad (4)$$

تابع چگالی آن‌ها به ترتیب عبارتند از:

$$g_F(x) = \frac{\gamma f(x)}{\Gamma(\alpha)\beta^\alpha \bar{F}(x)} e^{-\left(\frac{-\ln(\bar{F}(x))}{\beta}\right)^\gamma} (-\ln(\bar{F}(x)))^{\alpha-1},$$

$$g_{\bar{F}}(x) = \frac{\gamma f(x)}{\Gamma(\alpha)\beta^\alpha F(x)} e^{-\left(\frac{-\ln(F(x))}{\beta}\right)^\gamma} (-\ln(F(x)))^{\alpha-1}.$$

اگر $\alpha = \beta = \gamma = 1$ ، آنگاه $g_F(x) = g_{\bar{F}}(x) = f(x)$.

اگر Y دارای توزیع گاما تعمیم یافته باشد، آنگاه متغیرهای تصادفی $X_1 = F^{-1}(1 - e^{-Y})$ و $X_2 = F^{-1}(e^{-Y})$ دارای توزیع‌هایی به ترتیب با تابع توزیع G_F و $G_{\bar{F}}$ هستند.

- برای $\beta = 1$ ، زوگرافوس و بالاکریشن (۲۰۰۹) کلاس توزیع‌های تولید شده از هسته گاما را در (۳) معرفی کردند و آن را خانواده تولید شده از گاما تعمیم یافته نامیدند.

- برای $\gamma = 1$ ، آلزاتره و همکاران (۲۰۱۳) خانواده G_F در (۳) را معرفی کردند و آن را خانواده گاما- X نامیدند.

- برای $\alpha = 1$ ، که حالت خاص را آلزاتره و همکاران (۲۰۱۳a) معرفی نموده و آن را خانواده وایبول- X نامیدند.

- برای $\beta = \gamma = 1$ ، خانواده (۴) توسط ریستیک و بالاکریشن (۲۰۱۱) معرفی شده است.

ترابی و منتظری (۲۰۱۲a) با جایگذاری $\frac{F(x)}{1-F(x)}$ به عنوان کران بالای تابع توزیع گاما، کلاس دیگری از توزیع‌های تولید شده از هسته گاما را به صورت

$$G_F(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{F(x)/\bar{F}(x)} e^{-\frac{t}{\beta}} t^{\alpha-1} dt$$

و تابع چگالی آن‌را به صورت

$$g_F(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \frac{f(x)}{\bar{F}^\alpha(x)} \exp\left(-\frac{F(x)}{\beta\bar{F}(x)}\right) \left(\frac{F(x)}{\bar{F}(x)}\right)^{\alpha-1}$$

۱۷۴ توزیع نرمال-یکنواخت

معرفی کردند. اگر Y دارای توزیع گاما باشد، آنگاه متغیر تصادفی $F^{-1}(\frac{Y}{1+Y})$ دارای تابع توزیع G_F است. در این روش توزیع جدید G_F تعمیمی از توزیع F نیست.

۳.۱ توزیع‌های تولید شده از هسته لگ-گاما

تابع توزیع لگ-گاما به صورت

$$G(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^x t^{\frac{1}{\beta}-1} (-\ln t)^{\alpha-1} dt, \quad 0 < x < 1$$

است. امینی و همکاران (۲۰۱۲) دو نوع توزیع با هسته لگ-گاما با تابع توزیع‌های

$$G_F(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{F(x)} t^{\frac{1}{\beta}-1} (-\ln t)^{\alpha-1} dt$$

و

$$G_{\bar{F}}(x) = 1 - \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{\bar{F}(x)} t^{\frac{1}{\beta}-1} (-\ln t)^{\alpha-1} dt$$

را معرفی کردند، که تابع چگالی‌های آن‌ها به ترتیب به صورت زیر هستند:

$$g_F(x) = \frac{f(x)}{\Gamma(\alpha)\beta^\alpha} (F(x))^{\frac{1}{\beta}-1} (-\ln(F(x)))^{\alpha-1}$$

و

$$g_{\bar{F}}(x) = \frac{f(x)}{\Gamma(\alpha)\beta^\alpha} (\bar{F}(x))^{\frac{1}{\beta}-1} (-\ln(\bar{F}(x)))^{\alpha-1}.$$

اگر $\alpha = \beta = 1$ ، آنگاه $g_F(x) = g_{\bar{F}}(x) = f(x)$ هرگاه Y دارای توزیع لگ-گاما باشد، آنگاه متغیرهای تصادفی $X_1 = F^{-1}(Y)$ و $X_2 = F^{-1}(1-Y)$ دارای توزیع‌هایی به ترتیب با تابع توزیع‌های G_F و $G_{\bar{F}}$ هستند.

۴.۱ توزیع‌های تولید شده از هسته لوژستیک

تابع توزیع لوژستیک به صورت

$$G(x) = \frac{1}{\sigma} \int_{-\infty}^x e^{-\frac{t-\mu}{\sigma}} (1 + e^{-\frac{t-\mu}{\sigma}})^{-2} dt = (1 + e^{-\frac{x-\mu}{\sigma}})^{-1}$$

است. ترابی و منتظری (۲۰۱۲b) با جایگذاری $\ln(\frac{F(x)}{\bar{F}(x)})$ به عنوان کران بالای G کلاس توزیع‌های تولید شده از هسته لوژستیک را به صورت

$$G_F(x) = (1 + e^{\mu/\sigma} (\frac{F(x)}{\bar{F}(x)})^{-1/\sigma})^{-1}, \quad (5)$$

با تابع چگالی

$$g_F(x) = \frac{f(x)}{\sigma \bar{F}(x) F(x)} e^{\mu/\sigma} (\frac{F(x)}{\bar{F}(x)})^{-1/\sigma} (1 + e^{\mu/\sigma} (\frac{F(x)}{\bar{F}(x)})^{-1/\sigma})^{-2}$$

معرفی کردند. اگر $\mu = 0$ و $\sigma = 1$ ، آنگاه $g_F(x) = f(x)$. برای شبیه‌سازی از این خانواده، اگر متغیر تصادفی Y دارای توزیع لوژستیک باشد، آنگاه متغیر تصادفی $X = F^{-1}(\frac{1}{1+e^{-Y}})$ دارای تابع توزیع G_F است. ترابی و منتظری (۲۰۱۲b) با جایگذاری توزیع یکنواخت در رابطه (۵)، توزیع لوژستیک-یکنواخت را معرفی کردند.

۲ توزیع‌های تولید شده از هسته نرمال

در این کلاس جدید، F تابع توزیع یک متغیر تصادفی پیوسته را نشان می‌دهد. این کلاس از توزیع‌های تعمیم یافته به صورت

$$G_F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\Phi^{-1}(F(x))} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \Phi\left(\frac{\Phi^{-1}(F(x)) - \mu}{\sigma}\right) \quad (6)$$

تعریف می‌شود و تابع چگالی آن نیز به صورت

$$g_F(x) = \frac{f(x)}{\sigma} \phi\left(\frac{\Phi^{-1}(F(x)) - \mu}{\sigma}\right) \frac{1}{\phi(\Phi^{-1}(F(x)))}$$

است. اگر $\mu = 0$ و $\sigma = 1$ ، آنگاه $g_F(x) = f(x)$. برای شبیه‌سازی از این خانواده اگر متغیر تصادفی Y دارای توزیع نرمال باشد، آنگاه متغیر تصادفی $X = F^{-1}(\Phi(Y))$ دارای تابع توزیع G_F است.

برای حالت ویژه‌ای که $F(x)$ تابع توزیع یکنواخت پیوسته $U(a, b)$ است، گوییم X دارای توزیع نرمال-یکنواخت است و با نماد $NU(\mu, \sigma, a, b)$ نشان داده می‌شود.

۱.۲ توزیع نرمال-یکنواخت

تابع توزیع نرمال-یکنواخت (NU) با اختیار کردن F به عنوان توزیع یکنواخت $U(a, b)$ به صورت

$$G_F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\Phi^{-1}\left(\frac{x-a}{b-a}\right)} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \Phi\left(\frac{\Phi^{-1}\left(\frac{x-a}{b-a}\right) - \mu}{\sigma}\right), \quad a < x < b \quad (\text{V})$$

است. تابع چگالی احتمال و تابع نرخ شکست این توزیع به ترتیب به صورت

$$g_F(x) = \frac{1}{(b-a)\sigma} \phi\left(\frac{\Phi^{-1}\left(\frac{x-a}{b-a}\right) - \mu}{\sigma}\right) \frac{1}{\phi\left(\Phi^{-1}\left(\frac{x-a}{b-a}\right)\right)}, \quad a < x < b \quad (\text{A})$$

و

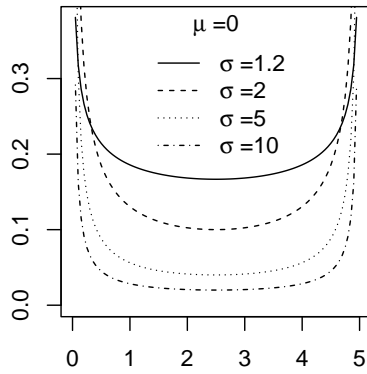
$$h_F(x) = \frac{g_F(x)}{1 - G_F(x)}$$

به دست می آیند. با توجه به رابطه (V) اگر $\mu = 0$ و $\sigma = 1$ ، آنگاه $g_F(x)$ همان تابع چگالی توزیع یکنواخت است. برای شبیه سازی از توزیع نرمال-یکنواخت اگر Y یک متغیر تصادفی از توزیع نرمال با پارامترهای μ و σ باشد، آنگاه $X = F^{-1}(\Phi(Y)) = a + (b-a)\Phi(Y)$ دارای توزیع $NU(\mu, \sigma, a, b)$ است.

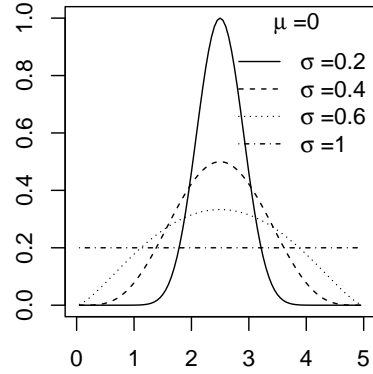
شکل ۱ برخی از توابع چگالی توزیع نرمال-یکنواخت را برای مقادیر گوناگون μ و σ به ازای $a = 0$ و $b = 5$ نشان می دهد.

با توجه به شکل ۷ انعطاف پذیری توزیع از نظر چولگی و کشیدگی و ساختار آن کاملاً مشهود است. خلاصه ای از نتایج شکل عبارتند از:

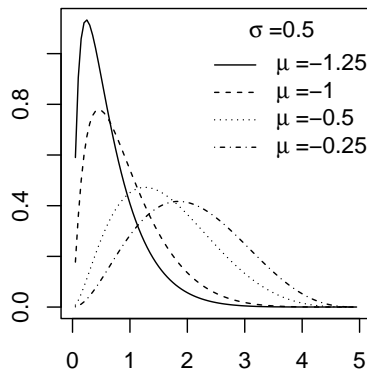
- برای $\mu = 0$ ، هنگامی که $\sigma \leq 1$ ، توزیع متقارن و تک مدی است؛ نمودار (الف) و هنگامی که $\sigma > 1$ توزیع گودالی شکل است؛ نمودار (ب).
- برای $\mu > 0$ و $\sigma < 1$ ، توزیع چوله به چپ است و با افزایش μ درجه چولگی چپ افزایش می یابد؛ نمودار (ج).
- برای $\mu < 0$ و $\sigma < 1$ ، توزیع چوله به راست است و با کاهش μ درجه چولگی راست افزایش می یابد؛ نمودار (د).



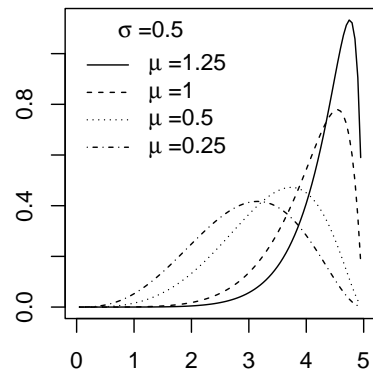
(ب)



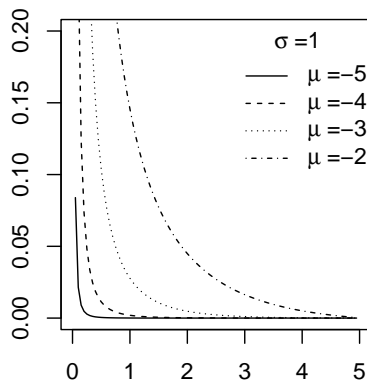
(الف)



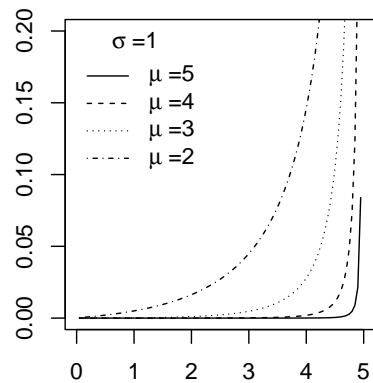
(د)



(ج)



(و)



(ه)

شکل ۱: تابع چگالی توزیع نرمال-یکنواخت برای $a = 0$ و $b = 5$

• برای $\sigma = 1$ هنگامی که $\mu < 0$ ، توزیع 'J' وارون شکل است؛ نمودار (ه) و هنگامی که $\mu > 0$ توزیع 'J' شکل است؛ نمودار (و).

شکل ۲ برخی از توابع نرخ شکست نرمال-یکنواخت را برای مقادیر گوناگون μ و σ به ازای $a = 0$ و $b = 5$ نشان می‌دهد. همان‌طور که ملاحظه می‌شود تابع نرخ شکست می‌تواند صعودی یا خم گودالی شکل باشد.

۳ گشتاورها و چندک‌ها

در حالت کلی برای گشتاورهای توزیع نرمال-یکنواخت نمی‌توان فرم بسته‌ای به دست آورد. مقادیر عددی واریانس، میانگین، چولگی، کشیدگی و میانه این توزیع در جدول ۱ ارائه شده‌اند.

برای محاسبه چندک این توزیع از معادله (۷) داریم: $p = \Phi\left(\frac{\Phi^{-1}\left(\frac{x-a}{b-a}\right) - \mu}{\sigma}\right)$ بنابراین چندک مرتبه p ام این توزیع برابر

$$x_p = a + (b - a)\Phi(\sigma\Phi^{-1}(p) + \mu).$$

با قرار دادن $p = 0.5$ ، میانه این توزیع $x_{0.5} = a + (b - a)\Phi(\mu)$ است، که به پارامتر σ بستگی ندارد.

۴ برآورد پارامترها

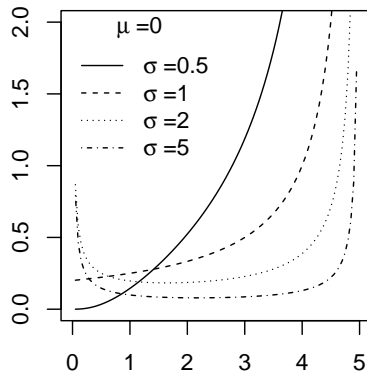
۱.۴ برآورد ماکسیمم درست‌نمایی

با توجه به فرم تابع چگالی تابع توزیع نرمال-یکنواخت تابع درست‌نمایی برای یک نمونه تصادفی X_1, \dots, X_n از چگالی (۸) به صورت

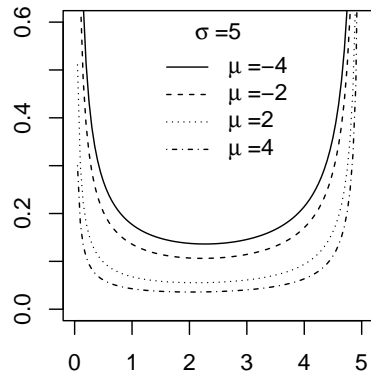
$$L(\mu, \sigma, a, b) = \frac{\prod_{i=1}^n \phi\left(\frac{\Phi^{-1}\left(\frac{x_i-a}{b-a}\right) - \mu}{\sigma}\right)}{(b-a)^n \sigma^n \prod_{i=1}^n \phi\left(\Phi^{-1}\left(\frac{x_i-a}{b-a}\right)\right)}, \quad a < x < b$$

است. بنابراین

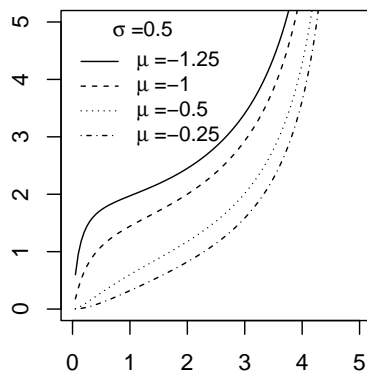
$$\log L(\mu, \sigma, a, b) = -n \log(b-a) - n \log \sigma + \sum_{i=1}^n \log \phi\left(\frac{\Phi^{-1}\left(\frac{x_i-a}{b-a}\right) - \mu}{\sigma}\right)$$



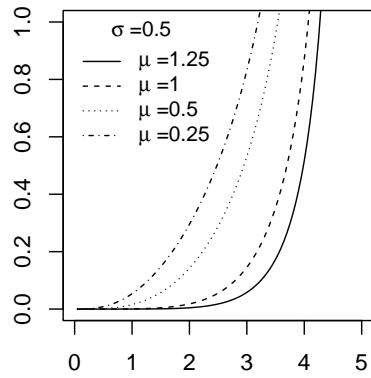
(ب)



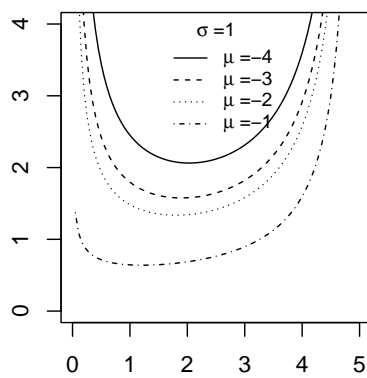
(الف)



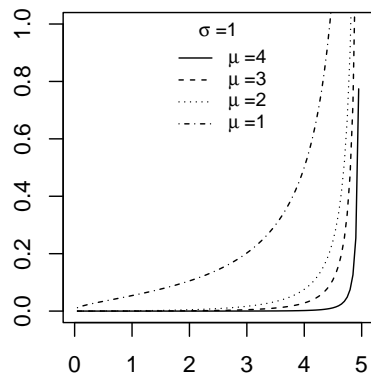
(د)



(ج)



(و)



(ه)

شکل ۲: تابع نرخ شکست نرمال-یکنواخت برای $a = 0$ و $b = 5$

جدول ۱: میانگین، واریانس، میانه، چولگی و کشیدگی توزیع $NU(\mu, \sigma, \circ, \delta)$

چولگی	کشیدگی	میانه	واریانس	میانگین	μ	σ
۰/۴۰۰	۱۰/۹۵۹	۲/۶۶۶	۰/۱۸۴	۰/۵۴۴	-۲	۰/۵
۰/۷۹۳	۱/۱۳۶	۱/۰۷۵	۰/۹۲۸	۰/۳۸۷	-۱	
۲/۵۰	-۰/۶۲۲	۰/۰۰۰	۲/۵۰۰	۰/۸۰۱	۰	
۴/۲۱۰	۱/۱۳۶	-۱/۰۷۵	۴/۰۲۷	۰/۳۸۷	۱	
۴/۸۷۰	۱۰/۹۵۹	-۲/۶۶۶	۴/۸۱۶	۰/۵۴۴	۲	
۰/۱۱۴	۸/۶۵۱	۲/۷۳۰	۰/۳۹۳	۰/۴۲۲	-۲	۱
۰/۷۹۳	۰/۳۴۷	۱/۰۹۹	۱/۱۹۹	۱/۳۹۳	-۱	
۲/۵۰۰	-۱/۲۰۰	۰/۰۰۰	۲/۵۰۰	۲/۰۸۳	۰	
۴/۲۱۰	۰/۳۴۷	-۱/۰۹۹	۳/۸۰۱	۱/۳۹۳	۱	
۴/۸۷۰	۸/۶۵۱	-۲/۷۳۰	۴/۶۰۷	۰/۴۲۲	۲	
۰/۱۱۴	-۱/۴۳۴	۰/۶۳۹	۱/۷۳۷	۴/۶۴۴	-۲	۵
۰/۷۹۳	-۱/۷۵۸	۰/۳۱۴	۲/۱۱۱	۵/۰۱۳	-۱	
۲/۵۰۰	-۱/۸۶۱	۰/۰۰۰	۲/۵۰۰	۵/۱۴۳	۰	
۴/۲۱۰	-۱/۷۵۸	-۰/۳۱۴	۲/۸۸۸	۵/۰۱۳	۱	
۴/۸۷۰	-۱/۴۳۴	-۰/۶۳۹	۳/۲۶۳	۴/۶۴۳	۲	

$$-\sum_{i=1}^n \log \phi(\Phi^{-1}(\frac{x_i - a}{b - a})), \quad a < x < b.$$

به دلیل وابسته بودن دامنه X به پارامترهای a و b ، تابع درست‌نمایی را نمی‌توان به صورت تحلیلی با مشتق گرفتن از آن به دست آورد و لازم است به صورت عددی حل شود. به این منظور برای محاسبه MLE از دستور mle و روش L-BFGS-B در بسته stats4 در نرم‌افزار R استفاده شده است.

۲.۴ برآورد مینیمم فاصله

فرض کنید X_1, \dots, X_n یک نمونه تصادفی به حجم n از تابع توزیع پیوسته $F(x; \theta)$ و $X_{i:n}$ ، i امین آماره ترتیبی این نمونه تصادفی باشد. مقدار θ که عبارت

$$\sum_{i=1}^{n+1} h(F(x_{i:n}; \theta) - F(x_{i-1:n}; \theta), \frac{1}{n+1}) \quad (9)$$

را مینیمم می‌کند، برآوردگر مینیمم فاصله توان‌های دوم (MSDE) است و با نماد $\hat{\theta}_{MSDE}$ نشان داده می‌شود و در آن $F(x_{n+1:n}; \theta) = 1$ و $F(x_{0:n}; \theta) = 0$ است. توجه کنید که $h(x, y)$ در رابطه (۹) یک فاصله است. برخی از انتخاب‌های مناسب

برای $h(x, y)$ ، فاصله قدر مطلق $|x - y|$ و فاصله قدر مطلق-لگاریتم $|\log x - \log y|$ است. برآوردگرهای متناظر با این فاصله‌ها به ترتیب به صورت برآوردگر مینیمم فاصله قدر مطلق (MSADE) و برآوردگر مینیمم فاصله قدر مطلق-لگاریتم (MSALDE) نام گذاری شده‌اند. این روش نخستین بار توسط ترابی (۲۰۰۸) معرفی شده است.

برای مقایسه نتایج روش‌های متفاوت برآورد، نمونه‌هایی به اندازه ۲۰، ۴۰، ۶۰ و ۱۰۰ از توزیع‌های $NU(-1, 0.5, 0, 5)$ ، $NU(1, 0.5, 0, 5)$ ، $NU(0, 0.5, 0, 5)$ و $NU(0, 2, 0, 5)$ شبیه‌سازی کرده و MSALDE، MSADE و MLE پارامترهای این توزیع به دست آمده است. سپس این عملیات ۱۰۰۰ بار تکرار شده و میانگین برآورد پارامترها (AE) و میانگین توان دوم خطاها (MSE) محاسبه و نتایج در جدول‌های ۲، ۳، ۴ و ۵ ارائه شده است. همان‌طور که ملاحظه می‌شود:

- برای $\sigma = 0.5$ و $\mu = -1$ که توزیع به ترتیب چوله به راست و چوله به چپ است، برآوردگرهای MSALDE و MSADE نسبت به MLE میانگین توان دوم خطای کوچکتری دارند.
- برای $\sigma = 0.5$ و $\mu = 0$ که توزیع متقارن است، برآوردگر MSALDE میانگین توان دوم خطای کوچکتری دارد.
- برای $\sigma = 2$ و $\mu = 0$ که توزیع خم گودالی است، برآوردگر MLE میانگین توان دوم خطای کوچکتری دارد.

۵ کاربرد توزیع نرمال - یکنواخت

در این بخش، برای نشان دادن کاربردی از توزیع جدید در تحلیل بقا، داده‌های طول عمر آرسن (Aarset: ۱۹۸۷) در نظر گرفته شده است. داده‌های آرسن مربوط به طول عمر ۵۰ قطعه است.

جدول ۲: مقادیر شبیه‌سازی شده میانگین انواع برآورد پارامترها و میانگین توان دوم خطاها از توزیع $NU(-1, 0/5, 0, 5)$ با تکرار ۱۰۰۰

MSALDE		MSADE		MLE		n
MSE	AE	MSE	AE	MSE	AE	
۰/۱۲۵	-۰/۹۷۵	۰/۰۸۲	-۰/۹۵۲	۰/۴۴۸	-۰/۶۶۷	$\hat{\mu}$ ۲۰
۰/۰۴۱	۰/۵۶۴	۰/۰۲۴	۰/۵۴۷	۰/۲۶۸	۰/۹۹۲	$\hat{\sigma}$
۰/۱۰۵	-۰/۰۷۰	۰/۰۵۹	-۰/۰۵۱	۰/۴۳۵	۰/۰۸۰	\hat{a}
۸/۵۱۹	۵/۵۳۳	۰/۹۴۱	۵/۰۳۷	۵۵۹۴/۱	۱۵/۳۸۶	\hat{b}
۰/۰۷۲	-۱/۰۱۲	۰/۰۴۱	-۰/۹۸۹	۰/۲۸۰	-۰/۸۲۳	$\hat{\mu}$ ۴۰
۰/۰۲۰	۰/۵۲۴	۰/۰۱۳	۰/۵۱۷	۰/۰۹۲	۰/۷۱۱	$\hat{\sigma}$
۰/۰۲۳	-۰/۰۳۶	۰/۰۲۲	-۰/۰۳۳	۰/۰۱۳	۰/۰۶۵	\hat{a}
۳۲/۰۱۲	۵/۸۰۰	۱/۸۸۷	۵/۱۸۰	۳۱۲۳/۴	۱۲/۹۶۰	\hat{b}
۰/۰۵۸	-۱/۰۲۶	۰/۰۲۹	-۱/۰۰۰	۰/۱۹۸	-۰/۸۶۶	$\hat{\mu}$ ۶۰
۰/۱۱۵	۰/۵۱۹	۰/۰۰۹	۰/۵۱۵	۰/۰۵۶	۰/۶۵۶	$\hat{\sigma}$
۰/۰۱۰	-۰/۰۲۰	۰/۰۱۱	-۰/۰۱۸	۰/۰۰۸	۰/۰۵۳	\hat{a}
۹/۶۲۰	۵/۸۰۹	۱/۰۹۷	۵/۱۸۸	۹۷۵/۴۹	۸/۹۲۴	\hat{b}
۰/۰۶۱	-۱/۰۵۰	۰/۰۲۶	-۱/۰۰۷	۰/۱۶۱	-۰/۹۴۵	$\hat{\mu}$ ۱۰۰
۰/۰۰۹	۰/۴۹۲	۰/۰۰۶	۰/۴۹۸	۰/۰۲۷	۰/۵۷۹	$\hat{\sigma}$
۰/۰۰۶	۰/۰۶۷	۰/۰۰۶	۰/۰۶۷	۰/۰۰۴	۰/۰۳۳	\hat{a}
۲۹/۳۵۵	۶/۳۲۹	۲/۱۸۵	۵/۳۱۸	۶۸۳/۸۷	۸/۹۸۰	\hat{b}

جدول ۳: مقادیر شبیه‌سازی شده میانگین انواع برآورد پارامترها و میانگین توان دوم خطاها از توزیع $NU(1, 0/5, 0, 5)$ با تکرار ۱۰۰۰

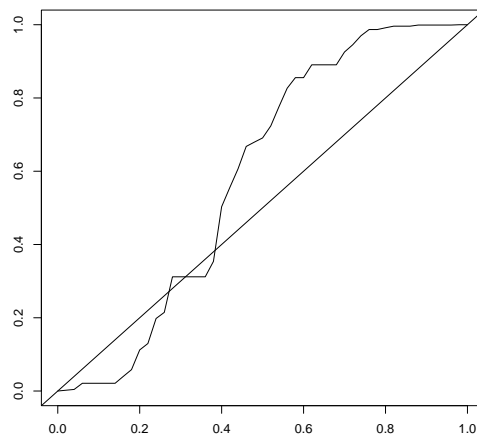
MSALDE		MSADE		MLE		n
MSE	AE	MSE	AE	MSE	AE	
۰/۱۳۲	۰/۹۶۰	۰/۰۹۴	۰/۹۴۳	۰/۴۵۴	۰/۶۴۲	$\hat{\mu}$ ۲۰
۰/۰۴۵	۰/۵۷۰	۰/۰۲۵	۰/۵۴۰	۰/۲۸۳	۰/۹۴۷	$\hat{\sigma}$
۵/۹۸۷	-۰/۳۹۵	۱/۴۶۶	-۰/۰۱۲	۴۰۵۷/۷	-۸/۴۱۴	\hat{a}
۰/۱۳۰	۵/۰۶۵	۰/۰۷۰	۵/۰۶۲	۰/۱۲۳	۴/۸۹۸	\hat{b}
۰/۰۷۱	۱/۰۰۷	۰/۰۳۸	۰/۹۸۶	۰/۳۱۰	۰/۸۴۰	$\hat{\mu}$ ۴۰
۰/۰۲۰	۰/۵۲۸	۰/۰۱۴	۰/۵۱۸	۰/۰۹۹	۰/۷۱۵	$\hat{\sigma}$
۸/۷۱۸	-۰/۷۴۴	۱/۴۷۶	-۰/۱۸۳	۲۰۶۰/۵	-۸/۱۷۴	\hat{a}
۰/۰۲۲	۵/۰۳۴	۰/۰۲۳	۵/۰۳۵	۰/۰۱۶	۴/۹۳۹	\hat{b}
۰/۰۶۶	۱/۰۲۴	۰/۰۲۳	۰/۹۹۰	۰/۲۵۲	۰/۸۹۴	$\hat{\mu}$ ۶۰
۰/۰۱۳	۰/۵۱۲	۰/۰۱۰	۰/۵۱۲	۰/۰۵۶	۰/۶۴۴	$\hat{\sigma}$
۲۴/۶۰۲	-۱/۰۵۱	۱/۵۴۹	-۰/۱۸۶	۱۹۷۲/۹	-۷/۸۲۶	\hat{a}
۰/۰۱۲	۵/۰۲۴	۰/۰۱۳	۵/۰۲۲	۰/۰۱۰	۴/۹۵۳	\hat{b}
۰/۰۴۹	۱/۰۳۰	۰/۰۲۵	۱/۰۰۸	۰/۱۳۳	۰/۹۰۸	$\hat{\mu}$ ۱۰۰
۰/۰۰۹	۰/۵۰۲	۰/۰۰۷	۰/۴۹۹	۰/۰۲۷	۰/۵۹۱	$\hat{\sigma}$
۱۲/۹۸۶	-۰/۹۱۴	۲/۶۵۸	-۰/۳۶۱	۹۱۳/۰	-۳/۰۰۶	\hat{a}
۰/۰۰۵	۵/۰۲۰	۰/۰۰۶۰	۵/۰۲۱	۰/۰۰۴	۴/۹۷۰	\hat{b}

جدول ۴: مقادیر شبیه‌سازی شده میانگین انواع برآورد پارامترها و میانگین توان دوم خطاها از توزیع $NU(0, 0/5, 0, 5)$ با تکرار ۱۰۰۰

MSALDE		MSADE		MLE		n
MSE	AE	MSE	AE	MSE	AE	
0/171	0/020	0/105	0/011	0/125	-0/001	$\hat{\mu}$ 20
0/061	0/575	0/036	0/571	0/343	0/994	$\hat{\sigma}$
13/784	-0/644	0/912	-0/091	34/275	0/014	\hat{a}
57/306	5/823	0/603	5/021	34/718	5/141	\hat{b}
0/092	0/011	0/067	0/000	0/069	0/001	$\hat{\mu}$ 40
0/040	0/522	0/023	0/528	0/077	0/740	$\hat{\sigma}$
7/286	-0/673	0/646	-0/125	8/0915	-0/021	\hat{a}
7/664	5/643	0/697	5/138	10/340	5/112	\hat{b}
0/066	-0/004	0/050	0/002	0/048	-0/001	$\hat{\mu}$ 60
0/028	0/493	0/016	0/513	0/065	0/647	$\hat{\sigma}$
9/276	-0/717	0/670	-0/188	4/993	-0/022	\hat{a}
7/742	5/706	0/749	5/189	6/569	5/072	\hat{b}
0/036	0/000	0/030	0/003	0/025	0/003	$\hat{\mu}$ 100
0/012	0/484	0/012	0/499	0/029	0/590	$\hat{\sigma}$
5/474	-0/480	0/514	-0/193	1/816	0/083	\hat{a}
4/364	5/503	0/534	5/197	1/446	4/908	\hat{b}

جدول ۵: مقادیر شبیه‌سازی شده میانگین انواع برآورد پارامترها و میانگین توان دوم خطاها از توزیع $NU(0, 2, 0, 5)$ با تکرار ۱۰۰۰

MSALDE		MSADE		MLE		n
MSE	AE	MSE	AE	MSE	AE	
0/164	-0/021	0/133	-0/027	0/132	-0/031	$\hat{\mu}$ 20
0/103	1/771	0/116	1/728	0/133	1/679	$\hat{\sigma}$
0/003	-0/010	0/006	-0/016	0/001	0/003	\hat{a}
0/0008	5/018	0/028	5/027	0/002	4/995	\hat{b}
0/164	-0/010	0/103	0/007	0/073	-0/008	$\hat{\mu}$ 40
0/103	1/767	0/164	1/624	0/124	1/675	$\hat{\sigma}$
0/000	-0/009	0/001	-0/011	0/000	-0/008	\hat{a}
0/000	5/009	0/001	5/011	0/000	5/008	\hat{b}
0/127	-0/007	0/080	-0/009	0/053	-0/013	$\hat{\mu}$ 60
0/068	1/836	0/182	1/596	0/116	1/677	$\hat{\sigma}$
0/000	-0/010	0/000	-0/010	0/000	-0/010	\hat{a}
0/000	5/009	0/000	5/010	0/000	5/009	\hat{b}
0/078	-0/002	0/060	-0/008	0/031	0/003	$\hat{\mu}$ 100
0/032	1/911	0/179	1/593	0/110	1/678	$\hat{\sigma}$
0/000	-0/010	0/000	-0/010	0/000	-0/010	\hat{a}
0/000	5/010	0/000	5/010	0/000	5/010	\hat{b}



شکل ۳: نمودار TTT داده‌های آراست.

نمودار TTT^۲ داده‌های آراست (۱۹۸۷) نشان دهنده شکل تابع نرخ شکست تجربی داده‌ها است. شکل‌های مختلف نمودار TTT عبارتند از: ثابت، مقعر، محدب، مقعر-محدب و محدب-مقعر که به ترتیب نشان دهنده تابع نرخ شکست ثابت، صعودی، نزولی، تک مدی و گودالی شکل است.

نمودار TTT داده‌های آراست در شکل ۳ محدب-مقعر است، که نشان دهنده یک تابع نرخ شکست گودالی شکل است. بنابراین توزیع‌هایی با تابع نرخ شکست گودالی شکل می‌توانند برای این داده‌ها مناسب باشند. از میان این توزیع‌ها، توزیع‌های گاما-یکنواخت (GU)، تراپی و منتظری (۲۰۱۲a) توزیع بتا-وایبول اصلاح شده (BMW)، سیلوا و همکاران (۲۰۱۰)، توزیع وایبول اصلاح شده تعمیم یافته (GMW)، کاراسکو و همکاران (۲۰۰۸)، توزیع بتا-وایبول (BW) لی و همکاران (۲۰۰۵) و توزیع وایبول اصلاح شده (MW) لی و همکاران (۲۰۰۳) برای برآزش دادن به داده‌ها انتخاب شده‌اند. برای مقایسه مدل‌ها از معیار آکاییک (AIC) و آماره کلموگروف-اسمیرنف (K-S) و P-مقدار نظیر آن استفاده شده است.

^۲ Total Time on Test

مقادیر کوچک آماره K-S و بزرگ P-مقدار دلالت بر پذیرفته شدن فرض مناسب بودن مدل دارند.

MLE پارامترها و مقادیر آماره AIC در جدول ۶ آورده شده است. این نتایج اشاره می کند که مدل نرمال-یکنواخت با انعطاف پذیری بیشتری داده ها را نسبت به بقیه مدل ها برازش می دهد.

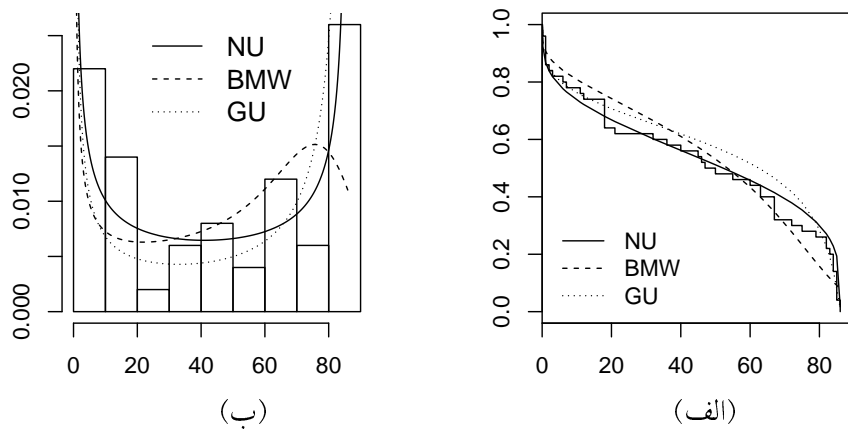
جدول ۶: برآورد ماکسیمم پارامترهای توزیع های NU، GU، BMW، GMW، BW و MW برای داده های آرست و مقایسه اندازه های AIC و آماره K-S آنها

P-value	K-S	AIC	MLE	توزیع
۰/۹۳۲	۰/۰۷۶	۴۰۹/۰۱۲	$\hat{\mu} = ۰/۱۲۸۵, \hat{\sigma} = ۱/۷۱۸۴$ $\hat{a} = ۰/۰۹, \hat{b} = ۸۶/۰۱$	NU
۰/۱۹۸	۰/۱۵۲	۴۱۸/۶۵۰	$\hat{\alpha} = ۰/۲۶۷, \hat{\beta} = ۵۱/۹۴۲$ $\hat{a} = ۰/۰۹, \hat{b} = ۸۶/۷۱۳$	GU
۰/۳۶۴	۰/۱۳۰	۴۵۰/۵۵۶	$\hat{\alpha} = ۰/۰۰۰۰۲۴, \hat{\lambda} = ۰/۰۵۴$ $\hat{\gamma} = ۱/۳۸, \hat{a} = ۰/۱۹۷$ $\hat{b} = ۰/۱۶۵$	BMW
۰/۲۳۴	۰/۱۴۶	۴۵۲/۸۰۷	$\hat{\alpha} = ۰/۰۰۰۰۰۱, \hat{\lambda} = ۰/۰۶۵$ $\hat{\gamma} = ۱/۳۷۳, \hat{a} = ۰/۲۲۳$	GMW
۰/۴۱۹	۰/۱۲۵	۴۵۴/۲۱۹	$\hat{a} = ۰/۱۲۹, \hat{b} = ۰/۰۷$ $\hat{\alpha} = ۰/۰۰۰۰۰۱, \gamma = ۳/۳۲$	BW
۰/۳۳۲	۰/۱۳۴	۴۶۰/۳۱۱	$\hat{a} = ۰/۰۶۲$ $\hat{\lambda} = ۰/۰۲۳, \hat{\gamma} = ۰/۳۵۵$	MW

در شکل ۴ نمودار (ب)، چگالی های برآورد شده توزیع های انتخاب شده با بافت نگار داده ها و نمودار (الف) تابع بقا برآورد شده توزیع های انتخاب شده با تابع بقا تجربی داده ها را نشان می دهد. همان طور که ملاحظه می شود توزیع نرمال-یکنواخت می تواند یک مدل تقریبی خوب برای این چنین داده هایی باشد.

بحث و نتیجه گیری

در این مقاله، خانواده توزیع های تولید شده از همسته نرمال معرفی گردید و یک عضو این خانواده به نام توزیع نرمال-یکنواخت مورد بررسی قرار گرفت. این



شکل ۴: نمودار توابع بقا برآورد شده و تجربی برای داده‌های آارست (الف) و نمودار چگالی‌های برآورد شده و تجربی (ب)

توزیع دارای انعطاف‌پذیری بالایی است و با تغییر پارامترهای توزیع، نمودار تابع چگالی آن می‌تواند متقارن، گودالی شکل، چوله به راست، چوله به چپ، 'J' و وارون و 'J' شکل باشد. همچنین این توزیع به دلیل دارا بودن تابع نرخ شکست صعودی و خم گودالی شکل در زمینه‌های قابلیت اعتماد مورد استفاده قرار می‌گیرد.

تقدیر و تشکر

نویسندگان مقاله بر خویش بایسته می‌دانند از داوران محترم که با پیشنهادهای سازنده خود باعث بهبودی و روانی مقاله شدند، سپاسگزاری نمایند.

مراجع

- Aarset, M. V. (1987), How to Identify Bathtub Hazard Rate, *IEEE Transactions on Reliability*, **36**, 106-108.

- Alexander, C., Cordeiro, G. M., Ortega, E. M. M. and Sarabia, J. M. (2012), Generalized Beta-Generated Distributions, *Computational Statistics and Data Analysis*, **56**, 1880-1897.
- Alzaatreh, A., Famoye, F. and Lee, C. (2013a), Weibull-Pareto Distribution and Its Applications, *Communications in Statistics-Theory and Methods*, **42**, 1673-1691.
- Alzaatreh, A., Lee, C. and Famoye, F. (2013). A New Method for Generating Families of Continuous Distributions, *METRION*, **71**, 63-79.
- Amini, M., MirMostafaei, S. M. T. K. and Ahmadi, J. (2012), Log-Gamma Generated Families of Distributions, *Statistics: A Journal of Theoretical and Applied Statistics*, 1-20.
- Carrasco, J. M. F., Ortega, E. M. M. and Cordeiro, G. M. (2008), A Generalized Modified Weibull Distribution for Lifetime Modeling, *Computational Statistics and Data Analysis*, **53**, 450-462.
- Cordeiro, G. M. and Castro, M. d. (2011), A New Family of Generalized Distributions, *Journal of Statistical Computation and Simulation*, **81**, 883-898.
- Eugene, N., Lee, C. and Famoye, F. (2002), Beta-Normal Distribution and Its Applications, *Communication in Statistics-Theory and Methods*, **31**, 497-512.
- Jones, M. C. (2004), Families of Distributions Arising From Distributions of Order Statistics, *TEST*, **13**, 1-43.
- Lai, C. D., Xie, M. and Murthy, D. N. P. (2003), A Modified Weibull Distribution, *Transactions on Reliability*, **52**, 33-37.

- Lee, C., Famoye, F. and Olumolade, O. (2005), The Beta-Weibull Distribution, *Journal of Statistical Theory Application*, **4**, 121-136.
- Ristic, M. M. and Balakrishnan, N. (2011), The Gamma-Exponentiated Exponential Distribution, *Journal of Statistical Computation and Simulation*, **82**, 1191-1206.
- Silva, G. O., Ortega, E. M. M. and Cordeiro, G. M. (2010), The Beta Modified Weibull Distribution, *Lifetime Data Analysis*, **16**, 409-430.
- Torabi, H. (2008), A General Method for Estimating and Hypotheses Testing Using Spacing, *Journal of Statistical Theory and Applications*, **8**, 163-168.
- Torabi, H. and Montazeri, H. N. (2012a), The Gamma-Uniform Distribution and Its Applications, *Kybernetika*, **48**, 16-30.
- Torabi, H. and Montazeri, H. N. (2012b), The Logistic-Uniform Distribution and Its Applications, *Communications in Statistics - Simulation and Computation*, DOI: 10.1080/03610918.2012.737491.
- Zografos, K. and Balakrishnan, N. (2009), On Families of Beta-and Generalized Gamma-Generated Distributions and Associated Inference, *Statistical Methodology*, **6**, 344-362.