

## برآورد بیزی پارامترهای توزیع چوله‌نرمال

محمد غلامی فشارکی<sup>۱</sup>، انوشیروان کاظم‌نژاد<sup>۱</sup>، فرید زایری<sup>۲</sup>

<sup>۱</sup>گروه آمار زیستی، دانشگاه تربیت مدرس

<sup>۲</sup>گروه آمار زیستی، دانشگاه علوم پزشکی شهید بهشتی

تاریخ دریافت: ۱۳۹۱/۱/۱۶ تاریخ آخرین بازنگری: ۱۳۹۱/۵/۲۶

**چکیده:** توزیع چوله‌نرمال، یکی از توزیع‌های مهم در تحلیل داده‌های غیرنرمال است. از آنجایی که تابع چگالی توزیع چوله‌نرمال حاوی تابع انگرال است محاسبه تابع چگالی این توزیع در رهیافت بیزی است در این مقاله با استفاده از تعریف شرطی توزیع چوله‌نرمال روشی برای برآورد بیزی پارامترهای این توزیع ارائه شده است. سپس در مطالعه‌ای شبیه‌سازی دقت این روش با روش معمولی مورد مقایسه قرار گرفته است.

**واژه‌های کلیدی:** رهیافت بیزی، روش شرطی، توزیع چوله‌نرمال.

### ۱ مقدمه

یکی از فرض‌های اساسی در بسیاری از مدل‌سازی‌های آماری، نرمال بودن توزیع مشاهدات است. متاسفانه این فرض در بسیاری از موارد قابل بررسی نبوده یا غیر واقع گرایانه است. به همین دلیل آماردانان به ناچار برای تحلیل داده‌ها از روش‌هایی مانند تبدیل داده‌ها یا روش‌های ناپارامتری بهره می‌جوینند (سندیوپدیهایی و همکاران، ۲۰۱۰). استفاده از روش‌های

آدرس الکترونیک مسئول مقاله: انوشیروان کاظم‌نژاد, Kazem\_an@modares.ac.ir  
کد موضوع بنای ریاضی (۲۰۰۰): ۶۲F۱۵, ۶۰E۰۵

نایپارامتری علی‌رغم کاهش اطلاعات از لحاظ اجرایی دارای محدودیت‌های زیادی هستند از طرف دیگر روش تبدیل متغیرها نیز، منجر به مشکلاتی چون تفسیر نامناسب پارامترهای تحت بررسی، احراز نشدن فرض نرمال توان متغیرها در حالات چندمتغیره (تسونگ و لی، ۲۰۰۳؛ لی و همکاران، ۲۰۰۵؛ جارا و همکاران، ۲۰۰۸)، پنهان شدن فرایند تولید داده‌ها (جارا و همکاران، ۲۰۰۸) و یکتا نبودن تبدیل در مطالعات مختلف می‌گردد. روش دیگر برای تحلیل داده‌های غیرنرمال استفاده از توزیع‌هایی است که به لحاظ نظری قادر به تبیین تغییرات مشاهدات بدون استفاده از فرض نرمال هستند. یکی از توزیع‌های معروف در این زمینه، توزیع چوله‌نرمال است. اما علی‌رغم مزایای این توزیع، به خاطر وجودتابع انتگرال در تابع چگالی این توزیع استفاده از آن در رهیافت بیزی با مشکل مواجه می‌شود. از این رو برخی از محققان برای تعریف توزیع چوله‌نرمال از یک روش غیرمستقیم و ترکیب خطی دو توزیع نرمال و نیمنرمال به این هدف نائل می‌آیند. در این مقاله پس از بررسی خصوصیات توزیع در بخش ۳، در بخش دو و بیان رهیافت بیزی توزیع نرمال چوله با استفاده از نرم‌افزار WinBUGS پرداخته و سپس در بخش پنجم با انجام شبیه‌سازی به بررسی دقیق و صحیح هر یک از دو روش ارائه شده پرداخته و نهایتاً به بحث و نتیجه‌گیری ارائه شده است.

## ۲ توزیع چوله‌نرمال

آزالینی (۱۹۸۵) توزیع چوله‌نرمال<sup>۱</sup> (SN) با پارامتر چولگی  $\lambda$  را به صورت

$$f_\lambda(x) = \frac{1}{\sqrt{2\pi}} \phi(x) \Phi(\lambda x) \quad \lambda \in \mathbb{R} \text{ & } x \in \mathbb{R} \quad (1)$$

معرفی کرد، که در آن  $(\cdot)\phi$  تابع چگالی و  $(\cdot)\Phi$  تابع توزیع تجمعی نرمال استاندارد هستند. در این توزیع با در نظر گرفتن  $\lambda$  برابر با صفر، توزیع SN به توزیع نرمال تبدیل می‌شود، همچنین با در نظر گرفتن  $\lambda$  برابر با  $\pm\infty$ ، توزیع SN به توزیع نیمنرمال<sup>۲</sup> تبدیل می‌شود (اسپینوسا ۲۰۰۸). آزالینی نشان داده پارامتر چولگی  $\lambda$  در بازه  $(-0.995, 0.995)$  تغییر می‌کند (نادریج و کوتز، ۲۰۰۳). یعنی از توزیع SN در موقعی می‌توان استفاده کرد که چولگی داده‌ها زیاد نباشد.

---

<sup>۱</sup> Skew Normal

<sup>۲</sup> Half Normal

اگر  $X$  دارای توزیع چوله نرمال با پارامتر  $\lambda$  باشد، آنگاه میانگین، واریانس، ضریب چولگی و کشیدگی و تابع مولد گشتاور این توزیع به ترتیب عبارتند از (آزالینی، ۱۹۸۵)

$$\begin{aligned} E(X) &= \sqrt{\frac{2}{\pi}}\delta \quad \delta = \lambda/\sqrt{1+\lambda^2} \in [-1, +1] \\ Var(X) &= 1 - \frac{2}{\pi}\delta^2 \\ Skewness &= \frac{1}{\sqrt{2}}(\frac{4}{\pi} - 1) Sign(\lambda) \left( \frac{\lambda^2}{\frac{\pi}{4} + (\frac{\pi}{4} - 1)\lambda^2} \right)^{\frac{1}{2}} \\ Kurtosis &= 2(\pi - 2) Sign(\lambda) \left( \frac{\lambda^2}{\frac{\pi}{4} + (\frac{\pi}{4} - 1)\lambda^2} \right)^2 \\ m_X(t) &= 2e^{t^2} 2\Phi(\delta \times t) \end{aligned}$$

برخی از خصوصیات دیگر این توزیع که مشابه توزیع نرمال هستند عبارتند از (گنتون، ۲۰۰۴)

$$-X \sim SN(-\lambda), \quad |X| \sim 2\phi(x)I(x > 0), \quad X^2 \sim \chi^2(1)$$

در حالت کلی  $Y$  دارای توزیع چوله نرمال با پارامتر مکانی  $\mu$ ، پارامتر پراکندگی  $\sigma^2$  و پارامتر چولگی  $\lambda$  است، هرگاه تابع چگالی آن به صورت

$$f(y, \mu, \sigma, \lambda) = \frac{1}{\sigma} \phi\left(\frac{y-\mu}{\sigma}\right) \Phi\left(\lambda \frac{y-\mu}{\sigma}\right)$$

باشد (گنتون، ۲۰۰۴). از طرف دیگر در صورتی که  $X$  دارای توزیع چوله نرمال استاندارد با پارامتر چولگی  $\lambda$  باشد، هر توزیع چوله نرمال را می‌توان از توزیع چوله نرمال استاندارد و با استفاده از ترکیب خطی  $\sigma X + \mu$  به دست آورد. خصوصیت مهم دیگری که آزالینی (۲۰۰۵) ثابت کرد این است که اگر  $X_1$  و  $X_2$  مستقل و دارای توزیع نرمال استاندارد باشند، آنگاه توزیع، متغیر تصادفی

$$Y = \lambda |X_1| + X_2 \tag{۲}$$

دارای توزیع چوله نرمال استاندارد با پارامتر  $\lambda$  است.

### ۳ روش انتخاب مدل و تحلیل بیزی

آماردانان در تحلیل داده‌ها از دو رهیافت کلاسیک و بیزی بهره می‌جویند. در رهیافت کلاسیک، پارامتر مورد بررسی ثابت و نامعلوم بوده و بر اساس نمونه‌ای تصادفی در مورد آن

تصمیم‌گیری می‌شود. این در حالی است که در رهیافت بیزی، پارامتر مورد بررسی تحقیقی از یک متغیر تصادفی است که تغییرات آن توسط یک توزیع احتمال، تحت عنوان توزیع پیشین، توصیف می‌شود. این توزیع بر اساس تجربیات قبلی آزمایشگر تعیین شده و بر اساس نمونه تصادفی به روز می‌شود. اصطلاحاً توزیع پیشین اصلاح شده بر اساس مشاهدات جدید را توزیع پسین نامیده و تصمیم‌گیری در مورد پارامترهای مدل بر اساس آن انجام می‌شود (کونگ دان، ۲۰۰۳، بولستند، ۲۰۰۴؛ گلدستین و وف، ۲۰۰۷) فرض کنید  $y = (y_1, \dots, y_n)$  بردار مشاهدات نمونه‌ای تصادفی از توزیع  $f(y|\theta)$  باشد. توزیع پسین را می‌توان به صورت

$$\pi(\theta|\lambda) = \frac{f(y|\theta)\pi(\theta)}{m(y)} \propto f(y|\theta)\pi(\theta)$$

نوشت، که در آن  $m(y) = \int f(y, \theta) d\theta$  توزیع کناری  $y$  است. در رهیافت بیزی معمولاً توزیع پیشین از یک خانواده انعطاف‌پذیر که دارای یک یا چند پارامتر است، انتخاب می‌شود. یعنی توزیع پیشین  $\theta$  به صورت  $\pi(\theta|\gamma)$  است، که در آن  $\gamma$  پارامتری با مقدار حقیقی یا برداری است و می‌تواند معلوم یا نامعلوم باشد. در حالتی که نامعلوم باشد می‌تواند ثابت و یا متغیری تصادفی یا یک توزیع خاص باشد. بنابراین در بحث آمار بیزی سه روش بیز با تک پیشین<sup>۳</sup> (در این حالت مقدار  $\gamma$  معلوم است)، بیز سلسله مراتبی<sup>۴</sup> (در این حالت  $\gamma$  نامعلوم و متغیری تصادفی است که دارای یک توزیع خاص می‌باشد) و بیز تجربی<sup>۵</sup> (در این حالت  $\gamma$  نامعلوم و ثابت که توسط نمونه و فرم توزیع پیشین برآورد می‌گردد) مطرح می‌شود. تحلیل‌های بیزی دارای مزایایی چون استفاده از تجربیات و اعتقادات پیشین در مورد پارامتر مورد مطالعه، محاسبه توزیع برآورد به جای برآورد نقطه‌ای (کونگدون، ۲۰۰۳؛ بولستند، ۲۰۰۴) حساس نبودن ضرایب به حجم نمونه (کاس رفتی، ۱۹۹۵)، وجود خصوصیت اصل درستنمایی<sup>۶</sup> (ساویله، ۲۰۰۸). تعابیر احتمالی درست برای  $p$ -value، فواصل اطمینان و استفاده از نقش توابع زیان نسبت به روش کلاسیک هستند (رفتی، ۱۹۸۶).

#### ۴ توزیع‌های پیشین و توام پسین

**برآورد پارامتر مستقیم:** چون کار با تابع توزیع به خاطر وجود انتگرال در آن دشوار است، در تعیین چگالی پسین شرطی، به ناچار تخمین تابع توزیع نرمال استفاده می‌شود. علی‌رغم

<sup>۳</sup> Single Prior Bayes

<sup>۴</sup> Hierarchical Bayes

<sup>۵</sup> Empirical Bayes

<sup>۶</sup> Likelihood Principle

تنوع روش‌های متعدد در تخمین تابع تجمعی نرمال (بولینگ و همکاران ۲۰۰۹)، که در همه آنها تابع توزیع نرمال به صورت تقریبی  $\Phi(x) \approx k_0 - k_1 \exp(f(k_2, \dots, k_p))$  در نظر گرفته می‌شود، به طوری که  $k_0, \dots, k_1, k_2, \dots, k_p$  پارامترهای تعیین‌کننده تابع توزیع نرمال هستند، عملاً استفاده از این تقریب‌ها برای محاسبه چگالی پسین شرطی میسر نیست. از این رو با استفاده از رگرسیون کمترین توانهای دوم و قرار دادن  $\ln(\Phi(x))$  به عنوان متغیر پاسخ و استفاده از اسپلاین بهترین تخمین‌گر برای تابع توزیع نرمال محاسبه می‌شود. سپس با محاسبه  $\sum_{i=1}^n |\Phi(x_k) - \Phi(\hat{x}_k)| / n$  میزان خطای مدل برابر با ۰.۰۸۴ حاصل گردید.

$$\Phi(x) \approx \exp(-0/727z + 0/813zx - 0/243zx^2) \quad (3)$$

$$z = \begin{cases} 1 & x \leq 2 \\ 0 & x > 2 \end{cases}$$

با در نظر گرفتن  $(\mu, \sigma, \lambda) = \theta$  و انتخاب توزیع‌های پیشین به صورت

$$\mu \sim N(\mu_\mu, \sigma_\mu^2), \quad \lambda \sim N(\mu_\lambda, \sigma_\lambda^2), \quad \sigma^2 \sim IGamma(\alpha_\sigma, \beta_\sigma)$$

می‌توان تابع چگالی پسین را به صورت

$$\begin{aligned} \pi(\theta|y) &\propto f(y|\theta)\pi(\theta) \\ &\propto \prod_{i=1}^n \frac{1}{\sqrt{\pi}\sigma} \exp\left(-\frac{1}{\sigma^2}(y_i - \mu)^2\right) \Phi\left(\lambda \frac{(y_i - \mu)}{\sigma}\right) \\ &\times \frac{1}{\sqrt{\pi}\sigma_\mu} \exp\left(-\frac{1}{\sigma_\mu^2}(\mu - \mu_\mu)^2\right) \frac{1}{\sqrt{\pi}\sigma_\lambda} \exp\left(-\frac{1}{\sigma_\lambda^2}(\lambda - \mu_\lambda)^2\right) \\ &\times \frac{\beta_\sigma^{\alpha_\sigma} \left(\frac{1}{\sigma^2}\right)^{\alpha_\sigma+1} \times \exp\left(-\frac{\beta_\sigma}{\sigma^2}\right)}{\Gamma(\alpha_\sigma)} \end{aligned} \quad (4)$$

به دست آورد. حال با جایگذاری (۳) در (۴) داریم

$$\begin{aligned} \pi(\theta|y) &\propto \left(\frac{1}{\sqrt{n}}\right)^{n/2} \frac{1}{\sqrt{\pi}\sigma_\mu} \frac{1}{\sqrt{\pi}\sigma_\lambda} \frac{\beta_\sigma^{\alpha_\sigma}}{\Gamma(\alpha_\sigma)} \exp\left(-0/727 \sum_{i=1}^n z_i\left(\frac{1}{\sigma^2}\right)^{\alpha_\sigma+\frac{n}{2}+1}\right) \\ &\times \exp\left\{-\frac{(-0/5 \sum_{i=1}^n (y_i - \mu)^2 + 0/243 \sum_{i=1}^n z_i (y_i - \mu)^2 \lambda^2)}{\sigma^2}\right\} \\ &+ \frac{(0/813 \lambda \sum_{i=1}^n z_i (y_i - \mu))}{\sigma} - \frac{1}{\sigma_\mu^2} (\mu - \mu_\mu)^2 \\ &- \frac{1}{\sigma_\lambda^2} (\lambda - \mu_\lambda)^2 - \frac{\beta_\sigma}{\sigma^2} \end{aligned} \quad (5)$$

## برآوردهای بیزی پارامترهای توزیع چوله‌نرمال ۸۸

چون مقدار  $(\frac{1}{\sqrt{n}})^{n/2} \frac{1}{2\pi\sigma_\mu\sigma_\lambda} \frac{\beta_\sigma^\alpha\sigma}{\Gamma(\alpha\sigma)}$  به  $\theta$  بستگی ندارد با حذف آن از رابطه (۵) داریم

$$\begin{aligned} \pi(\theta|y) &\propto (\frac{1}{\sigma})^{\alpha\sigma + \frac{n}{2} + 1} \\ &\times \exp\left\{-\frac{(-\circ/\delta \sum_{i=1}^n (y_i - \mu)^2 + \circ/243 \sum_{i=1}^n z_i (y_i - \mu)^2 \lambda^2)}{\sigma^2}\right\} \\ &+ \frac{(\circ/813\lambda \sum_{i=1}^n z_i (y_i - \mu))}{\sigma} - \frac{1}{2\sigma_\mu^2}(\mu - \mu_\mu)^2 \\ &- \frac{1}{2\sigma_\lambda^2}(\lambda - \mu_\lambda)^2 - \frac{\beta_\sigma}{\sigma} \end{aligned} \quad (6)$$

با خلاصه‌سازی (۶) توزیع‌های شرطی کامل به صورت زیر حاصل می‌شوند

$$\begin{aligned} (\lambda|y, \sigma^2, \mu) &\sim N\left(\frac{\frac{\mu_\lambda}{\sigma_\lambda^2} + \circ/813 \frac{\sum_{i=1}^n z_i (y_i - \mu)}{\sigma}}{\circ/486 \frac{\sum_{i=1}^n z_i (y_i - \mu)^2}{\sigma^2} + \frac{1}{\sigma_\lambda^2}}, \circ/486 \frac{\sum_{i=1}^n z_i (y_i - \mu)^2}{\sigma^2} + \frac{1}{\sigma_\lambda^2}\right) \quad (7) \\ (\mu|y, \sigma^2, \lambda) &\sim N\left(\frac{\frac{\mu_\mu}{\sigma_\mu^2} - \circ/813 \frac{\lambda \sum_{i=1}^n z_i}{\sigma} - (1 + \circ/486\lambda^2) \frac{\sum_{i=1}^n y_i z_i}{\sigma^2}}{\frac{1 + \circ/486\lambda^2}{\sigma^2} \sum_{i=1}^n z_i + \frac{1}{\sigma_\mu^2}}, \frac{1 + \circ/486\lambda^2}{\sigma^2} \sum_{i=1}^n z_i + \frac{1}{\sigma_\mu^2}\right) \\ \left(\frac{1}{\sigma}|y, \lambda, \mu\right) &\sim Nad\left(2\alpha_\sigma + n, \circ/5 \sum_{i=1}^n (y_i - \mu)^2 + \circ/243 \sum_{i=1}^n z_i (y_i - \mu)^2 \lambda^2 + \beta_\sigma, \circ/813\lambda \sum_{i=1}^n z_i (y_i - \mu)\right) \quad (8) \end{aligned}$$

که در آن Nad(a,b,c) بیانگر توزیع نادر اجاه (نادر اجاه، ۲۰۰۸) با تابع چگالی

$$f(x; \alpha, a, b) = C x^{\alpha-1} \exp(-ax^2 - bx) \quad x > 0, a > 0, -\infty < b < +\infty$$

است، به طوری که  $C = \int x^{\alpha-1} \exp(-ax^2 - bx) dx$  ضریب نرمال‌ساز است. با استفاده از MCMC می‌توان برآوردهای بیزی پارامترهای  $(\mu, \sigma^2, \lambda)$  را محاسبه کرد.

**برآوردهای پارامتر غیرمستقیم:** در روش غیرمستقیم و با استفاده از

(۲) می‌توان تابع چگالی چوله‌نرمال را به فرم دسته معادلات

$$\begin{aligned} f_\theta(y|t) &\sim N(y - \mu - \lambda|t|, \sigma^2) \\ f(t) &\sim N(0, \sigma^2) \\ f_\theta(y, t) &= f_\theta(y|t) \times f(t) \end{aligned} \quad (4)$$

نوشت. همچنین با در نظر گرفتن  $(\mu, \sigma^2, \lambda) = \theta$  می‌توان توزیع پسین را به صورت

$$\begin{aligned} \pi(\theta|y, t) \propto & \left( \frac{1}{\sigma^2} \right)^{\alpha_\sigma + n + 1} \exp\left(-\frac{\sum_{i=1}^n (y_i - \mu - \lambda|t_i|)^2}{2\sigma^2}\right) \\ & - \frac{\sum_{i=1}^n |t_i|}{2\sigma^2} - \frac{(\mu - \mu_\mu)^2}{2\sigma_\mu^2} - \frac{(\lambda - \mu_\lambda)^2}{2\sigma_\lambda^2} - \frac{\beta_\sigma}{\sigma^2} \end{aligned} \quad (10)$$

به دست آورد. حال توزیع‌های شرطی کامل پارامترها عبارتند از

$$\begin{aligned} (\lambda|y, \sigma^2, \mu, t) &\sim N\left(\frac{\frac{\mu_\lambda}{\sigma_\lambda^2} + \frac{\sum_{i=1}^n |t_i|(y_i - \mu)}{\sigma^2}}{\frac{\sum_{i=1}^n |t_i|}{\sigma^2} + \frac{1}{\sigma_\lambda^2}}, \frac{\sum_{i=1}^n |t_i|}{\sigma^2} + \frac{1}{\sigma_\lambda^2}\right) \\ (\mu|y, \sigma^2, \lambda, t) &\sim N\left(\frac{\frac{\mu_\mu}{\sigma_\mu^2} + \frac{(n\bar{y} - \lambda \sum_{i=1}^n |t_i|)}{\sigma^2}}{\frac{n}{\sigma^2} + \frac{1}{\sigma_\mu^2}}, \frac{n}{\sigma^2} + \frac{1}{\sigma_\mu^2}\right) \\ \left(\frac{1}{\sigma^2}|y, \lambda, \mu, t\right) &\sim Gamma(\alpha_\sigma + n + 0/5 \\ &, 0/5 \sum_{i=1}^n (y_i - \mu - \lambda|t_i|)^2 + \sum_{i=1}^n |t_i|^2 + \beta_\sigma) \end{aligned}$$

## ۵ مطالعه شبیه‌سازی

با استفاده از بسته sn در نرم افزار R از توزیع چوله‌نرمال در چهار سطح چولگی ۰، ۱، ۲، ۳ و ۴ حجم نمونه ۱۰۰، ۲۰۰ و ۵۰۰ دو میانگین ۳ و ۳- و واریانس ۲ شبیه‌سازی انجام گرفته است. سپس با استفاده از بسته‌های sn و R2WinBUGS در نرم افزار R و نرم افزار WinBUGS برآوردهای ماکسیمم درستنمایی و بیزی به روش‌های مستقیم و غیرمستقیم محاسبه می‌شوند. توزیع پیشین برای  $\mu$  و  $\lambda$   $N(0, 100)$  و برای  $\sigma^2$   $IGamma(0/1, 0/1)$  در نظر گرفته شد. پس از ۱۰۰ بار تکرار شبیه‌سازی نتایج در جداول ۱ تا ۴ ارائه شده‌اند، که در آن‌ها حروف D، I و M به ترتیب نشان‌دهنده روش‌های مستقیم، غیرمستقیم و ماکسیمم درستنمایی هستند. برای مقایسه دقیق برآورد پارامترهای مدل از معیار متوسط قدر مطلق خطا به صورت  $MAE(\hat{\theta}_k) = \frac{\sum_{i=1}^n |\hat{\theta}_{ik} - \theta_k|}{n}$  استفاده می‌شود، که در آن  $\hat{\theta}_k$  پارامتر مورد نظر و  $\theta_k$

برآوردهای پارامترهای توزیع چوله‌نرمال ..... ۹۰

جدول ۱: برآوردهای پارامترها و میزان خطای آن‌ها با سه روش به ازای  $\mu = 3$

$\lambda$						$n$
-۲			-۵			
M	I	D	M	I	D	
۲/۶۴	۲/۰۷	۱/۹۹	۲/۸۸	۲/۳۳	۲/۱۴	برآوردهای خطای درصد
۰/۷۵	۰/۹۹	۱/۰۳	۰/۲۶	۰/۶۸	۰/۸۷	$\mu$
۴۹	۱۱	۷	۴۴	۶	۶	
-۱۰۰/۸	-۰/۹۹	-۰/۸۸	-۳۵۶/۹	-۲/۰۸	-۲۶/۳	برآوردهای خطای درصد
۱۸۵/۸۸	۱/۴۶	۱/۰۲	۳۵۳/۸۷	۲/۹۹	۲۶/۷۷	$\lambda$
۳۹	۷۷	۷۷۱	۴۰	۱۰۰	۹۲	۳۰
۱/۹۷	۱	۱/۸	۱/۸۷	۰/۷۵	۸/۳۶	برآوردهای خطای درصد
۰/۳۴	۱	۰/۳۲	۰/۲۹	۱/۲۵	۷/۱۳	$\sigma^2$
۳۹	۷۷	۷۷	۴۰	۱۰۰	۹۲	
۲/۸۶	۲/۴۹	۲/۰۴	۳	۲/۸۶	۲/۹۱	برآوردهای خطای درصد
۰/۱۲۹	۰/۰۵۵	۰/۰۱	۰/۱۱۳	۰/۱۸	۰/۱۵	$\mu$
۵۱	۴۱	۱۸	۵۶	۲۱۱	۷۷	
-۱/۹۶	-۱/۴۷	-۱/۰۷	-۳۷/۳۵	-۳/۸۵	-۴/۴۶	برآوردهای خطای درصد
۰/۶۶	۰/۰۵	۰/۸۳	۳۳/۱۸	۱/۲۲	۰/۹۵	$\lambda$
۴۶	۷۰	۶۹	۳۹	۹۰	۶۲	۱۰۰
۱/۹۳	۰/۹۹	۱/۸۷	۲/۰۱	۰/۰۲	۱/۹۳	برآوردهای خطای درصد
۰/۲۱	۱/۰۱	۰/۲۳	۰/۱۶	۱/۹۸	۰/۱۷	$\sigma^2$
۴۶	۷۰	۶۹	۳۹	۹۰	۶۲	
۲/۹۶	۲/۷۳	۲/۷۹	۲/۹۹	۲/۹۱	۲/۹۵	برآوردهای خطای درصد
۰/۲۰	۰/۳۷	۰/۳۴	۰/۰۸	۰/۱	۰/۰۹	$\mu$
۴۹	۷۷	۴۰	۴۴	۱۷	۳۱	
-۲/۰۹	-۱/۷۸	-۱/۸۹	-۵/۵۱	-۴/۱۹	-۴/۷۷	برآوردهای خطای درصد
۰/۰۲	۰/۷۸	۰/۶۷	۱/۲۵	۰/۹۱	۰/۷۳	$\lambda$
۵۰	۵۸	۵۶	۴۳	۸۷	۵۸	۲۰۰
۱/۹۸	۰/۹۶	۱/۹۳	۲	۰/۴۷	۱/۹۷	برآوردهای خطای درصد
۰/۱۳	۱/۰۴	۰/۱۷	۰/۰۹	۱/۰۳	۰/۰۹	$\sigma^2$
۵۰	۵۸	۵۶	۴۳	۸۷	۵۸	
۲/۹۶	۲/۹۲	۲/۹۶	۲/۹۹	۲/۹۵	۲/۹۹	برآوردهای خطای درصد
۰/۲۰	۰/۱۵	۰/۱۴	۰/۰۸	۰/۰۶	۰/۰۵	$\mu$
۴۹۱	۷۷	۴۶	۴۴	۲۵	۴۵	
-۲/۰۹	-۱/۸۹	-۲/۰۱	-۵/۵۱	-۴/۲۸	-۵/۰۱	برآوردهای خطای درصد
۰/۳۲	۰/۳۲	۱/۲۵	۲	۰/۷۷	۰/۶۰	$\lambda$
۵۰	۶۱	۴۴	۴۳	۹۳	۵۵	۵۰۰
۱/۹۸	۰/۹۴	۱/۹۹	۲	۰/۴۵	۱/۹۹	برآوردهای خطای درصد
۰/۱۳	۱/۰۶	۰/۱۱	۰/۰۹	۱/۰۵	۰/۰۷	$\sigma^2$
۵۰	۶۱	۴۴	۴۳	۹۳	۵۵	

۹۱ ..... محمد غلامی فشارکی و همکاران

جدول ۲: برآوردهای پارامترها و میزان خطای آن‌ها با سه روش به ازای  $\mu = ۳$

$\lambda$						$n$	
۵			۲				
M	I	D	M	I	D		
۴/۱۲	۳/۳۷	۲/۱۴	۳/۴۵	۳/۴۹	۳/۳۴	$\mu$	
۰/۲۷	۰/۳۹	۰/۲۹	۰/۷۹	۰/۵۶	۰/۵۴		
۵۵	۸۹۱	۶۵	۵۶	۷۹	۶۹		
۴۱۷/۸۷	۲/۷۷	۲/۴۷	۱۰۵/۶	۱/۶۸	۲/۰۶	$\lambda$	
۴۱۴/۱۳	۲/۲۳	۱/۵۴	۲۱۶/۲۶	۰/۹۹	۱/۰۵		
۵۸	۰	۲	۵۵	۴۲	۵۷		
۱/۸۹	۰/۶۸۷	۱/۹۸	۱/۹۶	۰/۹۵	۲/۰۲	$\sigma^2$	
۰/۲۵	۱۱۳۲	۰/۳۴	۰/۳۳	۱/۰۵	۰/۳۰		
۵۸	۰	۲	۵۵	۴۲	۵۷		
۲/۹۹	۳/۰۸	۳/۰۶	۳/۲۳	۳/۲۵	۳/۳۲	$\mu$	
۰/۱۲	۰/۱۲	۰/۱۱	۰/۳۹	۰/۳۹	۰/۴۳		
۴۵	۱۷۹	۶۸	۶۰	۶۷	۷۰		
۲۳/۰۲	۴/۰۹	۴/۰۸	۱/۹۲	۱/۹۳	۱/۹۶	$\lambda$	
۱۱/۰۹	۱/۱۳	۰/۸۹	۰/۹۲	۰/۸۲	۰/۸۶		
۶۴	۱۸	۳۶	۴۵	۴۶	۴۲		
۲/۰۱	۰/۰	۱/۹۰	۱/۹۳	۰/۹۳	۱/۹۳	$\sigma^2$	
۰/۱۶۱	۱/۰۰۱	۰/۱۰	۰/۲۴	۱/۰۷	۰/۲۳		
۶۴	۱۸	۳۶	۴۵	۴۶	۴۲		
۲/۹۹	۳/۰۸	۳/۰۶	۳/۲۳	۳/۱۴	۳/۱۴	$\mu$	
۰/۱۲	۰/۱۰	۰/۰۸	۰/۳۹	۳/۲۷	۳/۲۸		
۴۵	۸۰	۷۳	۶۰	۵۸	۶۱		
۲۲/۲	۴/۱	۴/۷	۱/۹۲۱	۱/۷	۲	$\lambda$	
۱۹/۰۹	۱/۰۲	۰/۸۴	۰/۹۲	۰/۵۸	۰/۶۱		
۶۴	۱۴	۳۲	۴۵	۵۱	۵۴		
۲/۰۱	۰/۴۸	۱/۹۶	۱/۹۳	۰/۹۸	۱/۹۶	$\sigma^2$	
۰/۱۶	۱/۰۲	۰/۱۰	۰/۲۴	۰/۰۷	۰/۱۸		
۶۴	۱۴	۳۲	۴۵	۵۱	۵۴		
۲/۹۹	۳/۰۳	۳/۰۲	۳/۲۳	۲/۹۸	۳	$\mu$	
۰/۱۲	۰/۰۵	۰/۰۵	۰/۳۹	۰/۱۱	۰/۱۰		
۴۵۱	۸۷	۵۸	۶۰	۴۲	۵۲		
۲۳/۰۲	۴/۷۷	۰/۰۱	۱/۹۲۱	۲/۱۹	۲/۱۳	$\lambda$	
۱۹/۰۹	۰/۰۴	۰/۶۰	۰/۹۲	۰/۳۳	۰/۳۱		
۶۴	۳۴	۴۳	۴۵	۷۱	۶۴		
۲/۰۱	۰/۴۲	۱/۹۹	۱/۹۳	۰/۸۶	۲/۰۲	$\sigma^2$	
۰/۱۶	۱/۰۸	۰/۰۶	۰/۲۴	۱/۱۴	۰/۰۹		
۶۴	۳۴	۴۳	۴۵	۷۱	۶۴		

جدول ۳: برآوردهای پارامترها و میزان خطای آنها با سه روش به ازای  $\mu = -3$

$\lambda$						$n$
-۲			-۵			
M	I	D	M	I	D	
-۳ ۰/۱۵ ۵۱	-۳/۰۸ ۰/۲۰ ۴۵	-۳/۰۶ ۰/۲۰ ۴۵	-۳/۰۱ ۰/۰۸ ۴۷	-۳/۰۷ ۰/۱۰ ۲۲	-۳/۰۴ ۰/۰۸ ۳۱	برآورد خطا درصد
-۲/۱۴ ۰/۱۵ ۴۶	-۲/۰۶ ۰/۴۹ ۴۵	-۲/۱۲ ۰/۰۲ ۴۴۱	-۳/۸۳ ۱۶/۵۹ ۴۸	-۴/۲۶ ۰/۹۰ ۸۲	-۴/۸۲ ۰/۸۰ ۶۱	برآوردهای خطا درصد
۲ ۰/۱۳ ۴۶	۰/۹۰ ۱/۱۰ ۴۵	۱/۹۹ ۰/۱۴ ۴۴	۱/۹۸ ۰/۱۱ ۴۸	۰/۴۷ ۱/۰۳ ۸۲	۱/۹۶ ۰/۱ ۶۱	برآوردهای خطا درصد
-۳/۱۴ ۰/۳۲ ۸۰	-۳/۲۴ ۰/۳۷ ۳۴	-۳/۲۲ ۰/۳۷ ۴۰	-۲/۹۸ ۰/۱۱ ۱۰	-۳/۰۸ ۰/۱۱ ۲۸	-۳/۰۳ ۰/۱۱ ۳۹	برآوردهای خطا درصد
-۲/۳۲ ۱/۱۰ ۴۹	-۱/۹۸ ۰/۸۳ ۵۶	-۲/۰۵ ۰/۹۰ ۵۰	-۳۷/۰۷ ۳۲/۷۸ ۳۵	-۴/۰۸ ۱/۰۵ ۸۷	-۴/۶۸ ۰/۱۵ ۶۳	برآوردهای خطا درصد
۱/۹۰ ۰/۱۴ ۴۹	۰/۹۲ ۱/۰۹ ۵۶	۱/۹۵ ۰/۲۳ ۵۰	۲/۰۴ ۰/۱۵ ۳۵	۰/۰ ۱/۰۰ ۸۷	۲ ۰/۱۵ ۶۳	برآوردهای خطا درصد
-۳/۰۳ ۰/۱۹ ۵۸	-۳/۱۴ ۰/۲۷ ۴۹	-۳/۱۰ ۰/۲۵ ۵۴	-۳/۰۲ ۰/۰۹ ۴۸	-۳/۰۸۱ ۰/۱۰ ۲۳	-۳/۰۵ ۰/۹۰ ۳۲	برآوردهای خطا درصد
-۲/۱۲ ۰/۰۸ ۴۴	-۲ ۰/۶۴ ۴۶	-۴/۰۲ ۰/۶۶ ۴۴	-۵/۴۴ ۱/۴۵ ۴۸	۴/-۱۳ ۰/۹۲ ۸۵	-۴/۶۹ ۰/۸۸ ۶۴	برآوردهای خطا درصد
۰/۹۹ ۰/۱۰ ۴۴	۰/۹۲ ۱/۰۸ ۴۶	۱/۹۹ ۰/۱۵ ۴۴	۱/۹۸ ۰/۱۱ ۴۸	۰/۴۸ ۱/۰۲ ۸۵	۱/۹۵ ۰/۱ ۶۴	برآوردهای خطا درصد
-۳/۰۳ ۰/۱ ۴۹۱	-۳/۰۸ ۰/۱۳ ۳۴	-۳/۰۴ ۰/۱۲ ۴۸	-۳ ۰/۰۵ ۵۴	-۳/۰۴ ۰/۰۶ ۲۲	-۳/۰۱ ۰/۰۵ ۴۹	برآوردهای خطا درصد
-۱/۹۹ ۰/۲۹ ۵۰	-۱/۸۹ ۰/۳۲ ۵۷	-۲ ۰/۳۲ ۴۷	-۵/۰۸ ۰/۶۱ ۴۶	-۴/۲۷ ۰/۷۴ ۹۳	-۴/۹۰ ۰/۵۱ ۵۵	برآوردهای خطا درصد
۱/۹۸ ۰/۰۹ ۵۰	۰/۹۳ ۱/۱۷ ۵۷	۱/۹۸ ۱/۰۹ ۴۷	۲ ۰/۰۶ ۴۶	۰/۴۵ ۱/۰۵ ۹۳	۱/۹۹ ۰/۰۶ ۵۵	برآوردهای خطا درصد

جدول ۴: برآوردهای پارامترها و میزان خطای آنها با سه روش به ازای  $\lambda = \frac{5}{\mu}$

$\lambda$						$n$
M	I	D	M	I	D	
-۲/۹۹	-۲/۹۳	-۲/۹۵	-۲/۹۷	-۲/۷۷	-۲/۷۷	برآورد
۰/۰۸	۰/۰۹	۰/۰۸	۰/۱۹	۰/۳۳	۰/۳۳	خطا
۴۹	۷۶	۶۶	۴۹	۷۰	۶۵	درصد
۵/۴۱	۴/۱۸	۴/۹۹	۲/۰۴	۱/۸۱	۱/۸۴	برآورد
۱/۱۷	۰/۹۴	۰/۷۵	۰/۱۴۷	۰/۶۲	۰/۶۳	خطا
۵۲	۱۵	۳۳	۵۱	۴۴	۴۴	درصد
۱/۹۹	۰/۴۸	۱/۹۶	۱/۹۸	۰/۹۶	۱/۹۳	برآورد
۰/۱۱	۱/۰۲	۰/۱۱	۰/۱۵	۱/۰۴	۰/۱۷	خطا
۵۲	۱۵	۳۳	۵۱	۴۴	۴۴	درصد
-۳/۰۱	-۲/۹۰	-۲/۸۹	-۲/۷۸	-۲/۵۶	-۲/۴۹	برآورد
۰/۱۲	۰/۱۳	۰/۱۶	۰/۴۳	۰/۵۷	۰/۶۱	خطا
۴۷	۷۲	۶۱	۵۳	۶۷	۷۴	درصد
۳۶/۸۹	۴/۱۴	۱۵/۶۲	۱/۳۹۹	۱/۷۰	۱/۵۸	برآورد
۳۲/۷۲	۱/۰۵	۱۲/۰۱	۰/۹۱	۱/۰۴	۱/۰۱	خطا
۵۹	۱۶	۳۰	۵۴	۳۸	۳۴	درصد
۲	۰/۰۰	۳/۴۹	۱/۹۵	۰/۹۵	۱/۸۹	برآورد
۰/۱۵	۱/۰۰	۱/۷۰	۱/۲۰	۰/۰۵	۰/۷۲	خطا
۵۹	۱۶	۳۰	۵۴	۳۸	۳۴	درصد
-۳	-۲/۹۴	-۲/۹۶	-۲/۹۷	-۲/۷۹۱	-۲/۷۸	برآورد
۰/۰۹	۰/۰۹	۰/۰۸	۰/۱۸	۰/۳۱	۰/۳۳	خطا
۴۴	۷۲	۶۲	۵۲	۶۸	۶۶	درصد
۶۵/۶۶	۴/۳۷	۴/۸۵	۲/۰۵	۱/۸۳	۱/۸۴	برآورد
۱/۳	۰/۸۶	۰/۷۴	۰/۴۸	۰/۶۰	۰/۶۴	خطا
۵۶	۲۱	۴۶	۵۲	۴۵	۴۶	درصد
۲	۰/۴۶	۱/۹۷	۱/۹۸	۰/۹۵	۱/۹۳	برآورد
۰/۱	۱/۰۴	۰/۱۰	۰/۱۴	۱/۰۵	۰/۱۶	خطا
۵۶	۲۱	۴۶	۵۲	۴۵	۴۶	درصد
-۳/۰۱	-۲/۹۸	-۲/۹۹	-۳	-۲/۹۹	-۲/۹۸	برآورد
۰/۰۵	۰/۰۵	۰/۰۵	۰/۴۳	۰/۴۴	۰/۴۹	خطا
۴۶	۴۹	۵۵	۴۳	۴۴	۴۹	درصد
۵/۱۸	۴/۷۱	۴/۹۳	۲/۰۶	۲/۰۸	۲/۰۴	برآورد
۰/۷۳	۰/۰۴	۰/۶۰	۰/۲۸	۰/۳۰	۰/۲۹	خطا
۵۳	۳۵	۴۲	۵۷	۶۰	۵۷	درصد
۲	۰/۴۲	۱/۹۹	۲	۰/۱۹	۱/۹۹	برآورد
۱/۰۶	۱۱/۰۸	۰/۰۶	۰/۰۹	۱/۱۱	۰/۱۰	خطا
۵۳	۳۵	۴۲	۵۷	۶۰	۵۷	درصد

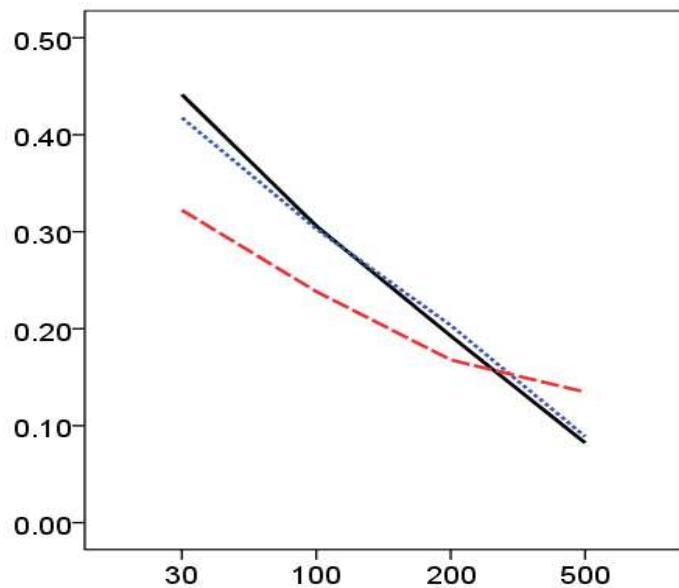
برآورده است. همچنین در جداول ۱ تا ۴، درصد بیش برآورده نیز گزارش شده است. برای مثال در جدول ۱، برای حجم نمونه  $30$  و  $-3 = \lambda$  میزان بیش برآورده میانگین  $6\%$  گزارش شده است. و این یعنی در هر  $100$  مورد،  $6\%$  موارد مقدار برآورده بالاتر از  $3$  و  $94\%$  موارد برآورده کمتر از  $3$  حاصل شده‌اند.

## ۶ بحث و نتیجه‌گیری

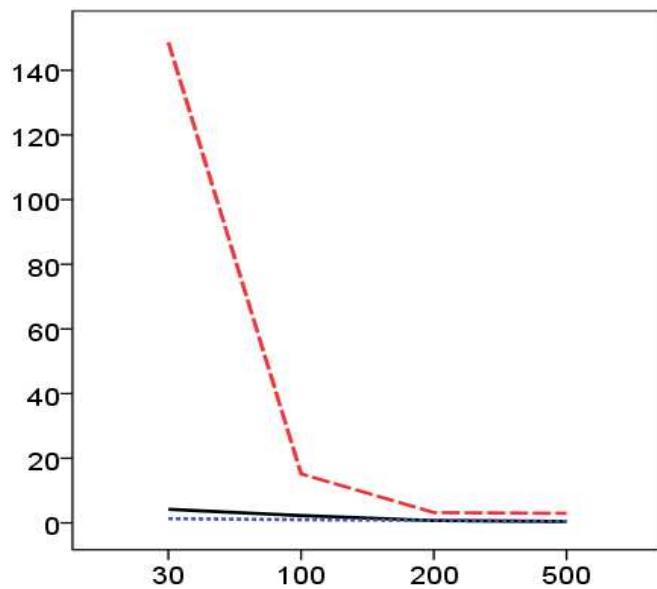
همان‌طور که در جدول‌های ۲ و ۳ ملاحظه می‌شود دقت برآورده روش بیزی غیرمستقیم برای پارامتر میانگین همانند دقت رهیافت بیزی مستقیم است. از این رو در برخی از کاربردها می‌توان از رهیافت بیزی غیرمستقیم به جای رهیافت بیزی مستقیم استفاده نمود. علت بیش برآورده و کم برآورده با توجه به رابطه (۲) قابل توجیه است. همان‌گونه که در حالت تعییم‌یافته (۲) می‌توان مشاهده کرد  $E(X) = \mu + \sqrt{\frac{2}{\pi} \delta \sigma}$ ، بنابراین با مشبت بودن  $\delta$  در برآورده انتظار بیش برآورده و با منفی بودن آن انتظار کم برآورده وجود دارد. نتایج بیانگر آنست که وقتی قدر مطلق چولگی بزرگ است برآورده ماقسیم درستنمایی نسبت به رهیافت بیزی دارای خطای بسیار فاحشی است. این نتیجه با مطالعه اسپینوسا (۲۰۰۸) نیز همخوانی دارد. همچنین از لحاظ شهودی شکل‌های ۱ تا ۳ نشان دهنده دقت یکسان برآورده پارامتر چولگی با دو رهیافت بیزی غیرمستقیم و مستقیم است. بعلاوه میزان خطای برآورده، علی‌الخصوص پارامتر چولگی، در حجم نمونه  $200$  و بالاتر کاهش چشمگیری داشته است. برآورده پارامتر پراکنده‌گی در روش بیزی غیرمستقیم دارای کم برآورده می‌باشد. علت آن را می‌توان به این خاطر دانست که در روش بیزی غیرمستقیم، از دو مولفه تصادفی استفاده شده که این مسئله باعث کاهش برآورده پارامتر پراکنده‌گی در روش بیزی غیرمستقیم و به تناسب آن افزایش واریانس برآوردها و کاهش دقت سایر برآوردها مانند میانگین و پارامتر چولگی می‌گردد (بندیوپدهایی و همکاران ۲۰۱۵).

## تقدیر و تشکر

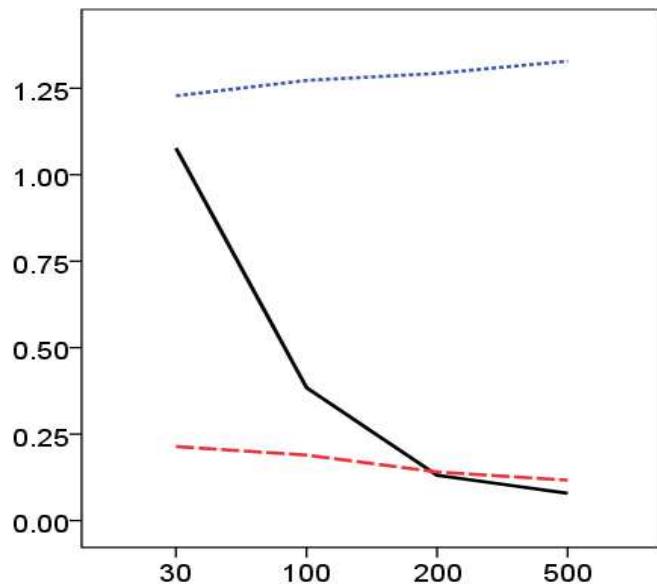
نویسنده‌گان از داوران و ویراستار محترم مجله بخاطر پیشنهادات ارزشمندانه‌ان که موجب ارائه بهتر مقاله شد، نهایت تشکر را دارند.



شکل ۱: متوسط خطابراورد میانگین توزیع چوله نرمال با روش های بیزی مستقیم (خط مستقیم)، غیرمستقیم (خط نقطه چین) و ماکسیمم درستنمایی (خط چین)



شکل ۲: متوسط خطابراورد پارامتر چولگی چوله نرمال با روش های بیزی مستقیم (خط مستقیم)، غیرمستقیم (خط نقطه چین) و ماکسیمم درستنمایی (خط چین)



شکل ۳: متوسط خطاب را برآورد واریانس توزیع چوله‌نرمال با روش‌های بیزی مستقیم (خط مستقیم)، غیرمستقیم (خط نقطه چین) و ماکسیمم درستنمایی (خط چین)

## مراجع

- Azzalini, A. (1985), A Class of Distributions Which Includes the Normal Ones.  
*Scandinavian Journal of Statistics, Theory and Applications*, **12**, 171-178.
- Azzalini, A. (2005), The Skew-Normal Distribution and Related Multivariate Families, *Scandinavian Journal of Statistics*, **32**, 159-188.
- Bandyopadhyay, D., Lachos, V. H., Abanatovalle C. A. and Ghosh, P. (2010),  
*Linear Mixed Models for Skew-Normal Independent Bivariate Responses with an Application to Periodontal Disease*,, *Statistics in Medicine*, **29**, 2643-2655.
- Bolstand, W. M. (2004), *Introduction to Bayesian Statistics*, Canada, John Wiley & Sons.

۴۷ ..... محمد غلامی فشارکی و همکاران

Bowling, S. R. K., and Mohammad T. (2009), A Logistic Approximation to the Cumulative Normal Distribution, *Journal of Industrial Engineering and Management*, **2**, 114-127.

Congdon, P. (2003), *Applied Bayesian Modeling*, England, John Wiley & Sons.

Genton, M. G. (2004), *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*, USA, Chapman & Hall/CRC.

Goldstein, M. and Wooff, D. (2007), *Bayes Linear Statistics Theory and Methods*, England, John Wiley & Sons.

Jara, A., Quintana, F. and Sanmartin E. (2008), Linear Mixed Models with Skew-Elliptical Distributions: a Bayesian Approach, *Computational Statistics and Data Analysis*, **52**, 5033-5045.

Kass, R. E. and Raftery, A. E. (1995), Bayes Factors, *Journal of the American Statistical Association*, **90**, 773-795.

Lee, J. C., Lin, T. I., Lee, K. J and Hsu, Y. L. (2005), Bayesian Analysis of Box-Cox Transformed Linear Mixed Models with ARMA(p,q) Dependence, *Journal of Statistical Planning and Inference*, **133**, 435-451.

Nadarajah, S. (2008), A New Model for Symmetric and Skewed Data, *Probability in the Engineering and Informatinal Sciences*, **22**, 261-271.

Nadarajah, S. and Kotz, S. (2003), Skewed Distributions Generated by the Normal kernel, *Statistics and Propability Letters*, **65**, 269-277.

Raftery, A. E. (1986), Choosing Models for Cross-Classifications, *American Sociological Review*, **51**, 145-146.

Raftery, A. E. (1993), *Bayesian Model Selection in Structural Equation Models*, In Testing Structural Equation Models, Beverly Hills: Sage.

Saville, B. R. (2008), *Bayesian Multilevel Models and Medical Applications*. School of Public Health, Chapel Hill, University of North Carolina, PHD of Biostatistics.

## برآوردهای بیزی پارامترهای توزیع چوله نرمال ..... ۴۸

Spinosa, J. S. (2008), *Inference for the Multiparameter Skew Normal Distribution*, Applied Statistics, California, University of California Riverside, PHD.

Tsung, I. L. and Lee, J. C. (2003), On Modelling Data From Degradation Sample Paths over Time, *Australian and New Zealand Journal of statistics*, **45**, 257-270.

### ضمیمه

کد برنامه برآوردهای بیزی با روش های مستقیم و غیر مستقیم در نرم افزار WinBUGS

کد نویسی مستقیم	کد نویسی غیر مستقیم
model	model
{	{
for(i in 1:n) {	for(i in 1:n) {
zeros[i] <- 0	mu2[i]<-mu+lambda*abs(k[i])
z[i]<-(x[i]-mu)/sigma	x[i]~dnorm(mu2[i],tau)
logf[i]<--0.5*pow(z[i],2)	k[i]~dnorm(0,tau)
w[i]<-lambda*z[i]	}
logG[i]<-log(phi(w[i]))	sigma <- 1 /sqrt(tau)
logLike[i]<--log(sigma)+logf[i]+logG[i]	mu~dnorm(0,0.1)
phi[i] <- -logLike[i]	tau~dgamma(0.1,0.1)
zeros[i] ~ dpois(phi[i])	lambda~dnorm(0,0.1)
}	}
mu~dnorm(0,0.1)	
tau~dgamma(0.1,0.1)	
lambda~dnorm(0,0.1)	
sigma <- 1 / sqrt(tau)}	