سیزدهمین کنفرانس آمار ایران
۲-۴ شهریور ۱۳۹۵

تحلیل کواریانس بیزی تحت مدل چوله نرمال

افشین فلاح [۱]، زهرا گودرزی [۲]
[۱]گروه آمار، دانشگاه بین‌المللی امام خمینی (ره) قزوین
[۲]گروه آمار، دانشگاه بین‌المللی امام خمینی (ره) قزوین

چکیده: در این مقاله برای تحلیل کواریانس در شرایطی که توزیع متغیر پاسخ چوله نرمال است، یک مدل بیزی توسعه داده شده است. چون توزیع پسین پارامترها در مدل پیشنهادی دارای فرم بسته نیست، توزیع‌های پسین شرطی کامل پارامترها محاسبه و یک الگوریتم گیبز برای نمونه‌گیری از توزیع پسین توسعه داده شده است. مدل پیشنهادی در قالب یک مطالعه‌ی شبیه‌سازی مورد ارزیابی قرار گرفته و با مدل‌های رقیب مقایسه شده است.

واژه‌های کلیدی: تحلیل کواریانس، توزیع چوله نرمال، تابع درستنمایی داده‌های کامل، متغیر تبیینی

کد موضوع‌بندی ریاضی (۲۰۱۰): 62c10، 62H25، 62F12.

---

۱ مقدمه

برای تعیین رابطه بین یک متغیر تصادفی پاسخ و مجموعه‌ای از متغیرهای تبیینی یا عامل‌ها می‌توان از روش‌های مختلف تحلیل رگرسیونی و تحلیل واریانس ($ANOVA$) استفاده کرد. تحلیل کواریانس، تلفیقی از تحلیل رگرسیونی و تحلیل واریانس است. ایده‌ی اولیه تحلیل کواریانس توسط فیشر (۱۹۳۲) طرح شده است. ککران (۱۹۵۷) مفاهیم و کاربردهای مختلف تحلیل کواریانس را مورد بحث قرار داده است. نظریه سنتی تحلیل کواریانس همانند تحلیل رگرسیونی و تحلیل واریانس، عمدتاً بر فرض نرمال بودن متغیر پاسخ بنا شده است. در بسیاری از این کاربردها این فرضیه به دلایل مختلف نقض می‌شود. راهکاری که می‌توان در چنین شرایطی در پیش گرفت، برازش مدل تحلیل کواریانس تحت فرض توزیع‌های دیگری است که بتوانند به خوبی عدم تقارن مشاهدات را در مدل‌سازی لحاظ نماید. در شرایطی که مشاهدات متقارن نیستند، جایگزین نمودن یک توزیع چوله به جای توزیع نرمال به عنوان توزیع متغیر پاسخ، می‌تواند به بهبود قابل توجه در کارایی مدل تحلیل کواریانس منجر شود. یکی از توزیع‌هایی که لزوماً متقارن نیست اما خواصی مشابه توزیع نرمال دارد، به توزیع چوله‌نرمال معروف است. این توزیع دارای پارامتری برای مدل‌بندی چولگی است و توزیع نرمال را به عنوان یک عضو شامل می‌شود.

---

[۲]نام ارائه دهنده مقاله : goodarzizahra67@yahoo.com

تعریف رایج این خانواده برای نخستین بار توسط آزالینی (۱۹۸۵) ارائه شد. در طول بیش از سه دهه گذشته توزیع‌های چوله متقارن و چوله نرمال متعددی توسط پژوهشگران مختلف معرفی و مورد مطالعه قرار گرفته است. در این مقاله، تحلیل کواریانس تحت مدل چوله نرمال از دیدگاه رویکرد بیزی مورد توجه قرار گرفته است. برای این منظور، با انتخاب توزیع‌های پیشین برای پارامترهای مدل و استفاده از تابع درستنمایی داده‌های کامل، توزیع پسین توام پارامترها به‌صورت متناسب محاسبه و از روش‌های مونت کارلو زنجیر مارکفی برای نمونه‌گیری از توزیع پسین توام پارامترها و استنباط‌های پسینی استفاده شده است. به منظور ارزیابی براوردگرهای پیشنهادی از یک مطالعه شبیه‌سازی استفاده است.

۲ مدل پیشنهادی

فرض کنید مدل تحلیل واریانس و رابطه رگرسیونی به ترتیب از طریق ساختار ماتریس طرح $\boldsymbol{X}$ و ماتریس متغیرهای تبیینی $\boldsymbol{Z}$ مشخص می‌شوند. در این صورت، مدل تحلیل کواریانس را می‌توان به شکل $\boldsymbol{\mu}=\boldsymbol{X\beta}+\boldsymbol{Z\gamma}=\boldsymbol{W\theta}$ نوشت، که در آن $\boldsymbol{\mu}$ بردار میانگین مشاهدات پاسخ، $\boldsymbol{\beta}=(\beta_0,\beta_1,\ldots,\beta_{s-1})'$ بردار عامل‌ها، $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_q)$ ضرایب رگرسیونی، $\boldsymbol{W}=[\boldsymbol{X}|\boldsymbol{Z}]$ ماتریس مشاهدات و $\boldsymbol{\theta}=(\boldsymbol{\beta},\boldsymbol{\gamma})'$ بردار ضرایب مدل است. هدف اصلی، براورد $\boldsymbol{\theta}$ با استفاده از مشاهدات $\boldsymbol{y}=(y_{11},\ldots,y_{sr})'$ و $\boldsymbol{W}$ است. در این بخش تحلیل کواریانس تحت فرض چوله نرمال بودن توزیع متغیر پاسخ مورد توجه قرار گرفته است. متغیر تصادفی Y دارای توزیع چوله نرمال است ($Y\sim SSN(\mu,\sigma^2,\lambda)$)، هرگاه تابع چگالی آن به‌صورت

$$f_Y(y)=2\phi(y|\mu,\sigma^2+\lambda^2)\Phi\Big(\frac{\lambda}{\sigma}\frac{(y-\mu)}{(\sigma^2+\lambda^2)^{\frac{1}{2}}}\Big),$$

باشد، که درآن $\phi(.|\mu,\sigma^2)$ و $\Phi(.|\mu,\sigma^2)$ به ترتیب تابع چگالی و تابع توزیع تجمعی توزیع نرمال با میانگین μ و واریانس σ^2 هستند ساهو و همکاران (۲۰۰۳). این توابع به ازای $(\mu,\sigma)=(0,1)$ به ترتیب با نمادهای $\phi(\cdot)$ و $\Phi(\cdot)$ نشان داده می‌شوند. این توزیع به ازای $\lambda<0$ دارای چولگی منفی، به ازای $\lambda>0$ دارای چولگی مثبت و به ازای $\lambda=0$ متقارن بوده و به توزیع نرمال تبدیل می‌شود.

مدل تحلیل کواریانس تحت فرض چوله نرمال بودن متغیر پاسخ را به‌صورت

$$Y_{ij}|\boldsymbol{w}_{ij}\sim SSN(\boldsymbol{w}_{ij}\boldsymbol{\theta}-\sqrt{\frac{2}{\pi}}\lambda,\sigma^2,\lambda),\qquad i=1,\ldots,s;\ j=1,\ldots,r, \tag{۱.۲}$$

در نظر بگیرید. تابع درستنمایی متناظر با مدل (۱.۲) به‌صورت

$$f(\boldsymbol{y},\boldsymbol{W}|\boldsymbol{\theta},\sigma^2,\lambda)=\prod_{i=1}^{s}\prod_{j=1}^{r}2\phi(y_{ij}|\boldsymbol{w}_{ij}\boldsymbol{\theta}-\sqrt{\frac{2}{\pi}},\sigma^2+\lambda^2)\Phi(\frac{\lambda}{\sigma}\frac{(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta})}{(\sigma^2+\lambda^2)^{\frac{1}{2}}}),$$

است. ملاحظه می‌شود به دلیل پیچیدگی تابع درستنمایی براوردگرهای ماکسیمم درستنمایی پارامترهای مدل تحلیل کواریانس تحت فرض چوله نرمال بودن توزیع متغیر پاسخ فرم بسته ندارند. به‌همین دلیل، فلاح و گودرزی (۲۰۱۶) از الگوریتم EM برای محاسبه مقادیر عددی براوردهای ماکسیمم درستنمایی پارامترها و فواصل اطمینان مجانبی متناظر آن‌ها استفاده کردند. توزیع چوله نرمال را می‌توان به‌صورت آمیخته‌ای از توزیع‌های نرمال و نیم‌نرمال به‌شکل $T_{ij}\sim HN(0,1)$ و $Y_{ij}|T_{ij}=t_{ij}\sim N(\boldsymbol{w}_{ij}\boldsymbol{\theta}+\lambda(t_{ij}-\sqrt{\frac{2}{\pi}}),\sigma^2)$ ، $i=1,\ldots,s;\ j=1,\ldots,r$ نوشت. بر این اساس اگر Y_{ij}ها داده‌های ناقص و T_{ij}ها داده‌های گم‌شده تلقی شوند، تابع چگالی توام (Y_{ij},T_{ij}) را می‌توان به‌صورت

$$f_{(Y_{ij},T_{ij})}(y_{ij},t_{ij})\ =\ \frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big\{-\frac{1}{2\sigma^2}(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}-\lambda(t_{ij}-\sqrt{\frac{2}{\pi}}))^2\Big\}\times\sqrt{\frac{2}{\pi}}\exp\Big\{-\frac{t_{ij}^2}{2}\Big\},$$

نوشت. بنابراین تابع درستنمایی داده‌های کامل به‌صورت

$$f_c(\boldsymbol{y},\boldsymbol{W},\boldsymbol{t}|\boldsymbol{\theta},\sigma^2,\lambda) = \exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^{s}\sum_{j=1}^{r}\Big((y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta})^2-2\lambda(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}+\sqrt{\frac{2}{\pi}}\lambda)t_{ij} + (\lambda^2+\sigma^2){t^2}_{ij}+2\lambda\sqrt{\frac{2}{\pi}}(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta})+\lambda^2\frac{2}{\pi}\Big)\Big\}\times(\pi)^{-sr}(\sigma^2)^{-\frac{sr}{2}}, \quad (2.2)$$

است. از آن‌جا که خواص بهینه براوردگرهای ماکسیمم درستنمایی تنها در حالت مجانبی بروز می‌یابند رویکرد فراوانی‌گرایانه مبتنی بر الگوریتم EM برای نمونه‌های کوچک از بهینگی دور است و از کارایی لازم برخوردار نیست. در رویکرد بیزی، پارامترهای مدل متغیرهای تصادفی تلقی و برای آن‌ها توزیع‌های پیشین مناسب در نظر گرفته می‌شود. حساس‌ترین مرحله تحلیل بیزی، تخصیص توزیع‌های پیشین به پارامترهای مدل است. با توجه به دامنه تغییرات پارامتر چولگی برای آن می‌توان از توزیع‌های پیشین متفاوتی استفاده کرد. بیز و برانکو (۲۰۰۷) تقریبی از توزیع پیشین جفریز را برای این منظور به کار برده‌اند. کانچو و همکاران (۲۰۰۸) از توزیع پیشین نرمال بریده شده استفاده کرده‌اند. در این مقاله، توزیع‌های پیشین پارامترها به‌صورت

$$\begin{aligned}\boldsymbol{\theta} &\sim N_P(\boldsymbol{\eta}_0,\boldsymbol{\Sigma}_0), \quad \boldsymbol{\eta}_0\in\mathbb{R}^p,\\ \sigma^2 &\sim IG(\frac{m_0}{2},\frac{h_0}{2}), \quad m_0,h_0>0,\\ \lambda &\sim N(\mu_0,\sigma_0^2), \quad \mu_0\in\mathbb{R},\ \sigma_0^2>0,\end{aligned} \quad (3.2)$$

انتخاب شده‌اند، که در آن‌ها $\boldsymbol{\eta}_0=(\eta_1,\dots,\eta_p)$، $\boldsymbol{\Sigma}_0=diag(\delta_1,\dots\delta_p)$، μ_0، σ_0^2، m_0 و h_0 ابر پارامترهای مدل هستند و توسط تحلیل‌گر تعیین یا براورد می‌شوند. با فرض استقلال پیشینی پارامترها، توزیع پیشین برای $\boldsymbol{\Psi}=(\boldsymbol{\theta},\sigma^2,\lambda)$ را می‌توان به‌صورت

$$\pi(\boldsymbol{\theta},\sigma^2,\lambda)=\phi_p(\boldsymbol{\theta};\boldsymbol{\eta}_0,\boldsymbol{\Sigma}_0)d.IG(\sigma^2;\frac{m_0}{2},\frac{h_0}{2})\phi(\lambda;\mu_0,\sigma_0^2),$$

نوشت، که در آن‌ها $\phi(\cdot)$، $\phi_p(\cdot)$ و $d.IG(\cdot)$ به ترتیب توابع چگالی نرمال یک متغیره، نرمال چند متغیره و گامای وارون را نشان می‌دهند. با توجه به توزیع‌های پیشین (۳.۲) و تابع درستنمایی داده‌های کامل (۲.۲)، توزیع پسین پارامترها به‌صورت

$$\pi(\boldsymbol{\theta},\sigma^2,\lambda|\boldsymbol{y},\boldsymbol{W},\boldsymbol{t}) \propto f_c(\boldsymbol{y},\boldsymbol{W},\boldsymbol{t}|\boldsymbol{\theta},\sigma^2,\lambda)\phi_p(\boldsymbol{\theta};\boldsymbol{\eta}_0,\boldsymbol{\Sigma}_0)d.IG(\sigma^2;\frac{m_0}{2},\frac{h_0}{2})\phi(\lambda;\mu_0,\sigma_0^2), \quad (4.2)$$

به‌دست می‌اید. به دلیل فرم پیچیده توزیع پسین (۴.۲)، کمیت‌های پسینی فاقد فرم بسته هستند. از این رو، در ادامه از روش‌های مونت کارلو زنجیر مارکفی برای نمونه‌گیری از این توزیع و انجام استنباط‌های پسینی استفاده می‌شود.

۱.۲ توزیع‌های پسین شرطی کامل

در این بخش نحوه کاربست الگوریتم گیبز برای نمونه‌گیری از توزیع پسین (۴.۲) شرح داده می‌شود. لازمه استفاده از الگوریتم گیبز شناخت توزیع‌های پسین شرطی کامل است. با اندکی محاسبات می‌توان نشان داد که توزیع‌های پسین شرطی کامل پارامترهای مدل و داده‌های گم‌شده به‌صورت

$$\begin{aligned}t_{ij}|\boldsymbol{\theta},\sigma^2,\lambda,\boldsymbol{y} &\sim HN(\frac{\lambda}{\sigma^2+\lambda^2}(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}+\lambda\sqrt{\frac{2}{\pi}})^2,\frac{\sigma^2}{\sigma^2+\lambda^2}),\\ \sigma^2|\boldsymbol{\theta},\lambda,\boldsymbol{y},t_{ij} &\sim IG(\frac{sr+m}{2},\frac{1}{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}-\lambda(t_{ij}-\sqrt{\frac{2}{\pi}}))^2+\frac{h}{2}),\end{aligned}$$

$$\lambda|\boldsymbol{\theta},\sigma^{2},\boldsymbol{y},t_{ij} \sim N\Big(\frac{\sigma^{2}\mu_{0}+\sigma_{0}^{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(t_{ij}-\sqrt{\frac{2}{\pi}})(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta})}{\sigma^{2}+\sigma_{0}^{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(t_{ij}-\sqrt{\frac{2}{\pi}})^{2}},\frac{\sigma^{2}\sigma_{0}^{2}}{\sigma^{2}+\sigma_{0}^{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(t_{ij}-\sqrt{\frac{2}{\pi}})^{2}}\Big),$$

$$\boldsymbol{\theta}|\sigma^{2},\lambda,\boldsymbol{y},t_{ij} \propto \phi_{P}(\boldsymbol{\theta};\boldsymbol{\eta}_{0},\Sigma_{0})\prod_{i=1}^{s}\prod_{j=1}^{r}\phi(y_{ij}|\boldsymbol{w}_{ij}\boldsymbol{\theta}+\lambda(t_{ij}-\sqrt{\frac{2}{\pi}}),\sigma^{2}), \tag{۵.۲}$$

هستند، از این‌رو، می‌توان از الگوریتم گیبز برای نمونه‌گیری از توزیع پسین استفاده کرد. همانطورکه مشاهده می‌شود توزیع پسین شرطی کامل (۵.۲) صورت بسته و شناخته شده‌ای ندارد. بنابراین لازم است از الگوریتم متروپولیس-هاستینگز برای نمونه‌گیری از این توزیع پسین شرطی کامل استفاده نمود. بر این اساس، برای نمونه‌گیری از توزیع پسین توام $(\boldsymbol{\theta},\sigma^{2},\lambda,\boldsymbol{t})$ به تلفیقی از الگوریتم های گیبز و متروپولیس-هاستینگز نیاز است.

مراحل کلی الگوریتم گیبز برای نمونه‌گیری از توزیع پسین (۴.۲) به صورت زیر است:

۱- مقادیر اولیه $\boldsymbol{\theta}^{(0)}$، $\sigma^{2(0)}$ و $\lambda^{(0)}$ را انتخاب کنید.

۲- در مرحله‌ی $(k+1)$ام، براساس مقادیر حاصل از مرحله‌ی kام، پارامترها و متغیر گم‌شده را به‌صورت زیر به روز کنید:

$$\begin{aligned}
t_{ij}^{(k+1)} &\sim HN\Big(\frac{\lambda^{(k)}}{\sigma^{2(k)}+\lambda^{2(k)}}(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}^{(k)}+\lambda^{(k)}\sqrt{\frac{2}{\pi}})^{2},\frac{\sigma^{2(k)}}{\sigma^{2(k)}+\lambda^{2(k)}}\Big),\\
\sigma^{2(k+1)} &\sim IG\Big(\frac{sr+m}{2},\frac{1}{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}^{(k)}-\lambda^{(k)}(t_{ij}^{(k+1)}-\sqrt{\frac{2}{\pi}}))^{2}+\frac{h}{2}\Big),\\
\lambda^{(k+1)} &\sim N\Big(\frac{\sum_{i=1}^{s}\sum_{j=1}^{r}(t_{ij}^{(k+1)}-\sqrt{\frac{2}{\pi}})(y_{ij}-\boldsymbol{w}_{ij}\boldsymbol{\theta}^{(k)})\sigma_{0}^{2}+\mu_{0}\sigma^{2(k+1)}}{\sigma_{0}^{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(t_{ij}^{(k+1)}-\sqrt{\frac{2}{\pi}})^{2}+\sigma^{2(k+1)}}\\
&\qquad,\frac{\sigma^{2(k+1)}\sigma_{0}^{2}}{\sigma_{0}^{2}\sum_{i=1}^{s}\sum_{j=1}^{r}(t_{ij}^{(k+1)}-\sqrt{\frac{2}{\pi}})^{2}+\sigma^{2(k+1)}}\Big),\\
\boldsymbol{\theta}^{(k+1)} &\propto \phi_{P}(\boldsymbol{\theta};\boldsymbol{\eta}_{0},\Sigma_{0})\prod_{i=1}^{s}\prod_{j=1}^{r}\phi(y_{ij}|\boldsymbol{w}_{ij}\boldsymbol{\theta}^{(k)}+\lambda^{(k+1)}(t_{ij}^{(k+1)}-\sqrt{\frac{2}{\pi}}),\sigma^{2(k+1)}).
\end{aligned}$$

۳- مرحله ۲ الگوریتم را تا زمان رسیدن زنجیر مارکف به توزیع مانای خود تکرار کنید.

۳ مطالعه شبیه‌سازی

در این بخش مدل بیزی پیشنهادی در چهارچوب یک مطالعه شبیه‌سازی مورد بررسی و ارزیابی قرار گرفته است. برای این منظور، یک مدل تحلیل کواریانس با یک متغیر تبیینی و یک عامل دوسطحی به‌صورت

$$\mu_{ij}=E(y_{ij}|\boldsymbol{X},\boldsymbol{Z})=\boldsymbol{w}_{ij}\boldsymbol{\theta}=\boldsymbol{\beta}_{0}+\boldsymbol{\beta}_{i}+\boldsymbol{\gamma}(\boldsymbol{z}_{ij}-\bar{\boldsymbol{z}}_{i}),\quad i=1,2\ ;\ j=1,\dots,r,$$

به ازای $\beta_{0}=2$ ، $\beta_{1}=1$ و $\gamma=1$ در نظر گرفته شده است. متغیرهای پاسخ y_{ij}، $i=1,2$ و $j=1,\cdots,r$، از توزیع $SSN(\boldsymbol{w}_{ij}\boldsymbol{\theta}-\sqrt{\frac{2}{\pi}},\sigma^{2},\lambda)$ شبیه‌سازی شده‌اند. برای ارزیابی تأثیر اندازه نمونه بر کارایی روش پیشنهادی، مطالعه‌ی شبیه‌سازی به ازای مقادیر مختلف اندازه نمونه $\{10,20,40,100\}$ تکرار شده است. چون رویکرد پیشنهادی بیزی است از در نظرگرفتن اندازه‌های

بزرگ‌تر پرهیز شده است، چرا که مزیت اصلی رویکرد مدل بیزی عمدتاً مربوط به اندازه‌های کوچک نمونه‌ای است و برای نمونه‌های بزرگ به دلیل غلبه تابع درستنمایی بر توزیع پیشین، نتایج بر نتایج رویکرد فراوانی‌گرایانه منطبق خواهد بود. هم‌چنین به منظور ارزیابی قابلیت مدل پیشنهادی در مدل‌بندی مشاهدات دارای وضعیت‌های مختلف تقارن و چولگی، مقادیر مختلفی برای پارامتر چولگی در محدوده $\{-2, -1, 0, 1, 2\}$ لحاظ شده است. تمامی محاسبات ۲۰۰۰ بار تکرار و مقادیر ریشه توان دوم خطا براوردگرهای ماکسیمم درستنمایی پارامترهای مدل تحلیل کواریانس تحت شرایط بالا محاسبه و در جداول ۱ الی ۴ نشان داده شده است. مقادیر متناظر برای براوردگرهای ماکسیمم درستنمایی پارامترهای مدل که از طریق کاربست الگوریتم EM براساس رویکرد پیشنهادی فلاح و گودرزی (۲۰۱۶) نیز محاسبه و در جداول ارائه شده‌اند. در تمامی موارد، مدل نرمال نیز به عنوان مدل سنتی و به منظور نشان دادن اثر مخرّب انحراف از فرض نرمال بودن مشاهدات پاسخ، مدنظر قرار گرفته است.

جدول ۱: مقادیر ریشه میانگین توان دوم خطا براوردگرهای ماکسیمم درستنمایی ضرایب مدل تحلیل کواریانس چوله نرمال در رهیافت بیزی و فراوانی‌گرایانه به همراه مقادیر متناظر برای توزیع نرمال به ازای اندازه نمونه ۱۰.

			پارامترها			
پارامتر چولگی	مدل	براوردگر	β_0	β_1	γ	λ
−۲	چوله نرمال	فراوانی‌گرایانه	۰/۷۰۰۷	۱/۰۱۳۴	۷/۵۰۴۳	۲/۹۸۳۶
		بیزی	۰/۴۲۷۲	۰/۵۳۴۰	۰/۳۲۹۱	۱/۹۴۵۶
	نرمال	فراوانی‌گرایانه	۰/۶۹۵۱	۱/۰۰۱۹	۷/۵۸۳۸	–
−۱	چوله نرمال	فراوانی‌گرایانه	۰/۵۱۱۰	۰/۷۴۱۲	۵/۴۲۷۹	۱/۷۱۵۶
		بیزی	۰/۳۲۹۷	۰/۴۱۳۷	۰/۲۷۱۸	۱/۰۸۰۵
	نرمال	فراوانی‌گرایانه	۰/۵۱۳۷	۰/۷۴۲۶	۵/۵۳۴۹	–
۰	چوله نرمال	فراوانی‌گرایانه	۰/۴۲۸۱	۰/۵۹۱۸	۴/۴۱۹۳	$< 1 \times 10^{-13}$
		بیزی	۰/۲۸۸۹	۰/۳۴۳۴	۰/۲۴۷۸	۰/۲۹۱۶
	نرمال	فراوانی‌گرایانه	۰/۴۲۸۱	۰/۵۹۱۸	۴/۴۱۹۳	–
۱	چوله نرمال	فراوانی‌گرایانه	۰/۵۱۶۳	۰/۷۲۶۳	۵/۱۹۸۹	۰/۴۸۵۸
		بیزی	۰/۳۴۴۱	۰/۴۲۱۱	۰/۲۷۰۷	۱/۰۲۴۸
	نرمال	فراوانی‌گرایانه	۰/۵۳۲۳	۰/۷۵۰۱	۵/۶۲۸۰	–
۲	چوله نرمال	فراوانی‌گرایانه	۰/۶۵۷۶	۰/۹۲۳۵	۶/۷۲۳۰	۰/۷۵۸۹
		بیزی	۰/۴۲۶۴	۰/۵۳۱۱	۰/۳۳۷۰	۱/۸۲۳۴
	نرمال	فراوانی‌گرایانه	۰/۶۹۶۲	۰/۹۹۷۰	۷/۵۲۰۷	–

جدول ۲: مقادیر ریشه میانگین توان دوم خطا براوردگرهای ماکسیمم درستنمایی ضرایب مدل تحلیل کواریانس چوله نرمال در رهیافت بیزی و فراوانی‌گرایانه به همراه مقادیر متناظر برای توزیع نرمال به ازای اندازه نمونه ۲۰.

			پارامترها			
پارامتر چولگی	مدل	براوردگر	$\beta_۰$	$\beta_۱$	γ	λ
−۲	چوله نرمال	فراوانی‌گرایانه	۰/۴۹۷۷	۰/۷۲۹۵	۵/۰۳۳۰	۳/۱۰۹۶
		بیزی	۰/۳۵۴۰	۰/۴۷۶۷	۰/۴۶۰۳	۱/۷۰۴۵
	نرمال	فراوانی‌گرایانه	۰/۴۹۶۱	۰/۷۲۳۸	۵/۰۶۲۸	–
−۱	چوله نرمال	فراوانی‌گرایانه	۰/۳۶۹۴	۰/۵۲۳۴	۳/۶۳۲۸	۱/۸۱۷۰
		بیزی	۰/۲۸۳۸	۰/۳۶۸۱	۰/۳۸۱۷	۱/۰۰۵۸
	نرمال	فراوانی‌گرایانه	۰/۳۷۲۶	۰/۵۲۷۸	۳/۷۱۰۹	–
۰	چوله نرمال	فراوانی‌گرایانه	۰/۳۲۳۸	۰/۴۵۳۹	۳/۲۱۵۵	$< ۱ \times ۱۰^{-۱۳}$
		بیزی	۰/۲۴۸۹	۰/۳۲۰۱	۰/۳۱۴۷	۰/۱۶۵۵
	نرمال	فراوانی‌گرایانه	۰/۳۲۳۸	۰/۴۵۳۹	۳/۲۱۵۵	–
۱	چوله نرمال	فراوانی‌گرایانه	۰/۳۴۸۰	۰/۴۸۸۶	۳/۴۳۹۴	۰/۳۹۷۲
		بیزی	۰/۲۷۷۸	۰/۳۶۱۴	۰/۳۹۴۴	۰/۹۸۲۷
	نرمال	فراوانی‌گرایانه	۰/۳۶۳۱	۰/۵۱۱۸	۳/۷۷۶۲	–
۲	چوله نرمال	فراوانی‌گرایانه	۰/۴۶۱۱	۰/۶۵۶۱	۴/۳۳۳۸	۰/۵۰۵۳
		بیزی	۰/۳۵۸۸	۰/۴۶۶۵	۰/۴۱۸۳	۱/۶۳۹۵
	نرمال	فراوانی‌گرایانه	۰/۴۹۵۳	۰/۷۱۰۸	۵/۰۷۴۲	–

جدول ۳: مقادیر ریشه میانگین توان دوم خطا براوردگرهای ماکسیمم درستنمایی ضرایب مدل تحلیل کواریانس چوله نرمال در رهیافت بیزی و فراوانی‌گرایانه به همراه مقادیر متناظر برای توزیع نرمال به ازای اندازه نمونه ۴۰.

			پارامترها			
پارامتر چولگی	مدل	براوردگر	$\beta_۰$	$\beta_۱$	γ	λ
−۲	چوله نرمال	فراوانی‌گرایانه	۰/۳۴۱۱	۰/۴۸۹۹	۲/۷۴۶۳	۳/۱۸۷۰
		بیزی	۰/۲۷۵۶	۰/۳۶۵۷	۰/۶۴۸۸	۱/۵۳۲۶
	نرمال	فراوانی‌گرایانه	۰/۳۴۲۰	۰/۴۸۵۶	۲/۷۷۴۶	–
−۱	چوله نرمال	فراوانی‌گرایانه	۰/۲۵۹۶	۰/۳۶۹۸	۱/۹۵۸۷	۱/۸۷۳۴
		بیزی	۰/۲۲۲۶	۰/۳۰۴۶	۰/۵۰۷۳	۰/۹۶۸۶
	نرمال	فراوانی‌گرایانه	۱/۸۷۳۴	۰/۳۷۲۶	۲/۰۰۵۹	–
۰	چوله نرمال	فراوانی‌گرایانه	۰/۲۲۴۱	۰/۳۰۹۴	۱/۷۱۰۵	$< ۱ \times ۱۰^{-۱۳}$
		بیزی	۰/۱۹۳۳	۰/۲۵۵۱	۰/۳۹۱۲	۰/۰۷۹۲
	نرمال	فراوانی‌گرایانه	۰/۲۲۴۱	۰/۳۰۹۴	۱/۷۱۰۵	–
۱	چوله نرمال	فراوانی‌گرایانه	۰/۲۴۷۹	۰/۳۴۶۶	۱/۷۸۳۹	۰/۳۶۳۵
		بیزی	۰/۲۲۲۰	۰/۲۹۸۷	۰/۴۹۱۷	۰/۹۵۶۲
	نرمال	فراوانی‌گرایانه	۰/۲۶۱۳	۰/۳۶۸۷	۲/۰۱۳۲	–
۲	چوله نرمال	فراوانی‌گرایانه	۰/۳۰۴۵	۰/۴۲۹۲	۲/۲۴۳۹	۰/۳۱۰۳
		بیزی	۰/۲۷۲۶	۰/۳۶۶۶	۰/۵۷۰۰	۱/۴۹۶۴
	نرمال	فراوانی‌گرایانه	۰/۳۳۶۶	۰/۴۸۲۴	۲/۶۶۴۷	–

جدول ۴: مقادیر ریشه میانگین توان دوم خطا براوردگرهای ماکسیمم درستنمایی ضرایب مدل تحلیل کواریانس چوله نرمال در رهیافت بیزی و فراوانی‌گرایانه به همراه مقادیر متناظر برای توزیع نرمال به ازای اندازه نمونه ۱۰۰.

پارامتر چولگی	مدل	براوردگر	پارامترها			
			β_0	β_1	γ	λ
-۲	چوله نرمال	فراوانی‌گرایانه	۰/۲۱۸۸	۰/۳۱۰۴	۱/۵۷۷۲	۳/۲۳۵۵
		بیزی	۰/۱۹۷۱	۰/۲۷۱۳	۰/۶۰۷۳	۱/۸۰۹۴
	نرمال	فراوانی‌گرایانه	۰/۲۲۰۱	۰/۳۱۰۴	۱/۶۱۱۰	–
-۱	چوله نرمال	فراوانی‌گرایانه	۰/۱۶۳۳	۰/۲۳۱۹	۱/۱۰۲۰	۱/۹۱۴۸
		بیزی	۰/۱۵۵۵	۰/۲۱۶۲	۰/۴۵۵۹	۰/۹۸۸۷
	نرمال	فراوانی‌گرایانه	۰/۱۶۶۰	۰/۲۳۶۰	۱/۱۴۱۹	–
۰	چوله نرمال	فراوانی‌گرایانه	۰/۱۴۲۶	۰/۲۰۳۵	۱/۰۰۹۷	$< ۱ \times ۱۰^{-۱۳}$
		بیزی	۰/۱۳۵۱	۰/۱۸۹۴	۰/۴۴۳۸	۰/۰۱۸۸
	نرمال	فراوانی‌گرایانه	۰/۱۴۲۶	۰/۲۰۳۵	۱/۰۰۹۷	–
۱	چوله نرمال	فراوانی‌گرایانه	۰/۱۶۱۹	۰/۲۱۸۰	۰/۹۸۵۵	۰/۳۶۰۶
		بیزی	۰/۱۵۹۶	۰/۲۱۸۸	۰/۴۸۰۷	۰/۹۸۴۷
	نرمال	فراوانی‌گرایانه	۰/۱۷۱۴	۰/۲۴۰۱	۱/۱۵۳۳	–
۲	چوله نرمال	فراوانی‌گرایانه	۰/۱۹۲۶	۰/۲۷۳۳	۱/۲۸۰۶	۰/۱۸۵۸
		بیزی	۰/۱۹۳۱	۰/۲۷۱۷	۰/۵۸۶۳	۱/۷۹۱۷
	نرمال	فراوانی‌گرایانه	۰/۲۱۸۱	۰/۳۱۴۵	۱/۵۹۵۷	–

همانطور که ملاحظه می‌شود به ازای همه‌ی اندازه‌های نمونه، براوردگرهای مدل چوله نرمال حاصل از رویکرد بیزی در مقایسه با براوردگرهای مدل چوله نرمال و نرمال حاصل از رویکرد فراوانی‌گرایانه براساس مقادیر ریشه میانگین توان دوم خطا از دقت بیشتری برخوردار است. گر چه کاربست مدل تحلیل کواریانس از هر دو دیدگاه فراوانی‌گرایانه و بیزی به دلیل پیچیدگی‌های ذاتی مدل‌های سلسله مراتبی نیازمند استفاده از الگوریتم‌های تکراری بوده و با پیچیدگی‌های خاص خود همراه است، اما نتایج حاصل از شبیه‌سازی ارائه شده در این بخش نشان می‌دهد که میزان کارایی رهیافت بیزی در براورد پارامترها، دست کم به ازای اندازه‌های کوچک، به طور قابل ملاحظه‌ای از کارایی رهیافت فراوانی‌گرایانه بیش‌تر است.

بحث و نتیجه‌گیری

فرض نرمال بودن توزیع مشاهدات پاسخ که یکی از فرضیات پایه‌ای در تحلیل کواریانس است، در بسیاری از کاربردهای واقعی نقض می‌شود. به صورت خاص در بسیاری موارد مشاهدات متغیر پاسخ ساختاری تک‌مدی و چوله را نمایش می‌دهند. در این موارد خانواده توزیع‌های چوله نرمال به دلیل انعطاف‌پذیری مطلوب، می‌تواند توصیف کامل‌تری از پدیده تحت بررسی فراهم می‌آورد. نتایج این پژوهش نشان می‌دهند که در چنین شرایطی مدل تحلیل کواریانس تحت فرض چوله نرمال بودن توزیع متغیر پاسخ در مقایسه با مدل‌های تحلیل کواریانس سنتی از کارایی بیشتری برخوردار است. در این مقاله، از بین خانواده‌های مختلف توزیع‌های چوله نرمال، توزیع چوله نرمال ساهو با توجه به خصوصیات توزیعی جالب آن، مورد استفاده قرار گرفته است. اما خانواده‌های دیگر از توزیع چوله نرمال نیز می‌توان به شیوه‌ی مشابه به کار گرفته شوند. افزون براین، در مقام مقایسه رهیافت‌های بیزی و فراوانی‌گرایانه می‌توان گفت که تنها برای نمونه‌های بزرگ به دلیل بروز خواص مجانبی براوردگرهای ماکسیمم درستنمایی، می‌توان از رهیافت فراوانی‌گرایانه استفاده کرد.

مراجع

Azzalini A. (1985), A Class of Distributions which Includes the Normal ones, *Scandinanavi Journal Statistics*, 12, 171–178.

Barr R., Donald, E. and sherril. (1999), Mean and Variance of Truncated Normal Distribution, *The American Statistician*, 53, 357–361.

Bayes C. and Branco M. (2007), Bayesian Inference for the Skewness Parameterof the Scalar Skew-normal Distribution, *Brazilian Journal of Probabilityand Statistics*, 21, 141–163.

Cancho V., Lachos, V. and Ortega, E. (2008), Maximum A Nonlinear Regression Model with Skew-Normal Errors, *Statistical Paper*, 51, 547–558.

Cochran W. (1957), Analysis of Covariance: Its Nature and Uses, *Biometrics*, 13, 261–281.

Fallah A. and Goodarzi Z. (2016), Analysis of Covariance by Assuming a Skew Normal Distribution for Response Variable, *Haccetep Journal of Mathematics and Statistic*, doi:10.15672/HJMS.20159314058.

Fisher R. A. (1932), *Statistical Methods for Research Workers*, 4th ed., Oliver and Boyed, Edinburgh.

Sahu, S. , Dey, D. and Branco, M. (2003), A New Class of Multivariate Distributions with Applications to Bayesian Regression Models, *Canadian Journal of Statistics*, 29, 129–150.