تحلیل پارامتری مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی تحت تابع زیان هابر نوع دو: رهیافت بیز سلسله‌مراتبی

افشین فلاح[1]، منیر میرزایی

گروه آمار، دانشگاه بین‌المللی امام خمینی

تاریخ دریافت: ۱۳۹۵/۵/۳۱ تاریخ پذیرش: ۱۳۹۶/۱۱/۲۸

چکیده: مدل رگرسیون چندکی و تعمیم‌های آن، از جمله مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، به‌صورت سنتی به شیوه ناپارامتری تحلیل و پارامترهای آن‌ها به کمک برخی الگوریتم‌های تکراری بهینه‌سازی برآورد می‌شوند. به همین دلیل، در این مدل‌ها برای ساختن بازه‌های اطمینان و انجام آزمون فرض‌ها به‌ناچار از روش‌های مبتنی بر رتبه مشاهدات یا خودگردانی، استفاده می‌شود. در این مقاله، تحلیل پارامتری مدل رگرسیونی چندکی از نوع ماکسیمم درستنمایی مدنظر قرار گرفته است. نشان داده شده است که رهیافت فراوانی‌گرای مبتنی بر برآورد ماکسیمم درستنمایی، نتایجی منطبق بر روش ناپارامتری دارد. ازاین‌رو، برای دستیابی به مدلی کاراتر، از نظریه بیز استفاده شده و یک مدل بیز سلسله‌مراتبی توسعه داده شده است. کارایی مدل پیشنهادی در قالب یک مطالعه شبیه‌سازی و یک مثال کاربردی مورد ارزیابی قرار گرفته و با مدل رقیب مقایسه شده است. نتایج نشان می‌دهند که رهیافت بیزی تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی، به ازای همه اندازه‌های نمونه‌ای در مقایسه با مدل رقیب فراوانی‌گرا کاراتر است. بعلاوه، مدل پیشنهادی به‌خوبی تأثیر وجود نقاط دورافتاده در مشاهدات را که موجب چولگی توزیع متغیر پاسخ می‌شوند، در مدل‌سازی لحاظ می‌کند.

واژه‌های کلیدی: رگرسیون چندکی، بیز سلسله‌مراتبی، برآوردگر از نوع ماکسیمم درستنمایی.

رده‌بندی ریاضی (۲۰۱۰): ۶۲J۰۲، ۶۲M۱۰.

۱- مقدمه

در روش‌های معمول تحلیل رگرسیونی، میانگین متغیر پاسخ یا تابعی از آن به‌صورت شرطی بر مجموعه‌ای از متغیرهای تبیینی مدل‌بندی می‌شود. این در حالی است که میانگین، به‌عنوان یکی از معیارهای تمرکز، به‌تنهایی نمی‌تواند توزیع متغیر پاسخ را در فرایند مدل‌سازی تبیین نماید.

۱- آدرس الکترویکی نویسنده مسئول مقاله: a.fallah@ikiu.ac.ir

بر این اساس، این نوع تحلیل رگرسیونی نمی‌تواند درباره‌ی شکل توزیع متغیر پاسخ در سطوح مختلف متغیر تبیینی چندان آگاهی‌بخش باشد. چندک‌ها معیارهایی هستند که در کنار هم می‌توانند شکل توزیع را به‌صورت جامع‌تری به تصویر بکشند. برای مثال، اگر دهک‌های یک توزیع دارای فاصله‌ی تقریباً برابری از یکدیگر باشند، انتظار می‌رود توزیع نسبتاً هموار یا یکنواخت باشد و اگر دهک‌های بالایی دارای فاصله‌ی زیاد و دهک‌های پایینی دارای فاصله‌ی کمی از یکدیگر باشند، توزیع به سمت راست چوله خواهد بود. از طرفی، هنگامی‌که داده‌های دورافتاده‌ای وجود داشته باشد، علاوه بر میانگین شرطی متغیر پاسخ، رفتار مقادیر کرانگین این متغیر نیز از اهمیت زیادی برخوردار است. در این‌صورت به دلیل تأثیرپذیری میانگین از نقاط دورافتاده، یک مدل رگرسیون معمولی برای مدل‌سازی روابط بین متغیر پاسخ و متغیرهای تبیینی کفایت نمی‌کند و استفاده از روش کمترین توان‌های دوم خطا برای برآورد پارامترهای مدل مناسب نیست. یک راه برای کاهش اثر این‌گونه داده‌ها استفاده از معیار قدرمطلق خطا است. برخلاف روش کمترین توان‌های دوم خطا، روش کمترین قدرمطلق خطا نسبت به داده‌های دور افتاده استوار است [۱]. به همین دلیل، هاگ [۲] تحلیل رگرسیون میانه را که در آن برآورد پارامترهای مدل رگرسیونی با هدف مینیمم ساختن مجموع قدرمطلق خطاها صورت می‌پذیرد، برای مدل‌سازی داده‌هایی که دارای نقاط دورافتاده هستند، پیشنهاد نمود. رگرسیون چندکی را می‌توان تعمیمی از رگرسیون میانه دانست که در آن به‌جای مدل‌بندی میانه (چارک دوم) متغیر پاسخ، چندک‌های این متغیر به‌صورت شرطی بر مجموعه‌ای از متغیرهای تبیینی مدل‌بندی می‌شوند [۳]. جنبه‌های مختلف نظری و کاربردی تحلیل رگرسیون چندکی در منابع [۴]، [۵]، [۶]، [۷]، [۸] و [۹] مورد مطالعه قرار گرفته است. برکلینگ و چمبرز [۱۰] ایده‌ی تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی را مطرح ساختند که مبتنی بر برآوردگرهایی از نوع ماکسیمم درستنمایی هستند و بر اساس یک تابع تأثیر تعریف می‌شوند. تزاویدیس و همکاران [۱۱] تحلیل طولی با استفاده از مدل‌های رگرسیون چندکی از نوع ماکسیمم درستنمایی و الفو و همکاران [۱۲] مدل‌های رگرسیونی آمیخته متناهی چندکی از نوع ماکسیمم درستنمایی را مورد مطالعه قرار داده‌اند. چمبرز و تزاویدیس [۱۳] مدل‌های رگرسیون چندکی از نوع ماکسیمم درستنمایی را برای برآورد کوچک ناحیه‌ای مورد استفاده قرار داده و نشان دادند که انتخاب تابع تأثیر تنها ممکن است بر زمان همگرایی الگوریتم‌های عددی مؤثر باشد و بر دقت برآوردهای حاصل مؤثر نیست. آن‌ها علاوه بر تابع تأثیر هابر نوع دو، تابع تأثیر همپل و تابع تأثیر دو مربعی را نیز مدنظر قرار داده و نشان دادند که وقتی از تابع تأثیر هابر استفاده می‌شود، الگوریتم کمترین توان‌های دوم باز وزنی شده تکراری به یک جواب یکتا همگرا می‌شود. اما وقتی از تابع تأثیر همپل استفاده می‌شود، برخی مشکلات همگرایی بروز می‌کنند و در صورت استفاده از تابع تأثیر دو مربعی، مشکلات همگرایی نمود بیشتری خواهند داشت. ازآنجاکه علاوه بر مشکلات همگرایی، کلیه روش‌های موردبحث ناپارامتری هستند و توزیع خاصی را برای متغیر پاسخ در نظر نمی‌گیرند. به‌صورت

طبیعی انتظار می‌رود که با استفاده از روش‌های پارامتری بتوان برآوردهای مطلوب‌تری برای پارامترهای مدل به‌دست آورد و کارایی مدل را افزایش داد. در این مقاله، تحلیل پارامتری مدل رگرسیونی چندکی از نوع ماکسیمم درستنمایی مدنظر قرار گرفته است. نشان داده شده است که رهیافت فراوانی‌گرای مبتنی بر برآورد ماکسیمم درستنمایی، نتایجی منطبق بر روش ناپارامتری دارد. به همین دلیل، برای دستیابی به مدلی کاراتر، از نظریه بیز استفاده شده و یک مدل بیزی سلسله‌مراتبی توسعه داده شده است. نشان داده شده است که به‌دلیل شکل پیچیده تابع درستنمایی مبتنی بر مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، توزیع پسین فاقد صورت بسته بوده و شکل پیچیده‌ای دارد. ازاین‌رو، توزیع‌های پسین شرطی کامل پارامترها محاسبه و الگوریتم گیبز مناسب برای نمونه‌گیری از توزیع پسین، توسعه داده شده است.

در بخش ۲ مدل رگرسیون چندکی و در بخش ۳ مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، به‌طور اجمالی معرفی و نحوه برآورد پارامترهای این مدل‌ها موردبحث قرار گرفته است. در بخش ۴ رویکرد پیشنهادی برای تحلیل پارامتری رگرسیون چندکی از نوع ماکسیمم درستنمایی، معرفی و از دو دیدگاه فراوانی‌گرا و بیزی، موردبحث قرار گرفته است. در بخش‌های ۵ و ۶ نحوه کاربست رهیافت بیزی پیشنهادی در قالب یک مطالعه شبیه‌سازی و یک مثال کاربردی شرح داده شده و با مدل رقیب مقایسه شده است.

۲- تحلیل رگرسیون چندکی

در تحلیل رگرسیونی معمول، گشتاور اول توزیع متغیر پاسخ یا تابعی از آن به‌صورت شرطی بر مجموعه‌ای از متغیرهای تبیینی مدل‌بندی می‌شود. این در حالی است که گشتاورهای یک توزیع نمی‌توانند آن توزیع را به‌صورت یکتا مشخص سازند. هم‌چنین روش‌های رگرسیونی مبتنی بر مدل‌بندی میانگین پاسخ، در مواجهه به مشاهدات دورافتاده به‌شدت آسیب‌پذیر هستند. ازاین‌رو، یافتن روش‌های رگرسیونی جایگزین برای رگرسیون میانگین همواره موردتوجه پژوهشگران بوده است. تحلیل رگرسیون میانه، تحلیل رگرسیون چندکی، و تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی، از جمله روش‌های جایگزین در این زمینه هستند. در تحلیل رگرسیون چندکی، مجموعه‌ای از چندک‌ها برای تبیین دقیق‌تر توزیع متغیر پاسخ به‌کار گرفته می‌شوند. بعلاوه، چندک‌ها و به‌طور خاص میانه، در برابر مشاهدات دورافتاده استوارند. فرض کنید y_i و $\boldsymbol{x}'_i=\left(x_1,\ldots,x_p\right)$ به ترتیب متغیر پاسخ و بردار متغیرهای تبیینی متناظر با مشاهده iام را نشان می‌دهند. در تحلیل رگرسیون میانه، تابع زیان به‌صورت

$$\rho_{\cdot/\text{۵}}\left(\mathrm{y_i}-\mathbf{x}'_\mathrm{i}\boldsymbol{\beta}\right)=\cdot/\text{۵}\left|\mathrm{y_i}-\mathbf{x}'_\mathrm{i}\boldsymbol{\beta}\right|$$

در نظر گرفته می‌شود و برآورد بردار ضرایب رگرسیونی، $\beta' = (\beta_1,...,\beta_p)$، از کمینه ساختن کمیت

$$\sum_{i=1}^{n} \rho_{0/5}\left(y_i - x_i'\beta\right) = \sum_{i=1}^{n} 0/5\left|y_i - x_i'\beta\right|,$$

نسبت به β به‌دست می‌آید. تابع زیان رگرسیون میانه را می‌توان به‌صورت

$$\rho_{0/5}\left(u_i\right) = 0/5 u_i I_{[0,\infty)}\left(u_i\right) - \left(1-0/5\right) u_i I_{(-\infty,0)}\left(u_i\right), \qquad (1)$$

نوشت، که در آن

$$I_A\left(u\right) = \begin{cases} 0 & u \in A \\ 1 & \text{در غیر این‌صورت} \end{cases}$$

تابع نشانگر مجموعه‌ی A است و $u = y - \mathbf{x}'\beta$. تابع زیان مورد استفاده در رگرسیون میانه در رابطه‌ی (۱) را می‌توان با جایگزینی۰/۵با $q \in (0,1)$ به‌صورت

$$\rho_q\left(u_i\right) = q u_i I_{[0,\infty)}\left(u_i\right) - \left(1-q\right) u_i I_{(-\infty,0)}\left(u_i\right),$$

تعمیم داد. اگر چندک qام متغیر تصادفی y_i با $Q_q(y_i \mid \mathbf{x}_i)$ نشان داده شود، در این صورت مدل رگرسیون چندکی متناظر با این چندک به‌صورت

$$Q_q(y_i \mid \boldsymbol{x}_i) = \boldsymbol{x}'_i \boldsymbol{\beta}_q$$

است، که در آن $\beta_q' = \left(\beta_{0q}, \beta_{1q}, \ldots, \beta_{pq}\right)$ بردار ضرایب رگرسیونی را نشان می‌دهد. به ازای یک مقدار مشخص $q \in (0,1)$این بردار با استفاده از روش کمترین قدرمطلق خطا و از طریق کمینه ساختن تابع زیان قدرمطلق خطای وزنی

$$\begin{aligned} Q_q\left(\beta_q\right) &= \sum_i w_i\left(q\right)\left|y_i - \boldsymbol{x}'_i\boldsymbol{\beta}_q\right| \\ &= \sum_{i=1}^{n}\left|y_i - \boldsymbol{x}'_i\boldsymbol{\beta}_q\right|\left\{\left(1-q\right)\mathrm{I}\left(y_i - \boldsymbol{x}'_i\boldsymbol{\beta}_q \le 0\right) + q\,\mathrm{I}\left(y_i - \boldsymbol{x}'_i\boldsymbol{\beta}_q > 0\right)\right\}, \end{aligned} \qquad (2)$$

نسبت به β برآورد می‌شود، که در آن

$$w_i(q) = \begin{cases} q & y_i > \boldsymbol{x}'_i\boldsymbol{\beta}_q \\ 1-q & y_i \le \boldsymbol{x}'_i\boldsymbol{\beta}_q \end{cases},$$

محاسبه جواب‌های این مسئله مینیمم‌سازی، نیازمند کاربست روش‌های برنامه‌نویسی خطی است. چون الگوریتم‌های مورد استفاده برای برازش مدل‌های رگرسیون چندکی لزوماً همگرایی و جواب

یکتا را تضمین نمی‌کنند، برکلینگ و چمبرز [۱۰] ایده‌ی رگرسیون چندکی از نوع ماکسیمم درستنمایی را مطرح ساختند که در ادامه موردبحث قرار می‌گیرد.

۳- تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی

تحلیل رگرسیون چندکی از نوع درستنمایی، به‌عنوان تعمیمی از رگرسیون چندکی، مبتنی بر برآوردگرهایی از نوع ماکسیمم درستنمایی است و بر پایه یک تابع تأثیر تعریف می‌شود. اگر $\rho(\theta;T_n)$ یک تابع زیان باشد، $\psi(\theta;T_n)=\frac{\partial}{\partial\theta}\rho(\theta;T_n)$ را تابع تأثیر متناظر با آن گویند. بر این اساس، هر برآوردگر T_n که از حل یک مسئله کمینه‌سازی به شکل

$$\arg min_{\theta\epsilon\Theta}\sum_{i=0}^{n}\rho\left(\boldsymbol{\theta},T_n\right)\ ,$$

حاصل شود، یا به‌عنوان جواب صریح معادله

$$\sum_{i=1}^{n}\psi\left(\theta,T_n\right)=0$$

تعریف شود، یک برآوردگر از نوع ماکسیمم درستنمایی نامیده می‌شود. وجه تسمیه این نوع برآوردگر آن است که اگر قرینه تابع لگ‌درستنمایی به‌صورت

$$\rho\left(\boldsymbol{\theta};T_n\right)=-\log f\ \left(\boldsymbol{\theta};T_n\right).$$

به‌عنوان تابع زیان مدنظر قرار گیرد، برآوردگر ماکسیمم درستنمایی معمولی حاصل می‌شود. برآوردگرهای از نوع ماکسیمم درستنمایی، رده‌ی گسترده‌ای از برآوردگرها را تشکیل می‌دهند و هر برآوردگر از نوع ماکسیمم درستنمایی را می‌توان هم برحسب یک تابع زیان و هم برحسب تابع تأثیر متناظر با آن تعریف کرد. این نوع برآوردگر درصورتی‌که تابع تأثیر کران‌دار باشد، در برابر نقاط دور افتاده استوار خواهد بود. در اینجا از ذکر جزئیات بیشتر درباره این نوع برآوردگرها خودداری کرده و خواننده را به هابر [۱] ارجاع می‌دهیم.

در تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی، به‌جای میانگین پاسخ، چندک‌های شرطی متغیر پاسخ در سطوح مختلف متغیرهای تبیینی بر پایه یک تابع تأثیر، مبنای مدل‌سازی قرار می‌گیرند. در حالت کلی، یک مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی به‌صورت

$$MQ_q=\boldsymbol{x}'\boldsymbol{\beta}_{\psi}\left(q\right).$$

است، به ازای یک q مشخص و تابع تأثیر از پیش انتخاب شده‌ی ψ، برآوردگر ضرایب رگرسیون چندکی از نوع ماکسیمم درستنمایی از حل معادله‌ی

$$\sum_{i=1}^{n}\psi_q\left(r_{iq\psi}\right)\boldsymbol{x}_i=0, \tag{۳}$$

به‌دست می‌آیند، که در آن

$$\begin{aligned} r_{iq\psi}&=y_i-\boldsymbol{x}_i'\boldsymbol{\beta}_\psi(q),\\ \psi_q\left(r_{iq\psi}\right)&=2\psi\left(s^{-1}r_{iq\psi}\right)\left\{qI_{[0,\infty)}\left(r_{iq\psi}\right)+(1-q)I_{(-\infty,0)}\left(r_{iq\psi}\right)\right\}\\ &=\begin{cases}2q\psi(s^{-1}r_{iq\psi}) & r_{iq\psi}\ge 0\\ 2(1-q)\psi(s^{-1}r_{iq\psi}) & r_{iq\psi}<0\end{cases}, \end{aligned} \tag{۴}$$

و s یک برآوردگر مناسب مقیاس، مانند میانگین قدرمطلق خطای نمونه‌ای است [۱۳]. گرچه هر تابع تأثیری را می‌توان در معادله‌ی (۳) به‌کار بست، اما در این پژوهش به پیروی از [۱۳] تابع تأثیر هابر نوع دو مورد استفاده قرار گرفته‌است. این تابع تأثیر به‌صورت

$$\begin{aligned} \psi(u)&=uI_{(-\infty,c]}\left(|u|\right)+c\,sgn\left(u\right)\\ &=\begin{cases}u & |u|\le c\\ \mathrm{sgn}\,(u) & |u|>c,\end{cases} \end{aligned} \tag{۵}$$

است، که در آن c کمیتی مثبت بوده و q ضریب چندکی نامیده می‌شود. در مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی تحت تابع زیان هابر نوع دو به ازای $c\to\infty$ مدل رگرسیون چندکی و در حالت مدل رگرسیون انتظاره به‌دست می‌آید [۷]. چمبرز و تزاویدیس [۱۳] نشان دادند که انتخاب تابع تأثیر تنها ممکن است بر زمان همگرایی الگوریتم‌های عددی مؤثر باشد و بر دقت برآوردهای حاصله مؤثر نیست. آن‌ها علاوه بر تابع تأثیر هابر نوع دو، تابع تأثیر همپل و تابع تأثیر دومربعی را نیز مدنظر قرار داده و نشان دادند که وقتی از تابع تأثیر هابر نوع دو استفاده می‌شود، الگوریتم کمترین توان‌های دوم باز وزنی‌شده به یک جواب یکتا همگرا می‌شود. اما وقتی از تابع تأثیر همپل استفاده می‌شود، برخی مشکلات همگرایی بروز می‌کنند و در صورت استفاده از تابع تأثیر دو مربعی، مشکلات همگرایی نمود بیشتری خواهند داشت. ازآنجاکه همه روش‌های موردبحث ناپارامتری هستند و توزیع خاصی را برای متغیر پاسخ در نظر نمی‌گیرند. به‌صورت طبیعی انتظار می‌رود که با استفاده از روش‌های پارامتری بتوان به مدلی کاراتر دست یافت.

۴- مدل پیشنهادی: رهیافت پارامتری

در این بخش نظریه لازم برای تحلیل پارامتری رگرسیون چندکی از نوع ماکسیمم درستنمایی، توسعه داده می‌شود. بدیهی است برای هرگونه استنباط پارامتری، فراوانی‌گرا یا بیزی، شناخت توزیع متغیر پاسخ تحت مدل یاد شده ضروری است. بر این اساس، ابتدا توزیع متغیر پاسخ را در مدل پیشنهادی به‌دست می‌آوریم.

یک مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی به‌صورت

$$y = x'\beta_{\psi}(q) + u_q,$$

با تابع زیان

$$\rho_q(u) = ۲\rho\left(\frac{u}{s}\right)\left(qI_{[۰,\infty)}(u) + (۱-q)I_{(-\infty,۰)}(u)\right),$$

را در نظر بگیرید، که در آن

$$\rho(u) = \begin{cases} \dfrac{u^۲}{۲} & |u| \le c \\ c|u| - \dfrac{u^۲}{۲} & |u| > c \end{cases}$$

نشان می‌دهیم که متغیر پاسخ تحت این شرایط دارای یک توزیع احتمالاتی پیوسته با تابع چگالی

$$f_{Y|X}(y; x, \beta, q) = \frac{۱}{sB(q)}\exp\left(-\rho_q\left(\frac{y - x'\beta}{s}\right)\right), \qquad y \in \mathbb{R} \qquad (۶)$$

است، که در آن s یک پارامتر مقیاس و $B(q)$ ثابت چگالی‌ساز است. برای این منظور، ثابت چگالی‌ساز $B(q)$ در رابطه (۶) با تغییر متغیر $u = \dfrac{y - \mathbf{x}'\beta}{s}$ به‌صورت

$$
\begin{aligned}
B(q) &= \int_{-\infty}^{\infty} \exp\left(-\rho_q(u)\right) du \\
&= \int_{-\infty}^{-c} \exp\left\{-\left[c(-u) - c^{-2}/2\right](1-q)\right\} du \\
&+ \int_{-c}^{0} \exp\{-[\frac{u^2}{2} 2(1-q)\} du + \int_{0}^{c} \exp\left\{-\left[u^2/2\right]2q\right\} du \\
&+ \int_{-\infty}^{-c} \exp\left\{-\left[c(-u) - c^{-2}/2\right](q)\right\} du \\
&= \int_{-\infty}^{-c} \exp\{[2cu + c^2](1-q)\} du + \int_{-c}^{0} \exp\left\{-\left[u^2\right](1-q)\right\} du \\
&+ \int_{0}^{c} \exp\left\{-[u^2]q\right\} du + \int_{c}^{+\infty} \exp\left\{[-2c(u) + c^2]q\right\} du \\
&= \sqrt{\frac{\pi}{q}}\left[\phi\left(c\sqrt{2q}\right) - 1/2\right] + \sqrt{\frac{\pi}{1-q}}\left[\phi\left(c\sqrt{2(1-q)}\right) - 1/2\right] \\
&+ \frac{1}{2cq}\exp\left\{-c^2 q\right\} + \frac{1}{2c(1-q)}\exp\left\{-c^2(1-q)\right\},
\end{aligned}
$$

محاسبه می‌شود. هم‌چنین، تابع مولد گشتاورهای متناظر با توزیع (۶) به‌صورت

$$
\begin{aligned}
M_u(t) &= \frac{1}{B_q\left[2c(1-q) + t\right]}\exp\left\{-c^2(1-q) - ct\right\} \\
&+ \frac{\exp\left\{\frac{t^2}{4(1-q)}\right\}}{B_q}\sqrt{\frac{\pi}{(1-q)}}\left[\Phi\left(-\frac{t}{\sqrt{2(1-q)}}\right) - \Phi\left(-\frac{-2c(1-q) - t}{\sqrt{2(1-q)}}\right)\right] \\
&+ \frac{\exp\left\{\frac{t^2}{4q}\right\}}{B_q}\sqrt{\frac{\pi}{q}}\left[\Phi\left(\frac{2cq - t}{\sqrt{2q}}\right) - \Phi\left(\frac{-t}{\sqrt{2q}}\right)\right] \\
&- \frac{1}{B_q(t - 2cq)}\exp\left\{-c^2 q + ct\right\},
\end{aligned}
$$

قابل محاسبه است. گشتاورهای توزیع را می‌توان با استفاده از تابع مولد گشتاورها به‌دست آورد. اکنون با در اختیار داشتن توزیع متغیر پاسخ (۶) امکان انجام استنباط‌های پارامتری، فراوانی‌گرا یا بیزی، درباره پارامترهای مدل وجود دارد.

۴-۱-رهیافت فراوانی‌گرا

مشاهدات $\mathbf{D}=\{(y_1,x_1),\dots,(y_n,x_n)\}$ را که در آن x_i و y_i به‌ترتیب نشان دهنده مقادیر متغیرهای تبیینی و متغیر پاسخ حاصل از یک نمونه تصادفی هستند، در نظر بگیرید. توابع درستنمایی و لگ‌درستنمایی متناظر با مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، بر پایه تابع زیان هابر نوع دو به ترتیب به‌شکل

$$\begin{aligned}
L(\beta,q) &= \prod_{i=1}^{n} f_{Y_i|x_i}(y_i|\mathbf{x},\beta,q)\\
&= \prod_{i=1}^{n}\left\{\frac{1}{s\,B(q)}\exp\left(-\rho_q\left(\frac{y_i-\mathbf{x}_i'\beta}{s}\right)\right)\right\}\\
&= s^{-n}B(q)^{-n}\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i-\mathbf{x}_i'\beta}{s}\right)\right),
\end{aligned} \tag{۷}$$

و

$$\ell(\boldsymbol{\beta},q) = -n\log s - n\log B(q) - \sum_{i=1}^{n}\rho_q\left(\frac{y_i-\mathbf{x}_i'\boldsymbol{\beta}}{s}\right), \tag{۸}$$

هستند. به ازای یک q مشخص، ماکسیمم ساختن تابع لگ‌درستنمایی (۸)، معادل مینیمم ساختن عبارت $\sum_{i=1}^{n}\rho_q(\frac{y_i-\mathbf{x}_i'\beta}{s})$ در تابع نمایی موجود در سمت راست رابطه‌ی (۸) با استفاده از روش کمترین توان‌های دوم وزنی است. بنابراین، برآوردگرهای حاصل از روش ماکسیمم درستنمایی به‌عنوان یک روش پارامتری، و روش کمترین توان‌های دوم وزنی به‌عنوان یک روش ناپارامتری، بر هم منطبق‌اند. این بدان معنی است که استفاده از اطلاعات توزیع متغیر پاسخ از طریق رهیافت فراوانی‌گرا، نمی‌تواند در بهبود نتایج تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی مفید واقع شود. از طرفی چنانچه در منابع کلاسیک استنباط آماری بحث می‌شود، خواص بهینگی برآوردگرهای حاصل از روش ماکسیمم درستنمایی نظیر نااریبی، کارایی و نتایج نظریه والد، تنها در حالت مجانبی بروز پیدا می‌کنند و برای نمونه‌های کوچک این برآوردگرها می‌توانند بسیار دور از بهینگی باشند. همچنین، چنانچه نشان داده شد، برآوردگرهای ماکسیمم درستنمایی و کمترین توان‌های دوم وزنی هر دو از حل یک مسئله بهینه‌سازی به‌دست می‌آیند. بنابراین آن دسته از مشکلات بهینه‌سازی که توسط پژوهشگران سرشناس این حوزه تحقیقاتی، از جمله چمبرز و تزاویدیس [۱۳] برای روش کمترین توان‌های دوم وزنی موردبحث قرار گرفته‌اند، برای روش ماکسیمم درستنمایی نیز وجود دارند. به‌دلیل همین دشواری‌ها است

که در رهیافت ناپارامتری نیز عمدتاً از فواصل اطمینان مبتنی بر رتبه مشاهدات یا فواصل اطمینان خودگردانی، استفاده می‌شود.

۴-۲- رهیافت بیز سلسله‌مراتبی

در این بخش تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی، از دیدگاه بیزی مورد بررسی قرار گرفته و یک مدل بیزی برای آن توسعه داده شده است. لازمه تحلیل بیزی مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی آن است که پارامترهای این مدل متغیرهایی تصادفی تلقی و برای آن‌ها توزیع‌های پیشین مناسب در نظر گرفته شود. این توزیع‌های پیشین می‌بایست بتوانند به خوبی باورهای پیشین تحلیل‌گر در خصوص اهمیت متغیرهای تبیینی و حضور یا عدم حضور آن‌ها در مدل را منعکس سازند. با توجه به این‌که پارامترهای q و β به ترتیب در بازه $(0,1)$ و فضای $\mathbb{R}^p$ تغییر می‌کنند، در مقاله حاضر، برای این پارامترها به ترتیب توزیع‌های پیشین بتا و نرمال چندمتغیره در نظر گرفته شده است. بسیاری از محققان، از جمله گلمن و همکاران [۱۴] استفاده از توزیع پیشین نرمال برای ضرایب مدل‌های رگرسیونی را در حالت کلی منطقی می‌دانند. بعلاوه، این دو خانواده از توزیع‌های پیشین به اندازه کافی غنی هستند تا بتوان انتظار داشت عضوهایی از آن خانواده‌ها بتواند به‌خوبی تغییرپذیری و عدم قطعیت پارامترهای یاد شده را تبیین نمایند. بر این اساس، توزیع‌های پیشین پارامترهای مدل به‌صورت

$$
\begin{array}{ll}
\beta \mid \theta,\Delta \sim N_p(\theta,\Delta), & \Delta = \mathrm{diag}(\delta_1,\ldots,\delta_p) \\
\theta \sim N_p(\mu_0,\Sigma_0), & \mu_0 \in \mathbb{R}^p \\
1/\delta_j \sim \mathrm{Gamma}(k_0,\ell_0), & k_0,\ell_0 \in \mathbb{R}, \quad j=1,\ldots,P \\
1/s \sim \mathrm{Gamma}(a_0,b_0), & a_0,b_0 \in \mathbb{R} \\
q \sim \mathrm{beta}(\eta,\xi), & \eta>0, \quad \xi>0 \\
\eta \sim E(\lambda_0), & \lambda_0>0 \\
\xi \sim E(\tau_0), & \tau_0>0
\end{array}
\tag{۹}
$$

در نظر گرفته شده است، که در آن‌ها θ، Δ، μ_0، Σ_0، k_0، ℓ_0، a_0، b_0، η، ξ، λ_0 و τ_0 ابرپارامترهای توزیع پیشین هستند. این توزیع‌های پیشین انتخاب‌های بی‌طرفانه‌ای محسوب می‌شوند که برخی از مزایای مدل‌سازی بیزی برپایه توزیع‌های پیشین مزدوج را نیز به‌همراه دارند. این انتخاب‌ها تحلیل بیزی را در حالتی مدنظر قرار می‌دهند که اطلاعات پیشین قابل‌توجه و قوی درباره پارامترهای مدل در دسترس تحلیل‌گر قرار ندارد. اگر اطلاعات پیشین قابل‌توجهی در دسترس باشند، می‌توان توزیع‌های پیشین آگاهی‌بخش‌تری را مدنظر قرار داد. در نظریه بیز

ابرپارامترهای توزیع‌های پیشین می‌بایست بر اساس باورهای پیشین تحلیل‌گر درباره عدم قطعیت پارامترهای مدل تعیین شوند. اگر چنین امکانی برای تحلیل‌گر وجود نداشته باشد، معمولاً از روش‌های بیز تجربی یا بیز سلسله مراتبی برای تعیین ابرپارامترها یا مدل‌بندی عدم قطعیت آن‌ها استفاده می‌شود، که در این مقاله از رویکرد بیز سلسله‌مراتبی استفاده شده است. به‌طورکلی، چنانچه در [۱۵] صفحه ۴۶۰ اشاره شده است، مهم‌ترین مزیت انتخاب رویکرد سلسله مراتبی برای مواجهه با ابرپارامترهای نامعلوم یک مدل بیزی، استوارساختن استنباط‌های پسینی بیزی است. در واقع در تحلیل بیزی سلسله مراتبی عدم قطعیت موجود در تعیین پارامترهای توزیع‌های پیشین توسط یک توزیع احتمالاتی به سطوح دیگر حواله شده و به‌این‌ترتیب، عدم قطعیت از سطحی به سطح دیگر کاسته و به‌تدریج تأثیر توزیع‌های پیشین انتخابی بر استنباط‌های پسینی کم می‌شود. ذکر این نکته ضروری است که اغلب در نظر گرفتن دو سطح سلسله‌مراتبی برای توزیع‌های پیشین در یک مدل سلسله مراتبی، کفایت می‌کند و به در نظر گرفتن سطوح بیشتر که محاسبات را به‌صورت قابل ملاحظه‌ای پیچیده می‌کند، نیازی نیست. با انجام برخی محاسبات، توزیع‌های پیشین بردارهای β و θ را می‌توان به ترتیب به‌صورت متناسب به‌شکل

$$\pi(\beta) \propto |\Delta|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\beta-\theta)' \Delta^{-1} (\beta-\theta)\right\}$$
$$\propto |\Delta|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\beta'\Delta^{-1}\beta - 2\beta'\Delta^{-1}\theta + \theta'\Delta^{-1}\theta\right)\right\}, \qquad \beta \in \mathbb{R}^P$$

و

$$\pi(\boldsymbol{\theta}) \propto |\Sigma_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\mu}_0)' \Sigma_0^{-1} (\boldsymbol{\theta}-\boldsymbol{\mu}_0)\right\}$$
$$\propto |\Sigma_0|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\theta}'\Sigma_0^{-1}\boldsymbol{\theta} - 2\boldsymbol{\theta}'\Sigma_0^{-1}\boldsymbol{\mu}_0\right)\right\}, \qquad \boldsymbol{\theta} \in \mathbb{R}^P,$$

بازنویسی نمود. اکنون با در نظر گرفتن تابع درستنمایی (۷) و توزیع‌های پیشین (۹)، توزیع پسین را می‌توان بر اساس قواعد بیز بازنویسی نمود. اکنون با در نظر گرفتن تابع درستنمایی (۷) و احتمال شرطی، به‌صورت

$$
\begin{aligned}
\pi\left(\beta,q,\theta,\Delta,\eta,\xi,s \mid D\right) &\propto s^{-n}B^{-n}(q)\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i - x_i'\beta}{s}\right)\right) \\
&\times \phi_p\left(\beta,\theta,\Delta\right)\times\phi_p\left(\theta,\mu_0,\Sigma_0\right)\times f_{IG}\left(\delta_j;k_0,l_0\right) \\
&\times f_{IG}\left(s;a_0,b_0\right)\times f_{beta}\left(q,\eta,\xi\right)\times f_{\exp}\left(\tau_0\right)\times f_{\exp}\left(\lambda_0\right), \\
&\propto s^{-n}B^{-n}\left(q\right)\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i - x_i'\beta}{s}\right)\right) \\
&\times\left|\Delta\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\beta'\Delta^{-1}\beta - 2\beta'\Delta^{-1}\theta + \theta'\Delta^{-1}\theta\right)\right\} \\
&\times q^{\eta-1}\left(1-q\right)^{\xi-1}\times\exp\left\{-\frac{1}{2}\left(\theta'\Sigma_0^{-1}\theta - 2\theta'\Sigma_0^{-1}\mu_0\right)\right\} \\
&\times\left(\prod_{j=1}^{p}\frac{1}{\delta_j}\right)^{k+1} e^{-\ell_0\sum_{j=1}^{p}\frac{1}{\delta_j}}\times\left(\frac{1}{s}\right)^{a_0+1} e^{-b_0/s}\times e^{-\lambda_0\eta}\times e^{-\tau_0\xi}
\end{aligned}
\qquad (۱۰)
$$

نوشت، که در آن $f_{IG}(\cdot)$، $f_{\exp}(\cdot)$ و $f_{beta}(\cdot)$ به ترتیب نشان‌دهنده توابع چگالی توزیع‌های گامای وارون، نمایی و بتا را نشان می‌دهند. ملاحظه می‌شود که به‌دلیل شکل پیچیده تابع درستنمایی مبتنی بر مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، توزیع پسین توأم پارامترها فاقد صورت بسته است. ازاین‌رو، در ادامه از الگوریتم‌های مونت‌کارلو زنجیر مارکفی برای استنباط‌های پسینی درباره پارامترها استفاده می‌شود.

۴-۲-۱-توزیع‌های پسین شرطی کامل

در این بخش الگوریتم گیبز مناسب برای نمونه‌گیری از توزیع پسین (۱۰) توسعه داده می‌شود. لازمه استفاده از الگوریتم گیبز، شناخت توزیع‌های پسین شرطی کامل پارامترها است. شکل متناسب این توزیع‌ها با انجام برخی محاسبات به‌صورت

$$\pi(\beta|others) \propto \exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i - x_i'\beta}{s}\right)\right)\times\exp\left\{-\frac{1}{2}\left(\beta\Delta^{-1}\beta - 2\beta'\Delta^{-1}\theta\right)\right\},$$

$$\pi(\theta|others) \propto \exp\left\{-\frac{1}{2}(-2\beta'\Delta^{-1}\theta + \theta'\Delta^{-1}\theta)\right\}\times\exp\left\{-\frac{1}{2}\left(\theta'\Sigma_0^{-1}\theta - 2\theta'\Sigma_0^{-1}\mu_0\right)\right\}$$

$$\propto \exp[\{-\frac{1}{2}(-2\beta'\Delta^{-1}\theta + \theta'\Delta^{-1}\theta)\} + \left\{-\frac{1}{2}\left(\theta'\Sigma_0^{-1}\theta - 2\theta'\Sigma_0^{-1}\mu_0\right)\right\}$$

$$= \exp\left\{-\frac{1}{2}\left[\theta'\left(\Sigma_0^{-1} + \Delta^{-1}\right)\theta - 2\left(\beta'\Delta^{-1} + \mu_0\Sigma_0^{-1}\right)\theta\right]\right\},$$

$$\pi(q|others) \propto B^{-n}(q)\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i - x_i'\beta}{s}\right)\right)\times q^{\eta_0 - 1}(1-q)^{\xi_0 - 1},$$

$$\pi\left(\delta_j|others\right) \propto \left(\prod_{j=1}^{p}\frac{1}{\delta_j}\right)^{k+1} e^{-\ell_0\sum_{j=1}^{p}\frac{1}{\delta_j}}$$

$$\times\exp\left\{-\frac{1}{2}\left(\sum_{j=1}^{p}\beta_j^2\delta_j^{-1} - 2\sum_{j=1}^{p}\beta_j\theta_j\delta_j^{-1} + \sum_{j=1}^{p}\theta_j^2\delta_j^{-1}\right)\right\}$$

$$\propto \left(\frac{1}{\delta_k}\right)^{k+1+\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\beta_k^2 - 2\beta_k\theta_k + \theta_k^2\right)\frac{1}{\delta_k}\right\}\times\exp\left\{-\ell_0\frac{1}{\delta_k}\right\}$$

$$= \left(\frac{1}{\delta_k}\right)^{k+1+\frac{1}{2}}\exp\left\{\left(\frac{-\left(\beta_k - \theta_k\right)^2}{2} + \ell_0\right)\frac{1}{\delta_k}\right\},$$

$$\pi(s|others) \propto s^{-n}\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i - x_i'\beta}{s}\right)\right)\times\left(\frac{1}{s}\right)^{a_0+1} e^{-b_0/s},$$

$$\pi(\eta|others) \propto q^{\eta-1}e^{-\lambda_0\eta}$$

$$\propto e^{\log q^{\eta-1}}e^{-\lambda_0\eta}$$

$$= e^{-\eta(\lambda_0 - \log(q))},$$

$$\pi(\xi|others) \propto (1-q)^{\xi-1}e^{-\tau_0\xi}$$

$$\propto e^{\log(1-q)^{\xi-1}}e^{-\tau_0\xi}$$

$$= e^{-\xi(\tau_0 - \log(1-q))}.$$

قابل محاسبه‌اند. بر این اساس، توزیع‌های پسین شرطی کامل پارامترهای θ، Δ، η و ξ دارای نمایش بسته بوده و به‌صورت

$$
\begin{aligned}
&\theta|others \sim N_p\left(\left(\Sigma_0^{-1}+\Delta^{-1}\right)^{-1}\left(\Delta^{-1}\beta'+\mu_0\Sigma_0^{-1}\right),\left(\Sigma_0^{-1}+\Delta^{-1}\right)^{-1}\right), \quad \theta\in\mathbb{R}^P\\
&\delta_j|others \sim IG\left(k_0+\frac{1}{2},\frac{\left(\beta_k-\theta_k\right)^2}{2}+\ell_0\right), \qquad \delta_j>0,\ j=1,\ldots,p \qquad (11)\\
&\eta|others \sim E\left(\lambda_0-\log q\right), \qquad \lambda_0>\log p\\
&\xi|others \sim E\left(\tau_0-\log(1-q)\right), \qquad \tau_0>\log(1-q),
\end{aligned}
$$

به‌دست می‌آیند. با در دست داشتن توزیع‌های پسین شرطی کامل پارامترها می‌توان از توزیع پسین توأم پارامترها به کمک الگوریتم گیبز نمونه‌گیری نمود. مراحل کلی الگوریتم گیبز برای نمونه‌گیری از توزیع پسین (۱۰) به شرح زیر است:

۱. مقدار اولیه‌ی $(\beta^{(0)},q^{(0)},\theta^{(0)},\Delta^{(0)},\eta^{(0)},\xi^{(0)})$ را برای پارامترهای مدل در نظر بگیرید.

۲. در تکرار $(t+1)$ام الگوریتم، براساس مقادیر حاصل از مرحله tام، پارامترها را به‌صورت زیر به‌روز کنید:

$$
\begin{aligned}
&\xi^{(t+1)}|q^{(t)}\sim E\left(\tau_0-\log(1-q)\right), \quad \tau_0>\log(1-q),\\
&\eta^{(t+1)}|q^{(t)}\sim E\left(\lambda_0-\log q\right), \qquad \lambda_0>\log p,\\
&\beta^{(t+1)}|\left(q,\theta,\Delta\right)^{(t)},\left(\xi,\eta\right)^{(t+1)}\propto\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i-x_i'\beta}{s}\right)\right)\times\exp\left\{-\frac{1}{2}\left(\beta'\Delta^{-1}\beta-2\beta'\Delta^{-1}\theta\right)\right\},\\
&\theta^{(t+1)}|\left(q,s,\Delta\right)^{(t)},\left(\beta,\xi,\eta\right)^{(t+1)}\sim N_p\left(\left(\Sigma_0^{-1}+\Delta^{-1}\right)^{-1}\left(\Delta^{-1}\beta'+\mu_0\Sigma_0^{-1}\right),\left(\Sigma_0^{-1}+\Delta^{-1}\right)^{-1}\right),\\
&\Delta^{(t+1)}|\left(q\right)^{(t)},\left(\beta,\xi,\eta,\theta\right)^{(t+1)}\sim IG\left(k_0+\frac{1}{2},\frac{\left(\beta_k-\theta_k\right)^2}{2}+\ell_0\right),\\
&s^{(t+1)}|\left(q\right)^{(t)},\left(\beta,\theta,\xi,\eta,\Delta\right)^{(t+1)}\propto s^{-n}\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i-x_i'\beta}{s}\right)\right)\times\left(\frac{1}{s}\right)^{a_0+1}e^{-b_0/s},\\
&q^{(t+1)}|\left(\beta,\theta,x,\eta,\Delta\right)^{(t+1)}\propto B^{-n}\left(q\right)\exp\left(-\sum_{i=1}^{n}\rho_q\left(\frac{y_i-x_i'\beta}{s}\right)\right)\times q^{\eta_0-1}\left(1-q\right)^{\xi_0-1},
\end{aligned}
$$

۳. مرحله ۲ را تا همگرا شدن الگوریتم تکرار کنید

همان‌گونه که مشاهده می‌شود، در مرحله ۲ الگوریتم گیبز، توزیع پسین شرطی کامل پارامترهای β، q و s، صورت بسته و شناخته شده‌ای ندارند. بنابراین از الگوریتم متروپولیس-هاستینگز برای نمونه‌گیری از این توزیع‌ها استفاده می‌شود. بر این اساس، برای نمونه‌گیری از توزیع پسین(۱۰)به تلفیقی از الگوریتم‌های گیبز و متروپولیس-هاستینگز نیاز است.

۵-مطالعه شبیه‌سازی

در این بخش مدل بیزی پیشنهادی در چارچوب یک مطالعه‌ی شبیه‌سازی مورد ارزیابی قرار گرفته و کارایی آن با مدل فراوانی‌گرای چندکی از نوع ماکسیمم درستنمایی مقایسه می‌شود. برای این منظور، یک مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی به‌شکل

$$MQ = x'\beta = \beta_{0\psi} + \beta_{1\psi}(q)x_1 + \beta_{2\psi}(q)x_2 . \qquad (12)$$

در نظر گرفته شده است. مقادیر متغیرهای تبیینی از توزیع نرمال شبیه‌سازی شده‌اند. سپس با فرض $\beta_0 = 1$، $\beta_1 = 2$ و $\beta_2 = 3$ بردار مشاهدات پاسخ با در نظر گرفتن تابع چگالی رابطه‌ی (۶) شبیه‌سازی شده‌اند.

به‌منظور لحاظ نمودن عدم قطعیت حاکم بر فرایند تولید نمونه‌های تصادفی، فرایند تولید نمونه‌های تصادفی و برآورد پارامترها۱۰۰۰بار تکرار و بر اساس آن‌ها مقادیر ریشه توان‌های دوم خطای برآوردگرهای حاصل از مدل‌های رقیب به ازای چندک‌های مختلف ۰/۲۵، ۰/۵۰ و ۰/۷۵، و اندازه‌های نمونه‌ای۱۰، ۲۰، ۴۰و ۸۰ محاسبه و در جدول ۱ ارائه شده است. چون رهیافت پیشنهادی بیزی است، از در نظر گرفتن اندازه نمونه‌های بزرگ‌تر پرهیز شده است، چرا که مزیت اصلی رهیافت بیزی عمدتاً مربوط به اندازه‌های کوچک نمونه‌ای است و برای نمونه‌های بزرگ به‌دلیل غلبه درستنمایی بر پیشین، نتایج دو رهیافت بیزی و فراوانی‌گرا مشابه خواهند بود. برای اطمینان از همگرایی زنجیرهای مارکف تولید شده به توزیع مانای متناظر خود، از آزمون گلمن و روبین [۱۶] استفاده شده است. برای این منظور، تابع *gelman.diag* در بسته کتابخانه‌ای *Coda* (پلومر و همکاران، [۱۷]) از نرم‌افزار R (گروه مرکزی R، [۱۸]) مورد استفاده قرار گرفته است. با توجه به مقادیر جذر میانگین توان‌های دوم خطای مربوط به ضرایب رگرسیونی، ملاحظه می‌شود که رهیافت بیزی تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی، به ازای همه اندازه‌های نمونه‌ای، در مقایسه با مدل رقیب فراوانی‌گرا به‌صورت قابل ملاحظه‌ای از کارایی مطلوب‌تری برخوردار است.

جدول (۱): مقادیر ریشه میانگین توان دوم خطای برآوردگرهای ضرایب رگرسیونی در مدل‌های فراوانی‌گرا و بیزی تحلیل رگرسیون چندکی از نوع ماکسیمم درستنمایی، برای اندازه نمونه۱۰، ۲۰، ۴۰، ۸۰ و به ازای چندک‌های ۰/۲۵، ۰/۵۰ و ۰/۷۵.

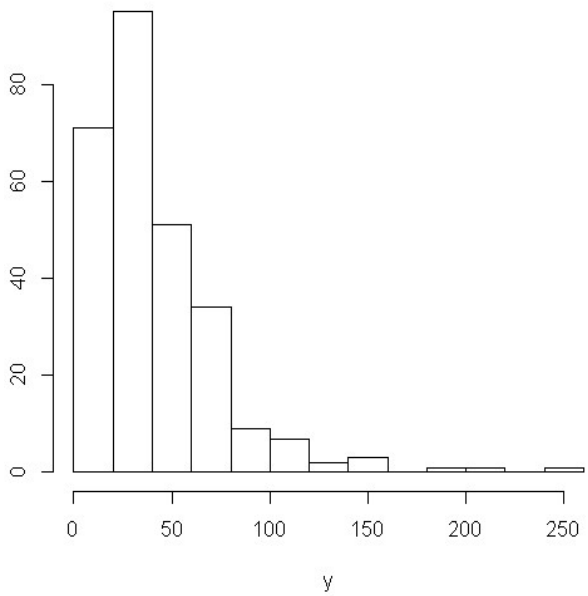
چندک	اندازه نمونه	پارامتر	مدل	
			بیزی	فراوانی‌گرا
۰/۲۵	۱۰	β_0	۰/۶۳۱۶	۳/۰۷۸۷
		β_1	۰/۱۴۳۲	۲/۰۶۳۵
		β_2	۰/۵۴۳۶	۳/۰۲۵۲
	۲۰	β_0	۰/۷۷۱۶	۱/۸۹۳۵
		β_1	۱/۵۶۷۳	۲/۰۰۰۴
		β_2	۱/۷۵۱۴	۳/۰۱۷۶
	۴۰	β_0	۱/۶۶۸۳	۲/۴۳۱۳
		β_1	۱/۸۷۴۹	۱/۹۹۸۳
		β_2	۲/۶۴۵۹	۳/۰۰۲
	۸۰	β_0	۰/۹۹۹۵	۱/۴۰۱۱
		β_1	۱/۹۶۶۳	۱/۰۰۲۳
		β_2	۲/۹۳۶۰	۲/۹۹۸۴
۰/۵	۱۰	β_0	۲/۰۹۷۳	۴/۰۹۱۴
		β_1	۰/۵۴۷۵	۲/۰۴۶۲
		β_2	۱/۵۰۹۰	۳/۰۰۴۷
	۲۰	β_0	۱/۳۵۳۴	۲/۰۱۰۹
		β_1	۱/۳۱۴۹	۲/۰۰۶۸
		β_2	۲/۰۰۰۷	۳/۰۰۲۲
	۴۰	β_0	۱/۲۸۱۴	۱/۲۰۸۳
		β_1	۱/۸۶۰۶	۲/۰۰۳۳
		β_2	۲/۸۱۰۵	۳/۰۰۳۵
	۸۰	β_0	۱/۰۱۳۷	۱/۱۸۵۱
		β_1	۱/۹۷۹۰	۲/۰۰۱۱
		β_2	۲/۹۳۳۹	۳/۰۰۰۷
۰/۷۵	۱۰	β_0	۰/۴۳۵۲	۰/۴۰۳۶
		β_1	۱/۲۷۹۰	۲/۰۰۹۲
		β_2	۲/۴۳۷۶	۳/۰۰۲۹
	۲۰	β_0	۳/۴۳۰۰	۳/۷۴۰۳
		β_1	۲/۰۰۲۱	۲/۰۱۱۵
		β_2	۲/۷۳۰۹	۳/۰۰۱۲
	۴۰	β_0	۱/۶۶۷۹	۱/۹۰۱۸
		β_1	۱/۸۷۱۶	۱/۹۹۹۱
		β_2	۲/۶۴۵۵	۳/۰۰۰۱
	۸۰	β_0	۲/۷۱۷۸	۲/۷۹۲۲
		β_1	۱/۹۸۹۲	۲/۰۰۸۳
		β_2	۱/۹۲۲۲	۲/۹۹۴۸

همان‌گونه که بر اساس ویژگی‌های مجانبی برآوردگرهای ماکسیمم درستنمایی انتظار می‌رود، با افزایش اندازه نمونه کارایی برآوردگرهای ماکسیمم درستنمایی افزایش می‌یابد، به‌گونه‌ای که برای اندازه نمونه‌های بزرگ تفاوت معنی‌داری بین نتایج حاصل از مدل‌های بیزی و فراوانی‌گرا وجود ندارد. لازم به ذکر است که چون در مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی توزیع متغیر پاسخ نامتقارن است، دقت مدل در برآورد پارامترهای مدل به ازای چندک‌های مختلف، یکسان نیست. در واقع بسته به اینکه توزیع متغیر پاسخ چوله به راست (چپ) باشد، دقت برآورد پارامترها به ازای چندک‌های بالایی (پایینی) کمتر خواهد بود. این یکی از مزایای مدل رگرسیونی چندکی از نوع ماکسیمم درستنمایی است که به خوبی تأثیر وجود نقاط دورافتاده در مشاهدات را که منجر به چولگی توزیع متغیر پاسخ می‌شوند، در مدل‌سازی لحاظ می‌کند. برنامه‌های نوشته‌شده برای تحلیل بیزی سلسله‌مراتبی مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، از طریق نشانی $https://sites.google.com/site/datasetandprograms$ در دسترس هستند.

۶- مثال کاربردی

در این بخش نحوه کاربست مدل بیزی پیشنهادی، در قالب یک مثال کاربردی شرح داده شده و کارایی آن مورد ارزیابی قرار گرفته است. برای این منظور، نمونه‌ای از مجموعه داده‌های مربوط به طرح آمارگیری هزینه و درامد خانوارهای شهری و روستایی مرکز آمار ایران، در سال ۱۳۸۸، مورد استفاده قرار گرفته است. طرح آمارگیری هزینه و درامد خانوار از سال۱۳۴۲در مناطق روستایی و از سال ۱۳۴۷در مناطق شهری به اجرا در آمده است. از سال۱۳۵۳علاوه بر هزینه‌های خانوار، اطلاعات درامد نیز گرداوری شده است. هدف کلی طرح آمارگیری از هزینه و درامد خانوار، برآورد میانگین هزینه‌ها و درامد خانوارهای شهری و روستایی در سطح کشور و استان‌ها است. در این طرح به شیوه سنتی میانگین متغیر پاسخ (درامد) مدل بندی می‌شود، حال آنکه میانگین نمی‌تواند توزیع را به‌صورت مناسبی تبیین نماید. ازاین‌رو، مدل رگرسیونی چندکی از نوع ماکسیمم درستنمایی که به‌جای میانگین، چندک‌های توزیع را مدنظر قرار می‌دهد، به‌عنوان مدل جایگزین پیشنهاد و بکار گرفته شده است. در این پژوهش، هزینه خالص خانوارهای روستایی به میلیون ریال به‌عنوان متغیر پاسخ، و بعد و تعداد افراد شاغل خانوار به‌عنوان متغیرهای تبیینی در نظر گرفته شده‌اند. گرچه همه متغیرهای تبیینی به‌کار گرفته شده کمّی هستند، اما اگر برخی متغیرهای تبیینی کیفی باشند نیز مدل به‌صورت مشابه قابل برازش است. در چنین شرایطی برای هر سطح از متغیرهای تبیینی کیفی، یک ضریب رگرسیون خاص آن سطح لحاظ و مدل حاصل به شکل یک مدل تحلیل کوواریانس برازش و تعبیر می‌شود. از الگوریتم گیبز شرح داده شده در بخش ۴-۲ برای نمونه‌گیری از توزیع‌های پسین شرطی کامل پارامترها استفاده شده

است. برآوردهای بیزی ضرایب مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، به همراه فواصل باورمند بیزی ۹۵٪ متناظر آن‌ها، در جدول ۲ ارائه شده‌اند. برآوردهای حاصل از مدل فراوانی‌گرا، به‌عنوان مدل رقیب، به‌همراه فواصل اطمینان خودگردان ۹۵٪ مربوطه نیز محاسبه و ارائه شده‌اند. محاسبات مربوط به مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی با استفاده بسته نرم‌افزاری *qrlm* انجام شده است.



شکل (۱): بافت نگار هزینه ۳۶ خانوار روستایی در ایران در سال۱۳۸۸.

جدول (۲): برآوردهای بیزی و ماکسیمم درستنمایی ضرایب مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی به همراه فواصل باورمند بیزی و بازه‌های اطمینان خودگردانی ۹۵٪ متناظر آن‌ها.

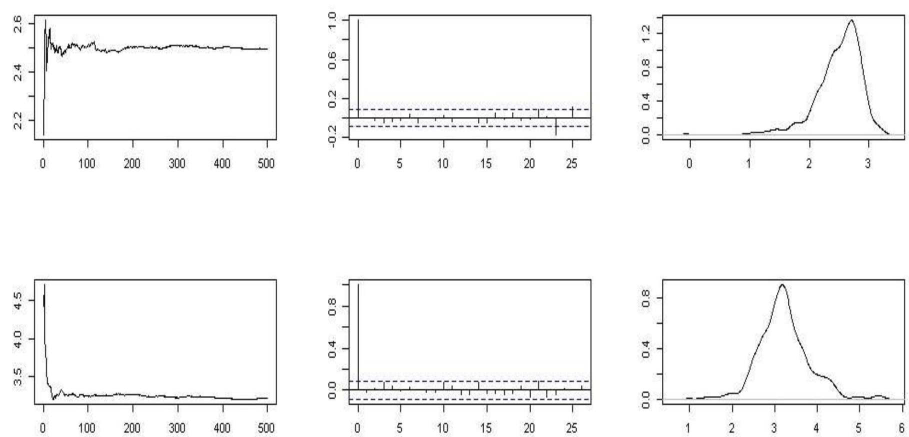
پارامتر	مدل	
	بیزی	فراوانی‌گرا
β_0	۰/۲۶۶۵ (-۲/۷۱ . ۳/۲۴)	۱۲/۹۱۵۷ (-۲۰/۷۹ . ۳۴/۹۳)
β_1	۴/۳۵۲۵ (۲/۳۳ . ۵/۳۳)	۸/۷۱۰۸ (۱/۶۵ . ۲۲/۸۹)
β_2	۴/۱۷۶۸ (۲/۱۴ . ۸/۷۴)	-۰/۳۵۴۹ (-۱۷/۹۷ . ۹/۷۰)

جدول (۳): برآوردهای بیزی و ماکسیمم درستنمایی ضرایب مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی بدون در نظر گرفتن عرض از مبدأ به همراه فواصل باورمند بیزی و بازه‌های اطمینان خودگردانی ٪۹۵ متناظر آن‌ها.

پارامتر	مدل	
	بیزی	فراوانی‌گرا
β_1	۲/۵۸۲۸ (۱/۵۱ ،۳/۰۱)	۱۲/۸۴۹۴ (۸/۱۳،۱۹/۴۷)
β_2	۳/۲۵۹۷ (۲/۰۷ ،۴/۴۵)	-۲/۸۷۹۸ (-۱۶/۲۳ ،۷/۵۶)

جدول (۴): مقادیر معیارهای اطلاع AIC، BIC، HQC و DIC برای مدل‌های بیزی و فراوانی‌گرا رگرسیون چندکی از نوع ماکسیمم درستنمایی، بدون در نظر گرفتن عرض از مبدأ.

معیار	مدل	
	بیزی	فراوانی‌گرا
AIC	۶۵/۳۱۶	۵۲/۳۳۳
BIC	۴۶/۳۱۹	۹۲/۳۳۴
HQC	۵۵/۳۱۷	۹۷/۳۳۳
DIC	۶۵/۳۱۲	۵۲/۳۳۱



شکل (۲): نمودارهای اثر (چپ)، خودهمبستگی(وسط) و چگالی(راست) زنجیرهای مارکف شبیه‌سازی شده از توزیع‌های پسین شرطی کامل پارامترهای ضرایب رگرسیونی β_1 (بالا) و β_2 (پایین) پس از طی شدن دوره داغیدن.

بر اساس فاصله اطمینان مربوطه به ضریب رگرسیونی $\beta_۲$ در مدل فراوانی‌گرا، حضور متغیر تبیینی $x_۲$ در مدل معنی‌دار نیست، درحالی‌که بر اساس فاصله باورمند بیزی متناظر با $\beta_۲$ این متغیر در سطح ۵٪ معنی‌دار است. بعلاوه، فواصل اطمینان حاصل از مدل فراوانی‌گرا در مقایسه با فواصل باورمند بیزی، طولانی‌تر هستند و از دقت کمتری برخوردارند. از طرفی با توجه به فواصل اطمینان حاصل از روش فراوانی‌گرا و فواصل باورمند حاصل از روش بیزی، ملاحظه می‌شود که پارامتر عرض مبدأ در هیچ‌یک از دو مدل معنی‌دار نیست. به همین دلیل این پارامتر از مدل حذف و مدل بدون عرض از مبدأ به مشاهدات برازش داده شده است. در مدل بدون عرض از مبدأ نیز ضریب رگرسیونی $\beta_۲$ در مدل فراوانی‌گرا، برخلاف مدل بیزی، معنی‌دار نیست. لازم به ذکر است که به‌منظور بررسی همگرایی زنجیرهای مارکف شبیه‌سازی شده به توزیع مانای متناظر خود، نمودارهای اثر، خودهمبستگی و چگالی زنجیرهای مارکف شبیه‌سازی شده از توزیع‌های پسین شرطی کامل ضرایب رگرسیونی، پس از طی شدن دوره داغیدن، در شکل ۲ رسم شده است. هرچند این نمودارها حاکی از همگرایی زنجیرهای مارکف شبیه‌سازی شده به توزیع مانای خود هستند، اما برای اطمینان از همگرایی زنجیرهای مارکف، در اینجا نیز از آزمون پیشنهادی گلمن و روبین [۱۶]استفاده شده است. برای این منظور، تابع *gelman.diag* در بسته کتابخانه‌ای *Coda* (پلومر و همکاران، [۱۷]) از نرم‌افزار R (گروه مرکزی R، [۱۸]) مورد استفاده قرار گرفته است. برای بررسی میزان برازش هر یک از مدل‌ها به مشاهدات موجود از معیارهای اطلاع آکائیک (AIC)، شوارتز (BIC)، هانان-کوئین (HQC) و DIC استفاده شده است. سه معیار اول به ترتیب به‌صورت

$$\begin{aligned} AIC &= -۲\ell(\hat{\theta}) + ۲k\,, \\ BIC &= -۲\ell(\hat{\theta}) + k\log n, \\ HQC &= -۲\ell(\hat{\theta}) + ۲k\,\log(\log(n)), \end{aligned}$$

محاسبه می‌شوند، که در آن‌ها k تعداد پارامترهای هر مدل و n اندازه نمونه را نشان می‌دهد. برای ملاحظه جزئیات مربوط به نحوه محاسبه و تعبیر معیار DIC، اشپیگل هارتر و همکاران [۱۹] را ببینید. جدول ۴ مقادیر این معیارهای اطلاع را برای مدل‌های بیزی و فراوانی‌گرا، نشان می‌دهد. همان‌گونه که ملاحظه می‌شود، مدل بیزی پیشنهادی در مقایسه با مدل فراوانی‌گرا برازش به‌صورت قابل‌توجه مناسب‌تری به مشاهدات دارد. این موضوع با توجه به این‌که بافت نگار مشاهدات پاسخ مربوط به هزینه ۳۰ خانوار روستایی که در شکل ۱ رسم شده است، نشان‌دهنده‌ی وجود نقاط دورافتاده در این مشاهدات و چولگی توزیع آن است. از دیدگاه نظری کاملاً قابل توجیه است. بعلاوه، برخلاف رویکردهای فراوانی‌گرای مبتنی بر روش ماکسیمم درستنمایی و کمترین توان‌های دوم وزنی، که به دلیل در اختیار نداشتن توزیع برآوردگرها قادر به ساختن

فواصل اطمینان پارامتری نیستند و به همین دلیل از فواصل اطمینان مبتنی بر روش‌های خودگردانی استفاده می‌کنند، از طریق نظریه بیزی به‌سادگی می‌توان برای پارامترهای مدل، فاصله‌های باورمند بیزی فراهم ساخت.

۷- بحث و نتیجه‌گیری

میانگین به‌عنوان یکی از معیارهای تمرکز به‌تنهایی نمی‌تواند توزیع یک متغیر را به‌صورت مناسبی تبیین نماید. ازاین‌رو، روش‌های مرسوم تحلیل رگرسیونی که میانگین متغیر پاسخ را مدل‌بندی می‌کنند، در بسیاری از کاربردها دارای کارایی لازم نیستند. تحلیل رگرسیونی چندکی از نوع ماکسیمم درستنمایی، به‌عنوان یکی از تعمیم‌های رگرسیون چندکی، مجموعه‌ای از چندک‌ها را بر پایه یک تابع تأثیر، برای تبیین تغییرات متغیر پاسخ مورد استفاده قرار می‌دهد. رویکرد سنتی در برازش این نوع مدل رگرسیونی ناپارامتری بوده و پارامترهای مدل را بدون استفاده از اطلاعات توزیع، برپایه برخی الگوریتم‌های تکراری برآورد می‌نماید. بعلاوه برای محاسبه فواصل اطمینان و انجام آزمون فرض درباره ضرایب رگرسیونی، به دلیل در اختیار نداشتن توزیع دقیق برآوردگرها از بازه‌های اطمینان خودگردانی و آزمون‌های ناپارامتری استفاده می‌کند. درحالی‌که در رهیافت بیزی می‌توان برای پارامترهای مدل، بازه‌های باورمند بیزی تدارک دید. ذکر این نکته ضروری است که متغیر پاسخ در مدل رگرسیون چندکی از نوع ماکسیمم درستنمایی، بر اساس تابع زیان هابر نوع دو به‌دست می‌آید. درصورتی‌که در مدل رگرسیون چندکی، توزیع متغیر پاسخ بر اساس تابع زیان قدرمطلق خطا به دست می‌آید. به عبارت روشن‌تر، وقتی ثابتcدر تابع زیان هابر نوع دو به سمت صفر میل می‌کند، مدل رگرسیون پیشنهادی در این مقاله به مدل رگرسیون چندکی و در حالتی کهcبه بینهایت میل می‌کند، به مدل رگرسیون انتظاره تبدیل می‌شود.

سپاس‌گزاری

نویسندگان از داوران گرامی برای ارائه پیشنهادهای سازنده، در راستای بهبود این پژوهش، سپاس‌گزارند.

منابع

[1] Huber, P. J. (2009). *Robust Statistics.* Wiley. New York.

[2] Hogg, R. V. (1975). Estimates of Precentile Regression Lines Using Salary Data. *Journal of the American Statistical Association*, **70**, 349, 56-59.

[3] Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica*, **46**, 33-50.

[4] Chamberlin, G. (1994). *Quantile Regression Censoring and Structure of Wages in Advances in Ecommetrics*, ed. by C.sims. *Elsevier*, New York.

[5] Koenker, R. and Bilias, Y. (2001). Quantile regression for Duration Data. *A Reapprasial of the Pennsylvania Re Employment Bouns Experiments Emprical Economics*, **26**, 199-230.

[6] Koenker, R. and Hallock, K. (2001). Qouantile regression: An Introduction for duration data. *Journal of Economics*, **26**, 15, 143-156.

[7] Koenker, R. and Hallock, K. (2005). *Qouantile regression*. *Cambridge University Press*, 26, 38.

[8] Koenker, R. and Dorey, V. (2016). Computing Regression Quantiles. *Journal of Applied Statistical Sciences*, **36**, 383-393.

[9] Hao, L. and Naiman D. Q. (2007). *Quantile Regression*. *Sage Publications*, Thousand Oaks, 515-533.

[10] Breckling, J. and Chambers, R. (1988). M-Quantiles. *Biometrika*, **75**, 761-771.

[11] Tzavidis, N., Salvati, N., Schmid, T., Flouri, E., and Midouhas, E. (2016). Longitudinal Analysis of the Strengths and Difficulties Questionnaire Scores of the Millennium Cohort Study Children in England Using M-quantile Random Effects Regression. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**, 427-452.

[12] Alfo, M., Salvati, N. and Ranallli, M. G. (2016). Finite Mixtures of Quantile and M-quantile Regression Models. *Statistics and Computing*, **27**, 547-570.

[13] Chambers, R. and Tzavidis, N. (2006). M-Quantiles Models for Small Area Estimation. *Biometrika*, 328, **93**, 255-268.

[14] Gelman, A., Jakulin, A. Grazia P. and Su, Y. (2006). Prior Distribution for Variance Parameters in Hierachical Models. *Bayesian Analysis*, **3**, 515-533.

[15] Robert, C. (2001). *Bayesian Choice*, Second Edition, Springer Verlag. New York.

[16] Gelman A. and Rubin D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, **7**, 457-511.

[17] Plummer, M., Best, N. Cowles, K. and Vines, K. (2006). CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, **6**, 1, 7-11.

[18] R Core Team (2013). *R: A Language and Environment for Statistical Computing. RFoundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[19] Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of Royal Statistical Society, Series B*, **59**, 731–792.

Paremetric Analysis of M-Quantile Regression Models Under Type 2 Huber Loss: Hierarchical Bayes Approach

Afshin Fallah and Monir Mirzaee

Department of Statistics. Imam Khomeini International University, Ghazvin, Iran

Abstract

Quantile regression model and its generalizations, including M-quantile regression model, are analyzed usually via a nonparametric approach and their parameters are estimated using some iterative optimization algorithms. For these reason, in these models confidence intervals and hypotheses testing have done perforce using rank-based or bootstrapping approaches. In this paper, we consider parametric analysis of M-quantile model. It is shown that, the frequentist based approach of maximum likelihood estimation leads to results that are similar to the nonparametric approach. Hence, in order to achieve a more afficient model, we have been used the Bayes theory and a hierarchical Bayes model has been developed. The efficiency of the proposed model has been assessed via a simulation study and real word example. The results show that the Bayesian approach of m-quantile regression analysis is more efficient than the correspond frequantist approach, for all sample sizes. In addition, the proposed model truly takes into account the effect of the outlier observation, which causes skewness in response variable distribution, in modeling.

Keywords: Quantile Regression, Hierarchical Bayes, Maximum Likelihood Type Estimator.

Mathematics Subject Classification (2010): 62J02, 62M10.