رگرسیون چندکی بیزی با تاوان لاسو سازوار برای داده‌های پانلی پویا

علی آقامحمدی،[۱] سکینه محمدی [۲]

تاریخ دریافت: ۱۳۹۳/۱۲/۲۳
تاریخ پذیرش: ۱۳۹۵/۱۲/۲۵

چکیده:

مدل‌های داده‌های پانلی پویا قسمت مهمی از مطالعات حوزه‌های پزشکی، اجتماعی و اقتصادی را شامل می‌شوند. ویژگی بارز این مدل‌ها وجود متغیر وابستهٔ تأخیری به‌عنوان متغیر توصیفی است. مشکل برآورد در این مدل‌ها از همبستگی بین متغیر وابستهٔ تأخیری و مؤلفهٔ خطای فعلی ناشی می‌شود. اخیراً رگرسیون چندکی تاوانیده برای تحلیل داده‌های پانلی پویا مورد توجه قرار گرفته است. در این مقاله نخست مدل رگرسیون چندکی با ایجاد تاوان لاسو سازوار روی اثرهای تصادفی برای داده‌های پانلی پویا با فرض وابستگی اثرهای تصادفی و مشاهدات اولیه ارائه می‌شود. همچنین این مدل با فرض استقلال بین اثرهای تصادفی و مشاهدات اولیه نیز بررسی خواهد شد. هر دو مدل از دیدگاه آمار بیزی بیان شده، مورد تحلیل قرار می‌گیرند. چون در این دو روش، توزیع پسین پارامترها به شکل بسته قابل حصول نیست، توزیع‌های پسین شرطی کامل پارامترها محاسبه و از الگوریتم نمونه‌گیری گیبز برای استنباط استفاده می‌شود. برای مقایسهٔ کارایی روش‌های بیزی ارائه‌شده با روش‌های متداول، مطالعهٔ شبیه‌سازی انجام شده و در پایان نیز روش استفاده از مدل‌ها در قالب مثال کاربردی شرح داده خواهد شد.

واژه‌های کلیدی: استنباط بیزی، تاوان لاسو سازوار، توزیع لاپلاس نامتقارن، داده‌های پانلی پویا، رگرسیون چندکی، نمونه‌گیری گیبز.

۱ مقدمه

مطالعهٔ داده‌های پانلی پویا [۳] طرحی در تحقیقات علوم مختلف نظیر علوم اقتصادی و اجتماعی است که به‌سبب نیاز محققان به این نوع مطالعه‌ها مورد استفاده قرار می‌گیرد. در این نوع مطالعات، متغیر پاسخ در چندین نوبت متوالی مشاهده می‌شود، یا به عبارت دیگر یک واحد آزمایشی تحت اندازه‌گیری مکرر در طول زمان قرار می‌گیرد. ویژگی بارز مدل داده‌های پانلی پویا وجود متغیر تأخیری به‌عنوان یک متغیر توصیفی است. چون متغیر پاسخ تحت تأثیر اندازهٔ خود در دورهٔ قبل است، این متغیر در دورهٔ قبل به‌عنوان متغیر توصیفی برای پاسخ‌های فعلی وارد مدل می‌شود. واضح است تحلیل این نوع داده‌ها با توجه به نوع متغیر پاسخ، روش‌های خاص خود را می‌طلبد. معمولاً برای تحلیل این نوع داده‌ها از مدل‌های خطی پانلی پویا با اثرهای تصادفی استفاده می‌شود. اما در این مدل‌ها برآوردگرهای کمترین مربعات معمولی اغلب ناسازگار و دارای اریبی قابل ملاحظه‌ای هستند. روش‌های برآوردی مختلفی که برآوردگرهای سازگار و نااریب در این مدل‌ها دارند نیز پیشنهاد شده است [۹]. اما در سال‌های اخیر مدل‌های رگرسیونی چندکی برای داده‌های پانلی پویا مورد توجه قرار گرفته‌اند [۶و۷].

رگرسیون چندکی [۴] اولین بار در [۱۲] ارائه شد. این روش نسبت به‌روش رگرسیون میانگین [۵] دارای دو مزیت مهم است. نخست این‌که رگرسیون چندکی نسبت به ناپایداری واریانس و داده‌های دورافتاده حسّاس نیست و دوم این‌که اطلاعات

[۱] استادیار گروه آمار دانشگاه زنجان، ایران
[۲] دانش‌آموختهٔ کارشناس ارشد آمار دانشگاه زنجان،ایران

[۳] Daynamic Panel Data
[۴] qauntile regression
[۵] mean regression

جزئی‌تری نسبت به تأثیر متغیرهای توصیفی در چندک‌های مختلف توزیع متغیر پاسخ ارائه می‌کند. همان‌طور که عنوان شد در مطالعهٔ داده‌های پانلی، علاوه بر اثرهای ثابت که تغییرات درون گروه‌ها (تغییرات در طول زمان) را کنترل می‌کنند، اثرهای تصادفی بین متغیرهای پاسخ نیز در مدل لحاظ می‌شود که تغییرات بین‌گروهی را کنترل می‌کنند. اما در بسیاری از مسائل، تعداد مشاهدات از متغیر پاسخ نسبت به مشاهدات انجام‌گرفته در طول زمان از متغیر مربوط بسیار زیاد است. بنا بر این در این مدل تعداد اثرهای تصادفی (تعداد پارامترها) نیز خیلی زیاد می‌شود و لذا تعبیر و تفسیر مدل، مشکل می‌شود و دقت آن نیز کاهش می‌یابد. برای رفع این مشکل، کوئنکر [۱۱] رگرسیون چندکی را برای تحلیل داده‌های پانلی با افزودن تاوان لاسو[۶] (تاوان با نرم l_1) روی اثرهای تصادفی بررسی کرد. در واقع این روش با ایجاد تاوان، اثرهای تصادفی کم‌اهمیت را به سمت صفر منقبض[۷] کرده و مدلی تنک[۸] ایجاد می‌کند. اولین بار تیبشیرانی [۱۷] رگرسیون لاسو را که با ایجاد تاوان لاسو روی پارامترها انجام می‌شود ارائه کرد. سپس زو [۲۰] رگرسیون لاسو سازوار[۹] را به‌منظور گسترش رگرسیون لاسو پیشنهاد کرد. این روش بر خلاف تاوان لاسو که برای همهٔ ضرایب، تاوان یکسانی در نظر گرفته می‌شود، اندازهٔ تاوان‌های مختلفی برای ضرایب رگرسیونی متفاوت لحاظ می‌کند. در واقع بر خلاف تاوان لاسو، این روش بر حسب میزان اهمیت هر یک از ضرایب رگرسیونی، تاوانی متناسب با آن را در نظر می‌گیرد. بررسی‌های زو [۲۰] مشخص کرد که در رگرسیون چندکی با تاوان لاسو سازوار، اریبی در برآورد پارامترها در مقایسه با روش لاسو کمتر است. الحمزوی و همکاران [۱و۲] مدل رگرسیون چندکی و لنگ و همکاران [۱۴] مدل رگرسیون میانگین با تاوان لاسو سازوار را از دیدگاه آمار بیزی معرفی کردند و نشان دادند که این تاوان در مقایسه با تاوان لاسو دقت زیادی دارد.

کوئنکر [۱۱] و گراسی و باتای [۸] با تعریف توزیع خاص برای اثرهای تصادفی و با در نظر گرفتن توزیع لاپلاس نامتقارن برای متغیرهای پاسخ، مدل رگرسیونی چندکی در داده‌های پانلی را از دیدگاه آمار بیزی بررسی کردند. در این روش، با فرض این‌که توزیع اثرهای تصادفی لاپلاس متقارن هستند، مدل کوئنکر [۱۱] حاصل خواهد شد. گالوائو و مونتس روخاس [۷] مدل ارائه‌شده توسط کوئنکر [۱۱] را برای داده‌های پانلی پویا با اثرهای ثابت از دیدگاه آمار بسامدی مطالعه کردند. گالوائو [۶] رگرسیون چندکی برای داده‌های پانلی پویا را با در نظر گرفتن اثرهای تصادفی و استفاده از متغیرهای ابزاری[۱۰] بدون در نظر گرفتن تاوان لاسو بررسی کرد. کوباباشی و کوزومی [۱۰] نیز داده‌های پانلی پویای سانسور شده را بدون در نظر گرفتن تاوان لاسو از دیدگاه آمار بیزی بررسی کردند.

هدف این مقاله تحلیل داده‌های پانلی پویا با استفاده از روش رگرسیون چندکی با در نظر گرفتن تاوان لاسو سازوار روی اثرهای تصادفی از دیدگاه آمار بیزی است. برای این منظور، تاوان لاسو سازوار از طریق تعریف توزیع پیشین مناسب روی اثرهای تصادفی در مدل لحاظ می‌شود و مدل از دیدگاه آمار بیزی تحلیل می‌شود. بنا بر این بخش دوم، شامل مطالب کلی در مورد مدل رگرسیون چندکی و رگرسیون چندکی در داده‌های پانلی پویا است. در بخش سوم، مدل رگرسیون چندکی با تاوان لاسو سازوار از دیدگاه آمار بیزی بررسی می‌شود و در بخش چهارم نیز با روش شبیه‌سازی، کارایی مدل ارائه‌شده با مدل‌های دیگر مورد مقایسه قرار می‌گیرد. بخش پنجم شامل تحلیل داده‌های واقعی و ارزیابی کارایی مدل‌ها است.

۲ رگرسیون چندکی

مدل رگرسیون خطی

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \ldots, n \qquad (1)$$

را در نظر بگیرید که در آن $\boldsymbol{x}_i$ بردار مربوط به متغیرهای توصیفی، $\boldsymbol{\beta}$ یک بردار $(k \times 1)$ بعدی از پارامترها و ϵ_i مؤلفهٔ خطا است. در مدل رگرسیونی میانگین، هدف استنباط در مورد متوسط مقدار پاسخ به‌ازای مقادیر متفاوت متغیرهای توصیفی است؛ لذا $E(y_i|\boldsymbol{x}_i) = \boldsymbol{x}_i'\boldsymbol{\beta}$ مورد توجه است. اما در رگرسیون

[۶] Lasso

[۷] shrink

[۸] spares

[۹] adaptive Lasso

[۱۰] instrumental variables

چندکی، تابع چندک شرطی

$$Q_\tau(y_i|\boldsymbol{x}_i) = \boldsymbol{x}_i'\boldsymbol{\beta}, \quad i = 1, \ldots, n$$

مورد توجه است که در آن $Q_\tau(\cdot)$ معکوس تابع توزیع تجمعی متغیر پاسخ y_i به‌شرط معلوم بودن بردار $\boldsymbol{x}_i$ است. بنا بر این $\boldsymbol{x}_i'\boldsymbol{\beta}$ همان چندک τام متغیر y_i را نشان می‌دهد. لذا برای مدل رابطهٔ ۱ در رگرسیون میانگین، فرض می‌شود که میانگین توزیع خطاها صفر است، امّا در رگرسیون چندکی، چندک τام ϵ_iها برابر با صفر است؛ یعنی $\int_{-\infty}^{\cdot} f_{\epsilon_i}(t)dt = \tau$. در مدل رگرسیون چندکی، ضرایب رگرسیونی یعنی $\boldsymbol{\beta}$ از مینیمم کردن عبارت

$$\sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}) \qquad (۲)$$

نسبت به $\boldsymbol{\beta}$ برآورد می‌شوند، که در آن $\rho_\tau(\cdot)$ نشان‌دهندهٔ تابع زیان و به‌صورت $\rho_\tau(u) = \frac{|u|+(2\tau-1)u}{2}$ تعریف می‌شود (کوئنکر و باست [۱۲]). چون رابطهٔ ۲ در مبدأ مشتق‌پذیر نیست، راه‌حل تحلیلی برای برآورد ضرایب رگرسیونی وجود ندارد. بنا بر این معمولاً از روش‌های بهینه‌سازی عددی برای برآورد $\boldsymbol{\beta}$ استفاده می‌شود و از این طریق هیچ‌گونه استنباط آماری نیز نمی‌توان در مورد برآوردگر $\boldsymbol{\beta}$ انجام داد. برای رفع این مشکل، یو و مویید [۱۹] رگرسیون چندکی را از دیدگاه آمار بیزی با استفاده از توزیع لاپلاس نامتقارن[۱۱] مورد بحث و بررسی قرار دادند. متغیر تصادفی y را دارای توزیع لاپلاس نامتقارن گویند ($y \sim ALD(\mu, \sigma, \tau)$) هرگاه تابع چگالی آن به‌صورت

$$f(y) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_\tau\left(\frac{y-\mu}{\sigma}\right)\right\}$$

باشد که در آن $0 < \tau < 1$ پارامتر چولگی، σ پارامتر مقیاس و $-\infty < \mu < \infty$ پارامتر مکان است. لذا با فرض $y_i \sim ALD(\boldsymbol{x}_i'\boldsymbol{\beta}, \sigma, \tau)$، تابع درست‌نمایی بردار $\boldsymbol{y} = (y_1, \ldots, y_n)$ به‌صورت

$$L(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{x}) \propto \sigma^{-n} \times \exp\left\{-\sum_{i=1}^{n} \rho_\tau\left(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)\right\} \qquad (۳)$$

به دست می‌آید. بنا بر این ماکسیمم کردن رابطهٔ ۳ در حضور پارامتر مزاحم σ با مینیمم کردن رابطهٔ ۲ نسبت به $\boldsymbol{\beta}$ معادل است. در واقع با در نظر گرفتن توزیع لاپلاس نامتقارن برای y_i، تابع درست‌نمایی برای مدل رگرسیون چندکی به دست می‌آید که در تحلیل‌های بیزی ضروری است. به همین دلیل در بسیاری از موارد، تحلیل رگرسیون چندکی از دیدگاه آمار بیزی با استفاده از این توزیع بررسی می‌شود. اولین بار یو و مویید [۱۹] با فرض $\sigma = 1$ و تعریف توزیع پیشین ناآگاهی‌بخش برای $\boldsymbol{\beta}$، ($\pi(\boldsymbol{\beta}) \propto 1$) رگرسیون چندکی بیزی را در داده‌های مقطعی معرفی و مطالعه کردند. هدف این مقاله تعمیم این روش به مدل داده‌های پانلی پویا با اثرهای تصادفی و با فرض تاوان لاسو سازوار روی اثرهای تصادفی از دیدگاه آمار بیزی است.

مدل رگرسیونی داده‌های پانلی پویا، یک مدل آمیختهٔ خطی به‌صورت

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \gamma y_{i(t-1)} + \alpha_i + \epsilon_{it},$$
$$i = 1, \cdots, n, \quad t = 1, \cdots T, \qquad (۴)$$

است که در آن y_{it} متغیر پاسخ برای iامین واحد آزمایشی در زمان t، $\boldsymbol{x}_{it}$ بردار $(k \times 1)$ بعدی از متغیرهای توصیفی، $y_{i(t-1)}$ متغیر پاسخ تأخیری، γ و $\boldsymbol{\beta}$ ضرایب رگرسیونی و α_i اثر تصادفی مربوط به پاسخ y_{it} را نشان می‌دهد. توجه کنید که در این مدل α_iها اثرهای بین گروهی و $\boldsymbol{\beta}$ اثرهای درون گروهی را کنترل می‌کنند. وجود $y_{i(t-1)}$ بیان‌کنندهٔ این است که متغیر پاسخ تحت تأثیر مقدار خود در دورهٔ قبل نیز قرار دارد. با فرض این‌که چندک τام ϵ_{it} برابر با صفر است، تابع چندک شرطی در این مدل به‌صورت

$$Q_\tau(y_{it}|\cdot) = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \gamma y_{i(t-1)} + \alpha_i \qquad (۵)$$

به دست می‌آید. کوئنکر [۱۱] مدل فوق را بدون در نظر گرفتن متغیرهای تأخیری و با فرض تاوان لاسو روی α_iها به‌صورت $\lambda \sum_{i=1}^{n} |\alpha_i|$ مورد بررسی قرار داد. او پارامترهای مدل یعنی $\boldsymbol{\alpha}$ و $\boldsymbol{\beta}$ را از مینیمم کردن تابع زیان تاوانیدهٔ

$$\sum_{k=1}^{q}\sum_{i=1}^{n}\sum_{t=1}^{T} w_k \rho_{\tau_k}(y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}_{\tau_k} - \alpha_i) + \lambda \sum_{i=1}^{n} |\alpha_i|,$$

نسبت به $\boldsymbol{\alpha}$ و $\boldsymbol{\beta}$ برآورد کرد و همچنین استدلال کرد که اگر n در مقایسه با T خیلی بزرگ باشد، تابع زیان تاوانیده فوق، نسبت به‌روش‌های غیر تاوانیده عملکرد بهتری دارد. در ادامه گالوائو و مونتس روخاس [۷] مدل رابطه را با فرض تاوان لاسو روی α_iها مورد بررسی قرار دادند. در واقع آن‌ها پارامترهای $\boldsymbol{\alpha}$ ، $\boldsymbol{\beta}$ و γ را از مینیمم کردن تابع زیان تاوانیدهٔ

$$\sum_{k=1}^{q}\sum_{i=1}^{n}\sum_{t=1}^{T} w_k \rho_{\tau_k}(y_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}_{\tau_k} - \gamma y_{i(t-1)} - \alpha_i) + \lambda \sum_{i=1}^{n} |\alpha_i|$$

[۱۱] asymmetric Laplace distribution

نسبت به این پارامترها برآورد کردند. آنها نشان دادند این روش، زمانی که متغیر پاسخ تحت تأثیر مقادیر دورهٔ قبل است، عملکرد بهتری دارد. هدف این مقاله تحلیل مدل رابطهٔ ۴ با تاوان لاسو سازوار روی اثرهای تصادفی از دیدگاه آمار بیزی است. با در نظر گرفتن این تاوان روی اثرهای تصادفی، تابع زیان تاوانیده به‌صورت

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\rho_\tau(y_{it}-\boldsymbol{x}'_{it}\boldsymbol{\beta}_\tau-\gamma y_{i(t-1)}-\alpha_i)+\sum_{i=1}^{n}\lambda_i|\alpha_i| \quad (۶)$$

به دست می‌آید که در آن عبارت $\sum_{i=1}^{n}\lambda_i|\alpha_i|$ تاوان لاسو سازوار است و به λ_i پارامتر تنظیم گویند. توجه کنیم که در این روش برای هر α_i اندازهٔ تاوان متفاوت λ_i در نظر گرفته شده است. هرچه مقدار این پارامتر بزرگتر باشد، انقباض به سمت صفر نیز بیشتر بوده، اثرهای تصادفی کم‌اهمیت از مدل حذف خواهند شد.

۳ مدل بیزی با تاوان لاسو سازوار

همان‌طورکه عنوان شد در مدل رگرسیون چندکی با تاوان لاسو سازوار در داده‌های پانلی پویا، تابع زیان تاوانیده با تاوان لاسو سازوار روی اثرهای تصادفی، به‌صورت رابطهٔ ۶ است. لذا اگر فرض کنیم در رابطهٔ ۴، $\epsilon_{ij}\sim ALD(0,\sigma,\tau)$ و توزیع پیشین برای $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_n)$ را به‌صورت

$$\pi(\boldsymbol{\alpha}|\lambda,\sigma)=\prod_{i=1}^{n}\frac{\lambda_i}{2\sqrt{\sigma}}\exp\left\{-\frac{\sum_{i=1}^{n}\lambda_i|\alpha_i|}{\sqrt{\sigma}}\right\}$$

در نظر بگیریم، توزیع پسین بردار $\boldsymbol{\alpha}$ به‌صورت

$$\pi(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{y},\boldsymbol{x},\sigma,\lambda)\propto\exp\{-\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{\sigma}\rho_\tau(y_{it}-\boldsymbol{x}'_{it}\boldsymbol{\beta}-\gamma y_{i(t-1)}-\alpha_i)\}\times\exp\{-\Sigma_{i=1}^{n}\nu_i|\alpha_i|\}, \quad (۷)$$

به دست می‌آید که در آن $\nu_i=\frac{\lambda_i}{\sqrt{\sigma}}$ است. بنا بر این مینیمم کردن تابع هدف رابطهٔ ۶ معادل ماکسیمم کردن توزیع پسین $\boldsymbol{\alpha}$ در رابطهٔ ۷ در حضور پارامتر مزاحم σ است. لذا می‌توان توزیع لاپلاس نامتقارن را برای این مدل مورد استفاده قرار داد. چون استفاده از تابع درست‌نمایی در توزیع لاپلاس نامتقارن به‌دلیل وجود قدر مطلق آسان نیست، در تحلیل‌های بیزی معمولاً از شکل آمیخته این توزیع استفاده می‌شود. اگر $u\sim ALD(0,\sigma,\tau)$ و متغیرهای z و e به‌ترتیب دارای توزیع نرمال استاندارد و نمایی با میانگین $\frac{\sigma}{\tau(1-\tau)}$ و مستقل از هم باشند،

$$u\overset{d}{=}k_1e+\sqrt{2\sigma e}z$$

که در آن $k_1=(1-2\tau)$ است [۱۳]. با استفاده از این خاصیت توزیع لاپلاس نامتقارن و با فرض این‌که $\epsilon_{it}\sim ALD(0,\sigma,\tau)$باشد، مدل رابطهٔ ۴ را می‌توان به‌صورت

$$y_{it}=\boldsymbol{x}'_{it}\boldsymbol{\beta}-\gamma y_{i(t-1)}-\alpha_i+k_1e_{it}+\sqrt{2\sigma e_{it}}z_{it} \quad (۸)$$

بازنویسی کرد، که در آن z_{it} و e_{it} به‌ترتیب دارای توزیع نرمال استاندارد و نمایی با میانگین $\frac{\sigma}{\tau(1-\tau)}$ و مستقل از هم هستند. بنا بر این توزیع شرطی y_{it} به‌شرط e_{it}، عضو خانوادهٔ توزیع‌های نرمال است و لذا تحلیل‌ها را می‌توان با استفاده از رگرسیون خطی میانگین انجام داد. به همین طریق برای سادگی، توزیع پیشین روی α_iها را که توزیع لاپلاس متقارن است به‌صورت توزیعی آمیخته از دو توزیع نرمال و نمایی به‌صورت

$$\alpha_i|s_i\sim N(0,s_i),\quad s_i|\nu_i^2\sim Exp(\frac{\nu_i^2}{2}) \quad (۹)$$

در نظر می‌گیریم که در آن $Exp(\frac{\nu_i^2}{2})$ توزیع نمایی با میانگین $\frac{2}{\nu_i^2}$ همان توزیع آمیختگی است [۱۶].

در برازش مدل‌های رگرسیونی پویا با اثرهای تصادفی به داده‌های پانلی پویا، وابستگی یا استقلال اثرهای تصادفی از مشاهدات اولیه موضوعی اساسی است که به آن مسئلهٔ شرایط اولیه گویند [۱۴و۱۸]. در این مقاله مدل را با هر دو فرض استقلال و وابستگی در نظر می‌گیریم و از دیدگاه آمار بیزی بررسی قرار می‌کنیم.

۱.۳ مدل با فرض استقلال y_i. و α_i

در این بخش فرض می‌کنیم که مشاهدات اولیه از اثرهای تصادفی مستقل هستند. چون هدف تحلیل بیزی است، برای ابرپارامتر ν_i^2 در رابطهٔ ۹ توزیع پیشین مزدوج را نمایی به‌صورت $\pi(\nu_i^2|\phi)=\phi\exp\{-\phi\nu_i^2\}$ و توزیع‌های پیشین برای پارامترهای σ و $\boldsymbol{\beta}$ را نیز مزدوج به‌ترتیب، معکوس گاما و نرمال چندمتغیره در نظر می‌گیریم. توزیع پیشین ϕ را به‌منظور کم کردن تأثیر آن بر استنباط‌ها، توزیع پیشین پخ[۱۲] به‌صورت $\pi(\phi)\propto\frac{1}{\phi}$ فرض می‌کنیم. حال با توجه به مدل رابطهٔ ۸ و ۹ و توزیع‌های پیشین

[۱۲] flat

تعریف شده، مدل سلسله‌مراتبی را می‌توان به‌صورت

$$y_{it}|e_{it},\boldsymbol{\beta},\sigma,\alpha_i \sim N(\boldsymbol{x}'_{it}\boldsymbol{\beta}+\gamma y_{i(t-1)}+\alpha_i + (1-2\tau)e_{it}, 2\sigma e_{it})$$

به‌قسمی که

$$\begin{aligned}
&\gamma|\psi^2 \sim N(0,\psi^2), \quad \psi^2 \sim IG(a,b),\\
&e_{it}|\sigma \sim Exp(\frac{\tau(1-\tau)}{\sigma}), \quad \boldsymbol{\beta} \sim N_k(\boldsymbol{b}.,\boldsymbol{B}.),\\
&\alpha_i|s_i \sim N(0,s_i), \quad s_i|\nu_i^2 \sim Exp(\frac{\nu_i^2}{2}),\\
&\nu_i^2|\phi \sim Exp(\phi), \quad \pi(\phi)\propto\frac{1}{\phi}, \quad \sigma|c.,d. \sim IG(c.,d.)
\end{aligned}$$

نوشت، که در آن $IG(a,b)$، نشان‌دهندهٔ توزیع گامای معکوس با پارامترهای a و b است. با توجه به مدل سلسله‌مراتبی فوق، توزیع پسین همهٔ پارامترها و ابرپارامترها به‌صورت بسته قابل حصول نیست. اما می‌توان توزیع‌های پسین شرطی کامل کلیه پارامترها و ابرپارامترها را محاسبه کرد که در پیوست الف ارائه‌شده است. همهٔ توزیع‌های پسین شرطی کامل پارمترها و ابرپارمترها، توزیع‌های شناخته‌شده‌ای هستند و تولید نمونه‌های تصادفی از آنها نیز با استفاده از بسته‌های نرم‌افزار R در دسترس اند؛ لذا با استفاده روش نمونه‌گیری گیبز می‌توان استنباط‌های بیزی را انجام داد.

۲.۳ مدل با فرض وابستگی $y_{i.}$ و α_i

در این بخش، مدلی را در نظر می‌گیریم که در آن، مشاهدات اولیه و اثرهای تصادفی وابسته هستند. در این مدل، تابع درست‌نمایی به‌صورت تابع چگالی توأم از مشاهدات اولیه و آتی خواهد بود. لذا توزیع مشاهدات اولیه به‌شرط اثرهای تصادفی را به پیروی از [۱۸] به‌صورت

$$y_{i.}|\alpha_i \sim N(\phi_{.}+\phi_1\alpha_i,\sigma_{.}^2)$$

در نظر می‌گیریم. حال برای تحلیل بیزی باید توزیع پیشین برای پارامترها و ابرپارامترها مشخص شود. توزیع‌های پیشین پارامترهای $\boldsymbol{\beta}$ و σ را مانند بخش ۱.۳ در نظر می‌گیریم. توزیع‌های پیشین پارامترهای ϕ_1، $\phi_{.}$ و $\sigma_{.}^2$ را نیز مانند [۱۰] به‌صورت

$$\phi_1 \sim N(\mu_{\phi_1},\sigma^2_{\phi_1}),\ \phi_{.} \sim N(\mu_{\phi_{.}},\sigma^2_{\phi_{.}}),\ \sigma_{.}^2 \sim IG(\frac{v_{.}}{2},\frac{\delta_{.}}{2})$$

در نظر می‌گیریم. حال با توجه به مدل رابطه‌های ۸ و ۹ و توزیع‌های پیشین تعریف‌شده، مدل سلسله‌مراتبی را می‌توان به‌صورت

$$y_{it}|e_{it},\boldsymbol{\beta},\sigma,\alpha_i \sim N(\boldsymbol{x}'_{it}\boldsymbol{\beta}+\gamma y_{i(t-1)}+\alpha_i + (1-2\tau)e_{it}, 2\sigma e_{it})$$

به‌قسمی که

$$\begin{aligned}
&y_{i.}|\alpha_i \sim N(\phi_{.}+\phi_1\alpha_i,\sigma_{.}^2), \quad \alpha_i|s_i \sim N(0,s_i),\\
&s_i|\nu_i^2 \sim Exp(\frac{\nu_i^2}{2}),\gamma|\psi^2 \sim N(0,\psi^2),\psi^2 \sim IG(a,b),\\
&e_{it}|\sigma \sim Exp(\frac{\tau(1-\tau)}{\sigma}), \quad \boldsymbol{\beta} \sim N_k(\boldsymbol{b}.,\boldsymbol{B}.),\\
&\nu_i^2|\phi \sim Exp(\phi),\pi(\phi)\propto\frac{1}{\phi},\sigma|c.,d. \sim IG(c.,d.),\\
&\phi_{.} \sim N(\mu_{\phi_{.}},\sigma^2_{\phi_{.}}), \quad \phi_1 \sim N(\mu_{\phi_1},\sigma^2_{\phi_1}),\\
&\sigma_{.}^2 \sim IG(\frac{v_{.}}{2},\frac{\delta_{.}}{2}),
\end{aligned}$$

نوشت. در این مدل نیز توزیع پسین کامل پارامترها و ابرپارامترها به‌صورت بسته در دسترس نیست. اما با اندکی محاسبات جبری، توزیع پسین شرطی کامل پارامترها و ابرپارامترها قابل حصول‌است، لذا می‌توان از روش نمونه‌گیری گیبز برای استنباط بیزی استفاده کرد. توزیع‌های پسین پارامترهای $\boldsymbol{\beta}$، γ، σ، $\boldsymbol{s}$، $\boldsymbol{\nu}^2$، ϕ، $\boldsymbol{e}$ و ψ^2 مانند بخش ۱.۳ است؛ اما توزیع پسین شرطی کامل سایر پارامترها در پیوست ب ارائه‌شده است.

۴ مطالعهٔ شبیه‌سازی

در این بخش برای مقایسهٔ کارایی روش‌های بیزی ارائه‌شده با روش‌های متداول دیگر به مطالعهٔ شبیه‌سازی، می‌پردازیم. در این شبیه‌سازی مدل را به‌صورت

$$\begin{aligned}
y_{it} = x_{it}^{(1)}\beta_1 + x_{it}^{(2)}\beta_2 + \gamma y_{i(t-1)} + \alpha_i + \epsilon_{it},\\
i=1,\dots,n, \quad t=1,\dots,T
\end{aligned}$$

در نظر می‌گیریم که در آن $n=100$ و $T=5$ است. توجّه کنیم که تعداد متغیرهای پاسخ در مقایسه با متغیّرهای توصیفی زیاد است. لذا استفاده از تاوان لاسو سازوار منجر به بهبود برآورد پارامترها و دقت مدل خواهد شد. متغیرهای توصیفی را به‌صورت $x_{it}^{(j)} \sim N(0,1)$، $j=1,2$ و α_iها را مانند لیو و همکاران [۱۵] از توزیع $N(0,4)$ تولید کرده، مؤلفه‌های خطا را نیز از سه توزیع به‌صورت

$$\epsilon_{it} \sim N(0,1), \quad \epsilon_{it} \sim t(3), \quad \epsilon_{it} \sim unif(-2,2)$$

تولید می‌کنیم که در آن $t = -۴۹, \ldots, T$ است. همچنین برای تولید متغیرهای پاسخ فرض می‌کنیم $y_{i,-۵۰} = ۰$ و پاسخ‌های y_{it} را برای $i = ۱, \ldots, n$ و $t = -۴۹, \ldots, T$ تولید می‌کنیم. سپس ۵۰ مشاهدهٔ اول را کنار می‌گذاریم و مشاهدات مربوط به $t = ۰, \ldots, T$ را برای تحلیل مورد استفاده قرار می‌دهیم. توجه کنید که این روش تولید داده‌ها برای اطمینان از عدم تأثیر مقدار اولیه بر داده‌های مورد تحلیل است. مقدار واقعی پارامتر $\boldsymbol{\beta}$ را برابر با $(۱, ۲)$ قرار می‌دهیم و شبیه‌سازی را برای مقادیر مختلف γ به‌صورت $\gamma = (۰, ۰/۱, ۰/۳, ۰/۵, ۰/۸, ۰/۹)$ انجام می‌دهیم. توزیع‌های پیشین را برای پارامترها و ابرپارامترها تقریبا ناآگاهی‌بخش به‌صورت

$\boldsymbol{\beta} \sim N(۰, ۱۰۰\boldsymbol{I})$، $\sigma \sim IG(۰/۰۱, ۰/۰۱)$، $\phi_۰ \sim N(۰, ۱۰۰)$، $\phi_۱ \sim N(۰, ۱۰۰)$ و $\sigma^۲ \sim IG(۰/۰۱, ۰/۰۱)$ در نظر می‌گیریم. برای برآورد پارامترها در مدل‌های بیزی با استفاده از روش نمونه‌گیری گیبز، ۸۰۰۰ نمونه از توزیع‌های پسین شرطی کامل پارامترها تولید کرده، ۲۰۰۰ نمونهٔ اول را به‌عنوان دورهٔ همگرایی مطلوب کنار می‌گذاریم. برای ارزیابی همگرایی زنجیر و تعیین تقریبی تعداد نمونه‌ها به عنوان همگرایی مطلوب از آمارهٔ $\hat{R}$ که در [۵] ارائه‌شده است، استفاده می‌کنیم. روش‌های بیزی، یعنی رگرسیون چندکی با تاوان لاسو سازوار از دیدگاه آمار بیزی با فرض وابستگی اثرات تصادفی و مشاهدات اولیه (BD) و با فرض استقلال اثرهای تصادفی و مشاهدات اولیه (BI) را با روش رگرسیون چندکی معمولی بدون در نظر گرفتن اثرهای تصادفی (QR) و روش رگرسیون چندکی در داده‌های پانلی با در نظر گرفتن تاوان لاسو روی اثرات تصادفی از دیدگاه آمار بسامدی (PQR) که توسط کوئنکر [۱۱] ارائه‌شده است، مقایسه می‌کنیم. همانند گراسی و باتای [۸] دو نوع تاوان $\lambda = ۰/۰۱$، (PQ1) و $\lambda = ۵$، (PQ2) را برای روش کوئنکر [۱۱] در نظر می‌گیریم. برای مقایسهٔ کارایی روش‌ها و به‌منظور لحاظ کردن تغییرپذیری نتایج حاصل از شبیه‌سازی، روند شبیه‌سازی را ۱۰۰ مرتبه تکرار می‌کنیم و دو معیار اریبی[۱۳] پارامترها و مجذور میانگین توان‌های دوم خطا[۱۴] را طبق روابط زیر برای پارامتر $\boldsymbol{\beta}$ و γ در نظر می‌گیریم.

$$\text{Bias}(\hat{\beta}_l) = \frac{۱}{۱۰۰}\sum_{r=۱}^{۱۰۰}(\hat{\beta}_l^r - \beta_l), \quad l = ۱, \ldots, k,$$

$$\text{RMSE}(\hat{\beta}_l) = \sqrt{\frac{۱}{۱۰۰}\sum_{r=۱}^{۱۰۰}(\hat{\beta}_l^r - \beta_l)^۲}, \quad l = ۱, \ldots, k,$$

$$\text{Bais}(\hat{\gamma}) = \frac{۱}{۱۰۰}\sum_{r=۱}^{۱۰۰}(\hat{\gamma}^r - \gamma),$$

$$\text{RMSE}(\hat{\gamma}) = \sqrt{\frac{۱}{۱۰۰}\sum_{r=۱}^{۱۰۰}(\hat{\gamma}^r - \gamma)^۲}.$$

واضح است هرچه اندازهٔ این معیارها کوچک‌تر باشند، حاکی از کارایی زیاد مدل است. چون نوشتن همهٔ مقادیر با توجه به وجود پنج مدل، دو روش برآورد برای سه نوع توزیع خطا و وجود سه پارامتر در هر شبیه‌سازی و در مجموع ۱۸ پارامتر، در قالب جدول امکان‌پذیر نیست، بنا بر این نتایج در قالب نمودارهای جعبه‌ای برای چندک‌های $\tau = (۰/۲۵, ۰/۵, ۰/۷۵)$، خلاصه شده و نتایج در شکل‌های ۱ و ۲ و ۳ ارائه‌شده است. هرچه نمودارهای جعبه‌ای کوچک‌تر و نزدیک به صفر باشند، نشان‌دهنده اریبی کمتر و برآورد بهتر پارامترها است. شکل ۱ عملکرد روش‌ها را از نظر اریبی و RMSE در چندک $\tau = ۰/۲۵$ مقایسه می‌کند. با توجه به این شکل در روش‌های بیزی BI و BD اریبی پارامتر بسیار نزدیک به صفر است. از نظر معیار RMSE نیز به‌غیر از توزیع نرمال استاندارد که روش PQ1 عملکرد نزدیک به‌روش‌های بیزی دارد، روش‌های بیزی عملکرد بهتری نسبت به سایر روش‌ها دارند. شکل‌های ۲ و ۳ نیز عملکرد روش‌ها را به‌لحاظ معیارهای ذکر شده در چندک‌های $\tau = ۰/۵$ و $\tau = ۰/۷۵$ مقایسه می‌کنند. با توجه به این دو شکل نیز مشاهده می‌شود که مانند شکل ۱، در این نمودارها نیز روش‌های بیزی برآورد بهتری از پارامترها در همهٔ توزیع‌های فرض‌شده برای خطا دارند. همچنین توجه کنیم روش QR به‌دلیل این‌که اثرهای تصادفی را در مدل لحاظ نمی‌کند، نسبت به‌روش‌های دیگر که اثرات تصادفی را در مدل لحاظ می‌کنند، عملکرد خوبی ندارد.

۵ تحلیل داده‌های واقعی

در این بخش به مطالعه و تحلیل داده‌های واقعی می‌پردازیم. این داده‌ها شامل رشد اندازهٔ وزن ۲۰۰ کودک از بدو تولد تا دوسالگی است که از سال ۱۳۹۱ تا سال ۱۳۹۲ (دو سال) از درمانگاه‌های شهرستان قم به‌تصادف انتخاب و گردآوری شده است. هدف بررسی تأثیر برخی عوامل بر میزان رشد وزن کودکان است. این

[۱۳] bias

[۱۴] root mean square errors

عوامل، یعنی متغیرهای توصیفی، عبارت‌اند از جنسیت کودک (β_1)، محل سکونت خانوار (β_2)، شاغل بودن مادر (β_3)، سن مادر (β_4)، تحت مراقبت ویژه بودن مادر (β_5)، وزن مادر (β_6)، قد مادر (β_7)، سن جنینی نوزاد (β_8)، چندقلو بودن نوزاد (β_9)، مدت شیردهی انحصاری مادر (β_{10})، مدت شیردهی مادر در کنار غذای کمکی (β_{11}) و متغیر پاسخ نیز اندازهٔ وزن کودک است که در فواصل زمانی بدو تولد و ۲، ۴، ۶، ۹، ۱۲، ۱۸و ۲۴ ماهگی اندازه‌گیری شده است. دو مدل برای تحلیل داده‌ها به‌صورت زیر در نظر می‌گیریم.

مدل ۱: مدل داده‌های پانلی بدون در نظر گرفتن متغیر تأخیری؛ یعنی

$$y_{it} = x_{it}^{(1)}\beta_1 + x_{it}^{(2)}\beta_2 + \ldots, x_{it}^{(11)}\beta_{11} + \alpha_i + \epsilon_{it},$$
$$i = 1, \ldots, 200, \quad t = 1, \ldots, 8.$$

مدل ۲: مدل داده‌های پانلی پویا با در نظر گرفتن متغیرهای تأخیری؛ یعنی

$$y_{it} = x_{it}^{(1)}\beta_1 + \ldots + x_{it}^{(11)}\beta_{11} + \gamma y_{i(t-1)} + \alpha_i + \epsilon_{it},,$$
$$i = 1, \ldots, 200, \quad t = 1, \ldots, 8.$$

معیارهای اطلاع آکائیکه[15] (AIC) و بیزی[16] (BIC) برای روش‌های مختلف عنوان شده و دو مدل ۱ و ۲ محاسبه شده است که نتایج آن برای چندک‌های $\tau = (0/25, 0/5, 0/75)$ در جدول ۱ ارائه‌شده است. نتایج جدول ۱ نشان می‌دهد که مقدار این دو معیار برای مدل دوم در همهٔ روش‌ها کوچک‌تر است، لذا مدل ۲ در همهٔ چندک‌های ذکر شده بهتر از مدل اول به داده‌ها برازنده شده است. همچنین با توجه به این جدول، دو روش بیزی در مدل دوم نیز نسبت به‌روش‌های بسامدی، بهتر به داده‌ها برازنده شده‌اند. جدول ۲ نیز برآورد ضرایب رگرسیونی را در مدل ۲ برای روش‌های مختلف در چندک ۰/۵ نشان می‌دهد. با توجه به این جدول برآورد پارامترها در همهٔ روش‌ها تقریباً یکسان است. همان‌طور که انتظار می‌رفت، ضریب متغیر تأخیری برای همهٔ روش‌ها مقدار بزرگ (نزدیک یک) است که حاکی از وابستگی متغیر پاسخ به اندازهٔ آن در زمان قبل است.

بحث و نتیجه‌گیری

در داده‌های پانلی پویا متغیر پاسخ تحت تأثیر مقدار خود در دورهٔ قبل است. در این نوع مطالعات، متغیر تأخیری به‌عنوان متغیر توصیفی وارد مدل می‌شود. در مدل‌های پانلی پویا علاوه بر اثرهای ثابت که تأثیر هر یک از متغیرهای توصیفی را بر متغیر پاسخ بیان می‌کند، اثرهای تصادفی نیز در مدل لحاظ می‌شود. این اثرها تغییرات بین متغیرهای پاسخ (تغییرات بین‌گروهی) را کنترل می‌کنند. در مطالعاتی که تعداد مشاهدات از متغیر پاسخ زیاد است، اثرهای تصادفی نیز زیاد می‌شود و پارامترهای زیادی وارد مدل شده، بنا بر این دقت مدل کم و تفسیر آن مشکل می‌شود. در این مقاله با ایجاد تاوان لاسو سازوار روی اثرهای تصادفی، اثرهای تصادفی را به سمت صفر منقبض کرده و اثرهای کم‌اهمیت از مدل حذف شدند. تاوان لاسو سازوار با تعریف توزیع پیشین مناسب روی اثرهای تصادفی، در مدل لحاظ شده و از دیدگاه آمار بیزی مورد بررسی قرار گرفت. نتایج حاصل از شبیه‌سازی نشان دادند که روش بیزی ارائه‌شده در تمامی چندک‌ها نسبت به‌روش‌های متداول دیگر در رگرسیون چندکی برای داده‌های پانلی پویا با اثرات تصادفی عملکرد بهتری دارد. نتایج حاصل از تحلیل داده‌های واقعی نیز مشخص کرد که در تحلیل داده‌های رشد وزن کودک، استفاده از مدل داده‌های پانلی پویا مناسب‌تر است. همچنین دو معیار AIC و BIC حاکی از برازش خوب مدل بیزی ارائه‌شده به داده‌ها، در مقایسه با روش‌های بسامدی است.

[15] Akaike Information Criterion

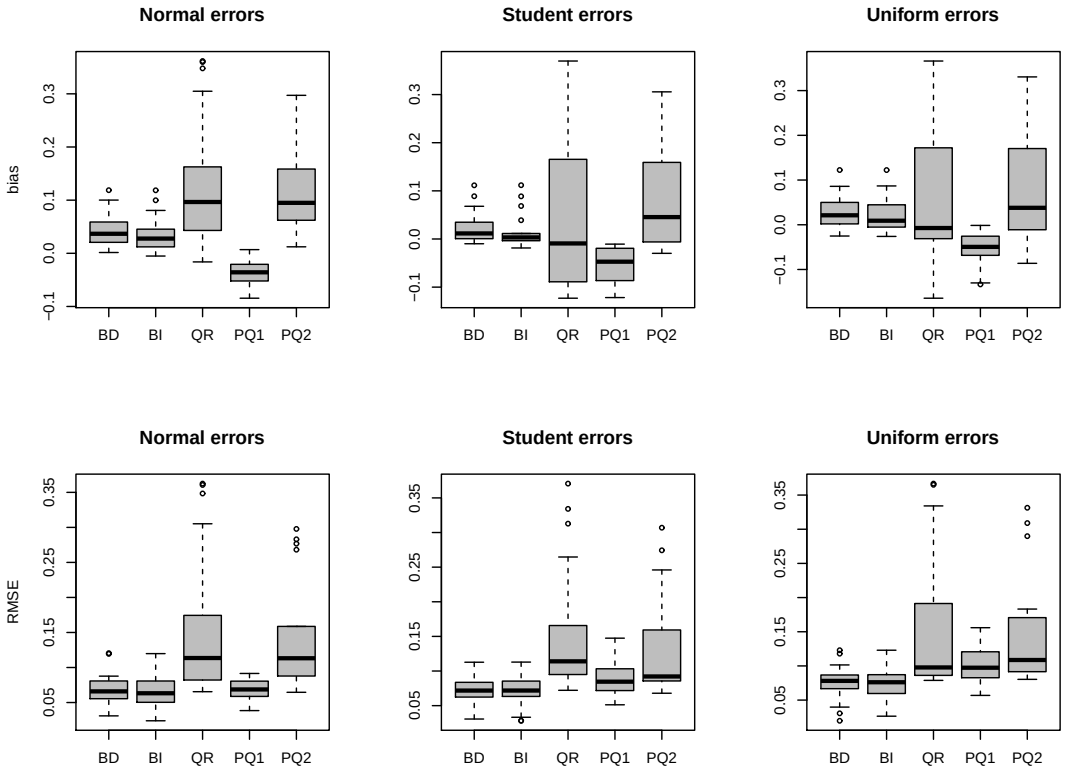
[16] Bayesian Inforrmation Criterion

جدول۱. برآورد معیارهای AIC و BIC برای دو مدل عنوان شده

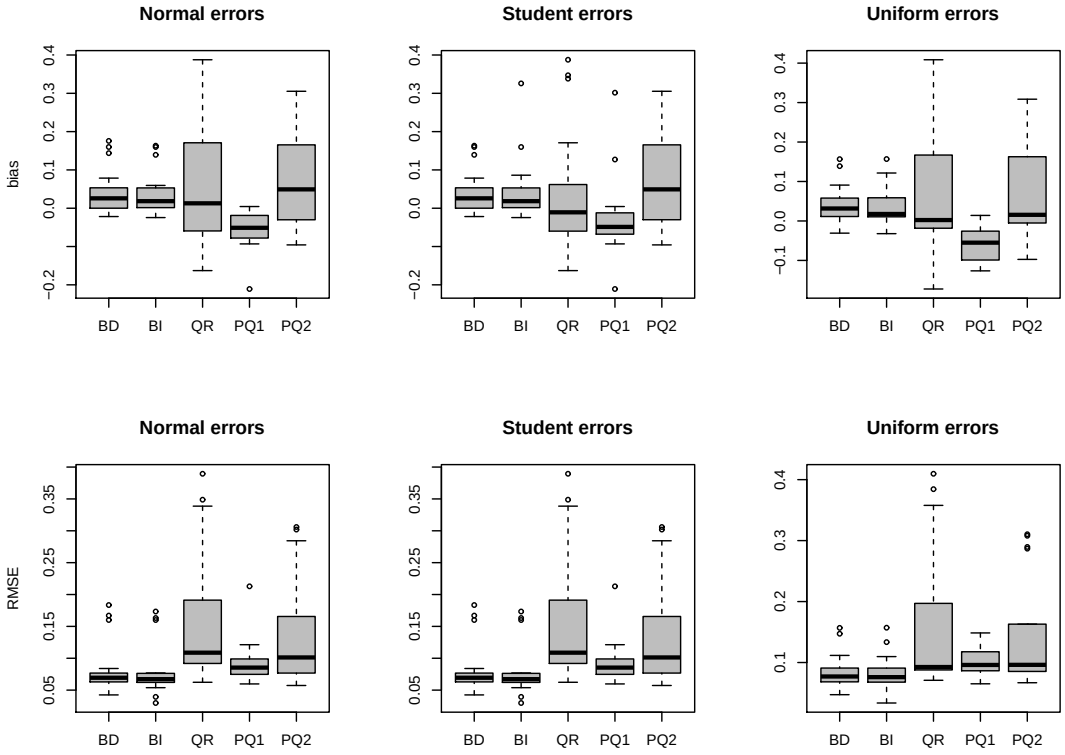
چندک	روش	مدل ۱		مدل ۲	
		AIC	BIC	AIC	BIC
۰/۲۵	BD	۲/۷۸۳۹	۲/۸۳۶۸	۰/۴۶۵۳	۰/۴۰۸۷
	BI	۲/۷۸۴۳	۲/۸۳۷۴	۰/۴۶۹۲	۰/۵۰۵۲
	PQ1	۳/۰۰۸۲	۳/۰۶۱۱	۱/۱۶۷۶	۱/۲۰۷۶
	PQ2	۲/۸۳۴۹	۲/۸۸۷۸	۱/۱۷۱۵	/۲۰۳۷
۰/۵	BD	۴/۰۲۲۱	۴/۰۹۴۷	۰/۴۴۶۴	۰/۴۱۸۸
	BI	۴/۰۳۳۵	۴/۱۰۶۲	۰/۴۵۲۸	۰/۴۸۸۹
	PQ1	۴/۲۲۴۸	۴/۲۹۷۴	۱/۰۰۹۲	۱/۰۴۵۳
	PQ2	۴/۱۰۹۳	۴/۱۸۱۹	۱/۱۶۰۷	۱/۱۹۶۸
۰/۷۵	BD	۱/۳۶۸۳	۱/۳۹۶۶	۰/۴۸۹۶	۰/۴۰۷۶
	BI	۱/۳۶۹۱	۱/۳۹۷۴	۰/۴۹۶۲	۰/۵۳۲۳
	PQ1	۱/۴۱۷۰	۱/۴۴۵۳	۱/۲۱۵۱	۱/۲۵۱۲
	PQ2	۲/۲۲۷۴	۲/۲۵۵۷	۱/۷۷۵۱	۱/۲۱۳۶

جدول۲. برآورد ضرایب رگرسیونی برای مدل ۲

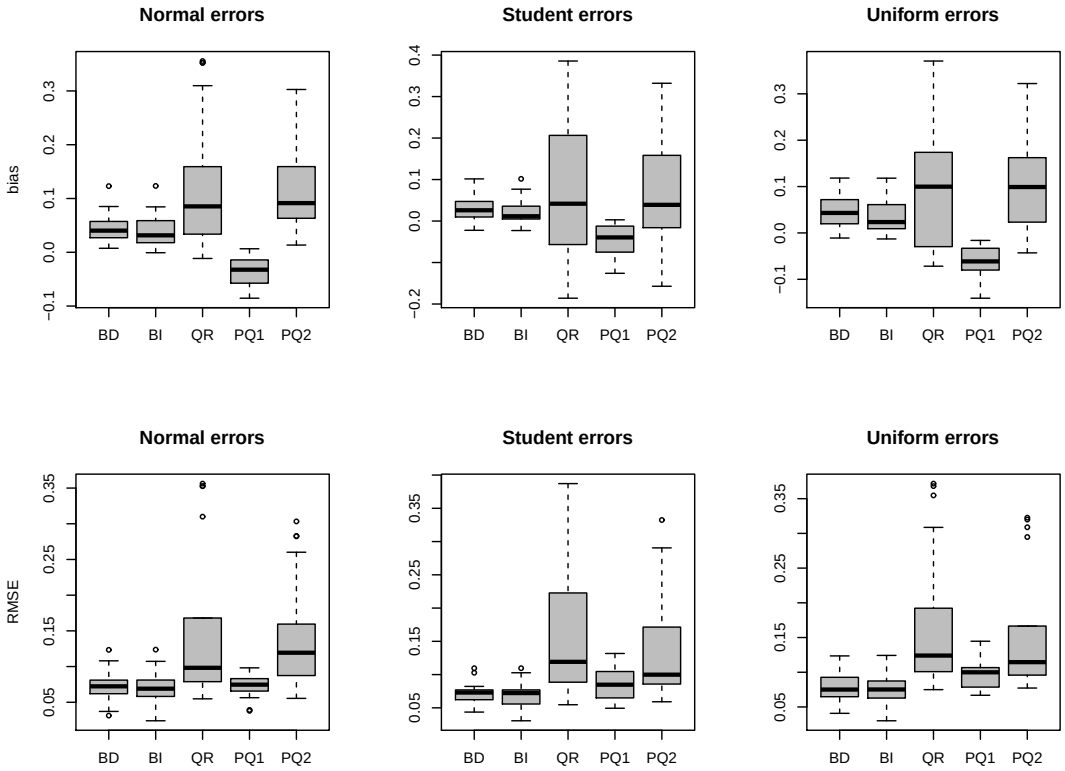
پارامتر	BD	BI	PQ1	PQ2
β_1	−۰/۰۰۲۴	−۰/۰۰۰۲	۰/۰۲۷۲	۰/۱۶۴۸
β_2	۰/۰۰۵۳	۰/۰۰۶۷	۰/۰۰۷۸	۰/۰۰۷۵
β_3	−۰/۰۰۳۷	−۰/۰۰۳۹	۰/۰۰۰۰	۰/۰۰۰۰
β_4	−۰/۰۲۴۰	−۰/۰۲۳۵	−۰/۰۲۵۰	−۰/۰۲۳۵
β_5	۰/۰۹۶۵	۰/۰۹۴۹	۰/۰۴۴۱	۰/۰۴۳۱
β_6	۰/۰۰۴۰	۰/۰۰۳۹	۰/۰۰۲۸	۰/۰۰۲۷
β_7	−۰/۰۰۲۱	−۰/۰۰۲۲	−۰/۰۰۰۳	−۰/۰۰۰۲
β_8	۰/۰۲۵۵	۰/۰۲۵۹	۰/۰۱۸۳	۰/۰۱۸۲۳
β_9	۰/۱۳۱۳	۰/۱۲۹۹۱	۰/۱۷۷۵	۰/۱۷۹۶
β_{10}	۰/۰۲۹۱	۰/۰۲۹۴	۰/۰۰۰۰	−۰/۱۳۸۱
β_{11}	۰/۰۰۰۱	۰/۰۰۰۵	−۰/۰۰۲۱	−۰/۰۰۲۰
γ	۰/۹۹۴۴	۰/۹۹۴۴	۰/۹۸۶۲	۰/۹۹۴۲



شکل ۱. عملکرد روش‌ها به‌لحاظ معیار اریبی و RMSE در چندک $\tau = ۰/۲۵$.



شکل ۲. عملکرد روش‌ها به‌لحاظ معیار اریبی و RMSE در چندک $\tau = ۰/۵$.



شکل ۳. عملکرد روش‌ها به‌لحاظ معیار اریبی و RMSE در چندک $\tau = ۰/۷۵$.

پیوست

الف) توزیع‌های پسین شرطی کامل پارامترها و ابرپارامترها با فرض استقلال α_i و y_i.

$$\begin{aligned}
\pi(\alpha_i|\cdot) &\propto f(\boldsymbol{y}_i|\boldsymbol{\beta},\alpha_i,\boldsymbol{e}_i,\gamma,\sigma,s_i,\nu^۲,\phi,)\times\pi(\alpha_i|s_i)\Longrightarrow\\
\alpha_i|\cdot &\sim N(\bar{\mu}_i,\tilde{\sigma}_i),\quad i=۱,\ldots,n\\
\bar{\mu}_i &= \tilde{\sigma}_i\left[\frac{۱}{۲\sigma}\sum_{t=۱}^{T}\frac{(y_{it}-\boldsymbol{x}'_{it}\boldsymbol{\beta}-\gamma y_{i(t-۱)}-k_۱e_{it})}{e_{it}}\right]\\
\tilde{\sigma}_i &= \left(\frac{۱}{s_i}+\frac{۱}{۲\sigma}\sum_{t=۱}^{T}\frac{۱}{e_{it}}\right)^{-۱}.
\end{aligned}$$

$$\begin{aligned}
\pi(\boldsymbol{\beta}|\cdot) &\propto f(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{e},\sigma,\boldsymbol{s},\nu^۲,\phi,)\pi(\boldsymbol{\beta})\Longrightarrow\boldsymbol{\beta}|\cdot\sim N_k(\boldsymbol{B}\boldsymbol{b},\boldsymbol{B})\\
\boldsymbol{b} &= \frac{۱}{۲\sigma}\sum_{i=۱}^{n}\sum_{t=۱}^{T}\frac{\boldsymbol{x}_{it}(y_{it}-\alpha_i-\gamma y_{i(t-۱)}-k_۱e_{it})}{e_{it}}+\boldsymbol{B}_{\cdot}^{-۱}\boldsymbol{b}.\\
\boldsymbol{B}^{-۱} &= \frac{۱}{۲\sigma}\sum_{i=۱}^{n}\sum_{t=۱}^{T}\frac{\boldsymbol{x}_{it}\boldsymbol{x}'_{it}}{e_{it}}+\boldsymbol{B}_{\cdot}^{-۱}.
\end{aligned}$$

$$\begin{aligned}
\pi(\sigma|\cdot) &\propto f(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{e},\sigma,\boldsymbol{s},\nu^۲,\phi,)\pi(\sigma)\Longrightarrow\sigma|\geq\sim IG(\nu_۱,w_۱)\\
w_۱ &= \sum_{i=۱}^{n}\sum_{t=۱}^{T}\{\frac{(y_{it}-\boldsymbol{x}'_{it}\boldsymbol{\beta}-\gamma y_{i(t-۱)}\alpha_i-k_۱e_{it})^۲}{۴e_{it}}+\tau(۱-\tau)e_{it}\}+d.\\
\nu_۱ &= \frac{۳N}{۲}+c.
\end{aligned}$$

$$\pi(e_{it}|\cdot) \propto f(y_{it}|\boldsymbol{\beta},\alpha_i,e_{it},\sigma,s_i,\nu^۲,\phi,)f(e_{it}),$$

$$\begin{aligned} &\propto (e_{it})^{-1}\exp\left\{-\frac{1}{2}(c_{it}^2 e_{it}^{-1}+d_{it}^2 e_{it})\right\}\Longrightarrow\\ e_{it}|\cdot &\sim GIG(0,c_{it},d_{it}),\quad i=1,\dots n,\quad t=1,\dots,T,\\ c_{it}^2 &= \frac{(y_{it}-\boldsymbol{x}_{it}'\boldsymbol{\beta}-\gamma y_{i(t-1)}-\alpha_i)^2}{2\sigma},\quad d_{it}^2=\frac{(1-2\tau)^2}{2\sigma}+\frac{2\tau(1-\tau)}{\sigma}, \end{aligned}$$

$GIG(r,c,d)$ نشان‌دهندهٔ توزیع گاوسی وارون تعمیم‌یافته[۱۷] با تابع چگالی

$$f(x|r,c,d)\propto x^{r-1}\exp\{-\frac{1}{2}(c^2x^{-1}+d^2x)\},x>0,-\infty<r<+\infty,\quad m,n>0$$

است. برای تولید نمونهٔ تصادفی از این توزیع می‌توان از تابع $rgig()$ استفاده کرد که در بستهٔ $ghyp$ در نرم‌افزار R موجود است [۳].

$$\begin{aligned} \pi(\gamma|\cdot) &\propto f(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{\alpha},\boldsymbol{e},\sigma,\boldsymbol{s},\nu^2,\phi,\gamma)\pi(\gamma|\psi)\\ &\propto (e_{it})^{-1}\exp\left\{-\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{4\sigma e_{it}}(y_{it}-\boldsymbol{x}_{it}^{'}\boldsymbol{\beta}-\gamma y_{i(t-1)}-\alpha_i-k_1e_{it})^2\right\}\\ &\quad\times\exp\left\{-\frac{\gamma^2}{2\psi^2}\right\}\Longrightarrow(\gamma|\cdot)\sim N(\bar{\mu}_\gamma,\tilde{\sigma}_\gamma),\\ \bar{\mu}_\gamma &= \tilde{\sigma}_\gamma\left[\frac{1}{2\sigma}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{y_{i(t-1)}(y_{it}-\boldsymbol{x}_{it}'\boldsymbol{\beta}-\alpha_i-k_1e_{it})}{e_{it}}\right],\\ \tilde{\sigma}_\gamma &= \left[\frac{1}{2\sigma}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{y_{i(t-1)}^2}{e_{it}}+\frac{1}{\psi^2}\right]^{-1}. \end{aligned}$$

$$\begin{aligned} \psi^2|\cdot &\sim IG(a+\frac{1}{2},b+\frac{\gamma^2}{2}),\quad s_i|\cdot\sim GIG(\frac{1}{2},\alpha_i,\nu_i),\quad i=1,\dots,n,\\ \phi|\cdot &\sim G(n,\sum_{i=1}^{n}\nu_i^2),\quad \nu_i^2|\cdot\sim G(2,\frac{s_i}{2}+\phi), \end{aligned}$$

ب): توزیع‌های پسین شرطی کامل پارامترها و ابرپارامترها با فرض وابستگی α_i و $y_{i\cdot}$

$$\begin{aligned} \pi(\alpha_i|\cdot) &\propto f(\boldsymbol{y}_i|\boldsymbol{\beta},\alpha_i,\boldsymbol{e}_i,\gamma,\sigma,s_i,\nu^2,\phi_1)f(y_{i\cdot}|\alpha_i,\phi_\cdot,\phi_1,\sigma_\cdot^2)\pi(\alpha_i|s_i)\\ &\propto \exp\left\{-\frac{1}{4\sigma}\sum_{t=1}^{T}\frac{(y_{it}-\boldsymbol{x}_{it}^{'}\boldsymbol{\beta}-\gamma y_{i(t-1)}-\alpha_i-k_1e_{it})^2}{e_{it}}-\frac{\alpha_i^2}{2s_i}\right\}\\ &\quad\times\exp\{-\frac{1}{\sigma_\cdot^2}(y_{i\cdot}-\phi_\cdot-\phi_1\alpha_i)^2\}\propto\exp\left\{-\frac{\tilde{\sigma}_i^2}{2}(\alpha_i-\bar{\mu}_i)^2\right\}\Longrightarrow\\ \alpha_i|\cdot &\sim N(\bar{\mu}_i,\tilde{\sigma}_i),\quad i=1,\dots,n\\ \bar{\mu}_i &= \tilde{\sigma}_i\left[\frac{1}{2\sigma}\sum_{t=1}^{T}\frac{(y_{it}-\boldsymbol{x}_{it}'\boldsymbol{\beta}-\gamma y_{i(t-1)}-k_1e_{it})}{e_{it}}+\frac{(y_{i\cdot}-\phi_\cdot)\phi_1}{\sigma_\cdot^2}\right],\\ \tilde{\sigma}_i &= \left(\frac{1}{s_i}+\frac{1}{2\sigma}\sum_{t=1}^{T}\frac{1}{e_{it}}+\frac{\phi_1^2}{\sigma_\cdot^2}\right)^{-1}. \end{aligned}$$

$$\begin{aligned} \pi(\phi_1|\cdot) &\propto \exp\{-\frac{1}{2\sigma_\cdot^2}\sum_{i=1}^{n}(y_{i\cdot}-\phi_\cdot-\phi_1\alpha_i)^2\}\times\exp\{-\frac{1}{\sigma_{\phi_1}^2}(\phi_1-\mu_{\phi_1})^2\}\Longrightarrow\\ \phi_1|\cdot &\sim N(\bar{\mu}_{\phi_1},\tilde{\sigma}_{\phi_1})\\ \bar{\mu}_{\phi_1} &= \tilde{\sigma}_{\phi_1}^2\left[\frac{\sum_{i=1}^{n}\alpha_i(y_{i\cdot}-\phi_\cdot)}{\sigma_\cdot^2}+\frac{\mu_{\phi_1}}{\sigma_{\phi_1}^2}\right],\quad \tilde{\sigma}_{\phi_1}^2=\left[\frac{\sum_{i=1}^{n}\alpha_i^2}{\sigma_\cdot^2}+\frac{1}{\sigma_{\phi_1}^2}\right]^{-1}. \end{aligned}$$

$$\begin{aligned} \pi(\phi_\cdot|\cdot) &\propto \exp\{-\frac{1}{2\sigma_\cdot^2}\sum_{i=1}^{n}(y_{i\cdot}-\phi_\cdot-\phi_1\alpha_i)^2\}\times\exp\{-\frac{1}{\sigma_{\phi_\cdot}^2}(\phi_\cdot-\mu_{\phi_\cdot})^2\}\Longrightarrow\\ \phi_\cdot|\cdot &\sim N(\bar{\mu}_{\phi_\cdot},\tilde{\sigma}_{\phi_\cdot}), \end{aligned}$$

[۱۷] Generalized Inverse Gaussian

$$\bar{\mu}_{\phi.} = \tilde{\sigma}^{۲}_{\phi.}\left[\frac{\sum_{i=۱}^{n}(y_{i.}-\phi_{۱}\alpha_i)}{\sigma^{۲}_{.}}+\frac{\mu_{\phi.}}{\sigma^{۲}_{\phi.}}\right], \quad \tilde{\sigma}^{۲}_{\phi.}=\left[\frac{n}{\sigma^{۲}_{.}}+\frac{۱}{\sigma^{۲}_{\phi.}}\right]^{-۱}.$$

$$\pi(\sigma^{۲}_{.}|.)\propto(\frac{۱}{\sigma^{۲}_{.}})^{\frac{n}{۲}}\exp\{-\frac{۱}{۲\sigma^{۲}_{.}}\sum_{i=۱}^{n}(y_{i.}-\phi_{.}-\phi_{۱}\alpha_i)^{۲}\}\times(\frac{۱}{\sigma^{۲}_{.}})^{(\frac{v_{.}}{۲}+۱)}\exp\{-\frac{\delta_{.}}{۲\sigma^{۲}_{.}}\}\Longrightarrow$$
$$\sigma^{۲}_{.}|.\sim IG(\frac{n+v_{.}}{۲},\frac{\sum_{i=۱}^{n}(y_{i.}-\phi_{.}-\phi_{۱}\alpha_i)^{۲}\}}{۲}+\frac{\delta_{.}}{۲})$$

مراجع

[1] Alhamzawi, R., Yu, K. and Benoit, D. F. (2012). Bayesian adaptive Lasso quantile regression, *Statistical Modeling*, **12**, 279-297.

[2] Alhamzawi, R., Benoit, D. F. and Yu, K. (2013). Binary quantile regression with adaptive lasso: A Bayesian approach, working paper.

[3] Breymann, W. and Luthi, D. (2014). ghyp: A package on the generalized hyperbolic distribution and its special cases, R Package, Version 1.5.6, URL: http:// www.r-project.org

[4] Crouchley, R. and Davies, R. B. (2001). A Comparison of GEE and random effects models for distinguishing heterogeneity nonstationarity and state dependent in collection of short binary event series, *Statistical Modelling*, **1**, 271-285.

[5] Gelman, A. Carlin, J. B. Stern, H. S. and Rubin, D. B. (1995). *Bayesian Data Analysis*, Chapman and Hall, London.

[6] Galvao, A. F. (2011). Quantile regression for dynamic panel data with fixed effects, *Journal of Applied Econometric*, **164**(**1**), 142-157.

[7] Galvao, A. F. and Montes-Rojas, G. V. (2010). Penalized quantile regression for dynamic panel data, *Journal of Statistical Planning and Inference*, **140**, 3476-3497.

[8] Geraci, M. and Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution, *Biostatistics*, **8**, 140-154.

[9] Hsiao, C. (2003). *Analysis of Panel Data, 2nd ed*. Cambridge University Press, Cambridge.

[10] Kobayashi, G., and Kozumi, H. (2012). Bayesian analysis of quantile regression for censored dynamic panel data, *Journal of Statistical Computation and Simulation*, **27**, 359-380.

[11] Koenker, R. (2004). Quantile regression for longitudinal data, *Journal of MultivariateAnalysis*, **91**, 74-89.

[12] Koenker, R. and Bassett, G. (1978). Quantile regression, *Econometrica*, **46**, 33-50 .

[13] Kozumi, H. and Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression, *Journal of Statistical Computation and Simulation*, **8**1,1565-1578.

[14] Leng, C., Tran, M. N. and Nott, D. A. (2014). Bayesian adaptive lasso, *Annals of the Institute of Statistical Mathematics*, **66**, 221-224.

[15] Luo, Y., Lian, H. and Tian, M. (2012), Bayesian quantile regression for longitudinal data model, *Journal of Statistical Computation and Simulation*, **82**, 1635-1649.

[16] Park, T. and Casella, T. (2008), The bayesian lasso. *Journal of the American Statistical Assosation*, **103**, 681-686.

[17] Tibshirani, R. (1996), Regression shrinkage and selection via the Lasso, *Journal of the Royal Statistical Society, Series B*, **58**, 267 – 288.

[18] Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic nonlinear panel data models with unobserved heterogeneity, *Journal of Applied Econometrics*, **20**, 30-54

[19] 13. Yu, K. and Moyeed, R.A. (2001). Bayesian quantile regression, *Statistics and Probability Letters*, **54**, 437-447.

[20] Zou, H. (2006). The adaptive Lasso and its oracle properties. *Journal of the American Statistical Association*, **101**, 1418-1429.