دانشکده علوم ریاضی
گروه آمار

پایان‌نامه

برای دریافت درجه کارشناسی ارشد در رشته
آمار، گرایش آمار ریاضی

عنوان

تحلیل بیزی مدل‌های رگرسیونی بتا با پاسخ‌های وابسته

استاد راهنما

دکتر حسین باغیشنی

پژوهشگر

مهناز عجم

شهریور ۱۳۹۳

نام خانوادگی دانشجو: عجم نام: مهناز

عنوان: تحلیل بیزی مدل‌های رگرسیونی بتا با پاسخ‌های وابسته

استاد راهنما: دکتر حسین باغیشنی

مقطع تحصیلی: کارشناسی ارشد رشته: آمار گرایش: آمار ریاضی

دانشگاه: شاهرود دانشکده علوم ریاضی

تاریخ فارغ‌التحصیلی: شهریور ۱۳۹۳ تعداد صفحات: ۱۰۵

واژگان کلیدی: تحلیل بیزی، رگرسیون بتا، مدل‌های آمیخته

چکیده

در بسیاری از مسایل تحلیل پاسخ‌های نرخ و نسبت مانند نرخ تورم، نرخ رشد، درصد مشتریان خوش قول در یک بنگاه اقتصادی و درصد یک بیماری در منطقه‌ای خاص، علاقه‌مند به کشف عوامل موثر بر متغیر پاسخ هستیم. این هدف، معمولا با مدل‌های رگرسیونی قابل دستیابی است. مقیاس این گونه پاسخ‌ها، پیوسته و محدود به فاصله $(۰, ۱)$ است. برای تحلیل رگرسیونی این نوع داده‌ها، اولین گزینه‌هایی که به ذهن می‌رسد استفاده از مدل‌های رگرسیونی لجستیک یا پروبیت است. اما از آن‌جا که توزیع این نوع پاسخ‌ها معمولا به شدت چوله است، این مدل‌ها برای آن‌ها مناسب نیستند. مدلی که با ماهیت داده‌های نرخ و نسبت سازگاری منطقی و مناسبی دارد، مدل رگرسیون بتا می‌باشد که به دلیل انعطاف بالای توزیع بتا در مدل‌بندی انواع چولگی‌ها، استفاده از آن در سال‌های اخیر پرطرفدار شده است. رگرسیون بتا یک عضو از رده مدل‌های خطی تعمیم‌یافته (GLM) است. یک پذیره اساسی در مدل‌های خطی تعمیم‌یافته، استقلال متغیرهای پاسخ است. اما در موارد متعددی، مانند داده‌های طولی، داده‌های فضایی و سری‌های زمانی، پاسخ‌ها به هم وابسته می‌باشند. در این موارد، از مدل‌های آمیخته خطی تعمیم‌یافته $(GLMM)$ استفاده می‌شود که وابستگی داده‌ها با وارد شدن متغیرهای پنهان به مدل لحاظ می‌شود. برازش این مدل‌ها از دیدگاه بیزی ساده‌تر است. این سادگی مرهون انقلاب روش‌های نمونه‌گیری $MCMC$ است. بنابراین علاقه‌مند شدیم پاسخ‌های وابسته با مقیاس پیوسته و تکیه‌گاه محدود را از دیدگاه بیزی تحلیل کنیم.

تقدیم به همه آنهایی که

می خواهند بیشتر بدانند

خدایا...

چه یافت آن که تو را گم کرد

و چه گم کرد آن که تو را یافت

سپاس گزاری...

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را زیور عقل آراست.
در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ استاد راهنمای خود، جناب آقای دکتر حسین باغیشنی، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید.

مهناز عجم
شهریور ۱۳۹۳

پیش‌گفتار

در بسیاری از مسایل تحلیل پاسخ‌های نرخ و نسبت مانند نرخ تورم، نرخ رشد، درصد مشتریان خوش قول در یک بنگاه اقتصادی و درصد یک بیماری در منطقه‌ای خاص، علاقه‌مند به کشف عوامل موثر بر متغیر پاسخ هستیم. این هدف، معمولا با مدل‌های رگرسیونی قابل دستیابی است. مقیاس این گونه پاسخ‌ها، پیوسته و محدود به فاصله $(۰,۱)$ است. برای تحلیل رگرسیونی این نوع داده‌ها، اولین گزینه‌هایی که به ذهن می‌رسد استفاده از مدل‌های رگرسیونی لجستیک یا پروبیت است. اما از آن‌جا که توزیع این نوع پاسخ‌ها معمولا به شدت چوله است، این مدل‌ها برای آن‌ها مناسب نیستند. مدلی که با ماهیت داده‌های نرخ و نسبت سازگاری منطقی و مناسبی دارد، مدل رگرسیون بتا می‌باشد که به دلیل انعطاف بالای توزیع بتا در مدل‌بندی انواع چولگی‌ها، استفاده از آن در سال‌های اخیر پرطرفدار شده است. رگرسیون بتا یک عضو از رده مدل‌های خطی تعمیم‌یافته (GLM) است. یک پذیره اساسی در مدل‌های خطی تعمیم‌یافته، استقلال متغیرهای پاسخ است. اما در موارد متعددی، مانند داده‌های طولی، داده‌های فضایی و سری‌های زمانی، پاسخ‌ها به هم وابسته می‌باشند. در این موارد، از مدل‌های آمیخته خطی تعمیم‌یافته $(GLMM)$ استفاده می‌شود که وابستگی داده‌ها با وارد شدن متغیرهای پنهان به مدل لحاظ می‌شود. استنباط مبتنی بر درستنمایی در مدل‌های آمیخته خطی تعمیم‌یافته، به دلیل وجود انتگرال‌های با بعد بالا برای محاسبه تابع درستنمایی، بسیار دشوار است. برازش این مدل‌ها از دیدگاه بیزی ساده‌تر است. این سادگی مرهون انقلاب روش‌های نمونه‌گیری $MCMC$ است که در آن برای محاسبه توزیع پسین نه نیازی به حل انتگرال‌های با بعد بالاست و نه محاسبه مشتق‌های پیچیده. بنابراین علاقه‌مند شدیم پاسخ‌های وابسته با مقیاس پیوسته و تکیه‌گاه محدود را از دیدگاه بیزی تحلیل کنیم. با توجه به این مقدمه، ساختار این پایان‌نامه به صورت زیر است:

- در فصل اول، تعاریف و مفاهیم مورد نیاز برای ورود به موضوع این پایان‌نامه را مطرح می‌کنیم.
- در فصل دوم، مدل رگرسیون آمیخته بتا را معرفی می‌کنیم و به برازش و استنباط آن از دیدگاه بیزی می‌پردازیم.

- در فصل سوم، با یک مطالعه شبیه‌سازی عملکرد مدل و روش پیشنهادی را مورد ارزیابی قرار می‌دهیم.
- در فصل چهارم، کاربرد مدل را با دو مثال واقعی شامل داده‌های نسبت بنزین تبدیل‌شده از نفت خام و میزان مقاومت بتن نمایش می‌دهیم.
- در پیوست آ، به معرفی نرم‌افزار $JAGS$ و بسته $rjags$ می‌پردازیم و سپس نحوه برازش و استنباط در مدل‌های رگرسیون آمیخته بتا را توسط $JAGS$ با یک مثال تشریح می‌کنیم. کدهای نوشته‌شده برای بازتولید مثال‌های پایان‌نامه، در محیط $JAGS$، نیز در پیوست ب آورده شده‌اند.

فهرست مطالب

فهرست تصاویر

فهرست جداول

فصل ۱

تعاریف و مفاهیم اولیه

۱.۱ مقدمه

یکی از پرکاربردترین روش‌ها در بین تکنیک‌های آماری، روش‌های تحلیل رگرسیون است. در حقیقت، تحلیل رگرسیونی روشی با بیشترین و وسیع‌ترین کاربرد در بین روش‌های آماری است. از دید آماری، رگرسیون به مفهوم بازگشت به یک مقدار متوسط یا میانگین می‌باشد. اما امروزه واژه تحلیل رگرسیونی جهت اشاره به مطالعات مربوط به روابط بین متغیرها به‌کار برده می‌شود.

تحلیل رگرسیونی ما را قادر می‌سازد که بین یک متغیر مورد نظر که آن را متغیر پاسخ می‌نامیم و یک یا چند متغیر تبیینی یا پیش‌بین، رابطه‌ای را به تجربه معین کنیم و از آن استفاده نماییم. به عنوان مثال، در یک مطالعه رگرسیونی، متغیر پاسخ، تعداد قوطی‌های روغن گریسی است که در طول یک دوره معین در منطقه‌ای مصرف شده است و متغیر تبیینی، شامل تعداد موسساتی در این منطقه است که این روغن گریس را مصرف می‌کنند.

تحلیل رگرسیونی را غالبا برای پیش‌گویی متغیر پاسخ از روی آگاهی‌هایی که از متغیرهای تبیینی دارند، به‌کار می‌برند. مثلا، در مطالعه روغن گریس، مایل بوده‌اند که مصرف روغن در یک ناحیه فروش جدید را از روی آگاهی‌هایی که از تعداد موسسات این ناحیه دارند، پیش‌گویی کنند. از این روش برای بررسی ماهیت وابستگی بین متغیرهای تبیینی و متغیر پاسخ نیز استفاده می‌کنند.

مدل‌های معمول رگرسیون، پذیره‌های خطی و نرمال بودن پاسخ را شامل می‌شوند. اما در بسیاری از موارد کاربردی، متغیرهای پاسخ دارای اندازه‌ای محدود به فاصله $(۰, ۱)$ هستند. این موارد معمولا شامل درصدها، نرخ‌ها و نسبت‌ها می‌باشند. به عنوان مثال نرخ تورم، نرخ رشد، درصد مشتریان خوش‌قول

در یک بنگاه اقتصادی و درصد یک بیماری در یک منطقه خاص. معمولا، برای مدل‌بندی این نوع داده‌ها، مدل‌های رگرسیونی لجستیک[۱] و پروبیت[۲] مورد استفاده قرار می‌گیرند. اما، به دلیل توزیع به شدت چوله این نوع پاسخ‌ها، این دو مدل نمی‌توانند به خوبی داده‌ها را برازش دهند. در این‌گونه موارد، مدل رگرسیونی بتا پیشنهاد می‌شود. مدل رگرسیون بتا عضوی از خانواده مدل‌های خطی تعمیم‌یافته[۳] ($GLMs$) است. در این خانواده استقلال متغیرهای پاسخ یک پذیره اساسی است. اما در موارد مختلفی مانند داده‌های طولی، سری‌های زمانی و داده‌های فضایی، پاسخ‌ها وابستگی دارند. در این‌گونه موارد، تعمیمی از $GLMs$ به نام مدل‌های آمیخته خطی تعمیم‌یافته[۴] ($GLMMs$) مورد استفاده قرار می‌گیرند. برازش این مدل‌ها از دیدگاه بسامدی سخت است. در حالی‌که در دیدگاه بیزی با مشکلات کمتری مواجه خواهیم شد. در ادامه به معرفی رده $GLMM$ و استنباط‌های بسامدی و بیزی می‌پردازیم.

۲.۱ مدل‌های آمیخته خطی تعمیم‌یافته

تقریبا در اغلب موارد استفاده از روش‌های آماری، مرکز توجه روی میانگین‌ها و تغییرات آن‌هاست، که منجر به ایجاد روش‌های استنباط درباره میانگین‌ها یا درباره ماهیت آن‌ها می‌شود. معمول‌ترین مدل رگرسیونی، مدل خطی[۵] (LM) است. در یک مدل خطی، با متغیر پاسخ Y و ماتریس طرح X، تابع رگرسیونی، که همان میانگین شرطی $Y|X$ است، بر حسب ترکیب خطی از X به‌صورت $E[Y|X] = X\beta$ در نظر گرفته می‌شود. در این مدل، ضرایب β بردار ثابت‌های نامعلوم، یا به عبارتی اثرات ثابت[۶] نامیده می‌شوند.

یکی از پذیره‌های مدل‌های خطی، استقلال بین مشاهدات است. زمانی که برای پاسخ‌ها این پذیره برقرار نیست، تعمیمی از این مدل‌ها معروف به مدل‌های آمیخته خطی[۷] ($LMMs$) مورد استفاده قرار

---

[۱] Logistic

[۲] Probit

[۳] Generalized Linear Models

[۴] Generalized Linear Mixed Models

[۵] Linear Model

[۶] Fixed Effects

[۷] Linear Mixed Models

می‌گیرد. برای در نظر گرفتن وابستگی بین پاسخ‌ها، استفاده از متغیرهای پنهان[۸] یا همان اثرات تصادفی[۹] یک راه رایج و مناسب است. میانگین پاسخ در یک LMM به صورت $E[Y|X,Z,b] = X\beta + Zb$ نمایش داده می‌شود که در آن X ماتریس طرح متناظر با اثرات ثابت β و Z ماتریس طرح متناظر با بردار اثرات تصادفی b می‌باشند (مک‌کالاک و سیرل، ۲۰۰۱). اثرات تصادفی معمولا به عنوان متغیرهای تصادفی نرمال (با میانگین صفر) تلقی می‌شوند و پارامترهایی که توزیع این اثرات را توصیف می‌کنند، برآورد می‌شوند.

در بسیاری از مسایل، مانند پاسخ‌های دودویی و شمارشی، استفاده از توزیع نرمال و مدل خطی مناسب نیست. زیرا مدل‌های خطی محدودیت موجود در میانگین پاسخ برای این داده‌ها را در نظر نمی‌گیرند. به عنوان مثال برای داده‌های دودویی، میانگین که همان احتمال موفقیت می‌باشد، در فاصله $(۰, ۱)$ قرار می‌گیرد و چنان‌چه از مدل خطی استفاده کنیم ممکن است برآورد میانگین پاسخ هر مقداری در مجموعه اعداد حقیقی باشد. در نتیجه برای این داده‌ها با فرض توزیع برنولی، معمولا، از رگرسیون پروبیت یا لجستیک استفاده می‌شود. برای مرتفع ساختن این مشکل نلدر و ودربرن (۱۹۷۲) مدل‌های خطی تعمیم‌یافته را معرفی کردند. این مدل‌ها تعمیمی از مدل‌های خطی هستند که برای تحلیل پاسخ‌های عضو خانواده توزیع‌های نمایی گسترش یافته‌اند. ذات این تعمیم در این‌گونه مدل‌ها دو چیز است: اول، متغیرهای پاسخ می‌توانند غیرنرمال باشند؛ دوم، میانگین پاسخ لزوما به صورت ترکیبی خطی از متغیرهای تبیینی نیست، بلکه تابعی از آن به نام تابع پیوند[۱۰] به صورت ترکیب خطی از متغیرهای تبیینی در نظر گرفته می‌شود.

تعریف ۱.۲.۱. خانواده‌ای از توزیع‌ها را زیررده‌ای از خانواده توزیع‌های نمایی گوییم، هرگاه بتوان تابع چگالی آن را به صورت زیر نوشت:

$$f_Y(y) = \exp\left\{\frac{y\gamma - \psi(\gamma)}{a(\phi)} - c(y,\phi)\right\}, \qquad (۱.۱)$$

که در آن ϕ و γ پارامترهای خانواده توزیع و $c(\cdot)$، $\psi(\cdot)$ و $a(\cdot)$ توابعی معلوم هستند. بنابراین، در یک خانواده نمایی، میانگین و واریانس به صورت

$$E(Y) = \frac{\partial\psi(\gamma)}{\partial(\gamma)} = \mu, \quad Var(Y) = a(\phi)\frac{\partial^۲\psi(\gamma)}{\partial(\gamma)^۲}, \qquad (۲.۱)$$

محاسبه می‌شوند.

---

[۸] Latent Variables

[۹] Random Effects

[۱۰] Link Function

در یک GLM، اگر η رابطه‌ای خطی از متغیرهای تبیینی و ضرایب رگرسیونی، به نام پیشگوی خطی[۱۱] باشد، آن‌گاه

$$g(\mu) = \eta = X\beta,$$

که $g(\cdot)$ یک تابع پیوند است. این تابع با پیوند میانگین پاسخ و پیشگوی خطی، میانگین را مدل‌بندی می‌کند. از جمله توابع پیوند مورد استفاده در $GLMs$، توابع پیوند لجیت، $g(\mu) = \ln \frac{\mu}{۱-\mu}$، پروبیت، $g(\mu) = \Phi^{-۱}(\mu)$ و لگاریتمی، $g(\mu) = \ln(\mu)$، می‌باشند. برای مقایسه این توابع به مک‌کالاک و نلدر (۱۹۸۹) مراجعه کنید.

مثال ۲.۲.۱. یک مثال بارز از خانواده توزیع‌های نمایی، توزیع نرمال است که چگالی آن به صورت زیر می‌باشد:

$$f(y_{ij}|b_i) = \exp\left\{\frac{y_{ij}\mu_{ij} - \mu_{ij}^{۲}/۲}{\sigma^{۲}} - \frac{۱}{۲}\left(\frac{y_{ij}^{۲}}{\sigma^{۲}} + \ln ۲\pi\sigma^{۲}\right)\right\},$$

که در آن، $\gamma_{ij} = \mu_{ij}$، $\phi = \sigma$، $\psi(\gamma_{ij}) = \frac{\mu_{ij}^{۲}}{۲}$ و $c(y_{ij}, \phi) = -\frac{۱}{۲}(\frac{y_{ij}^{۲}}{\sigma^{۲}} + \ln ۲\pi\sigma^{۲})$. همچنین، میانگین و واریانس توزیع نرمال طبق روابط (۲.۱) عبارتند از:

$$E(y_{ij}|\eta_{ij}) = \mu_{ij},$$

$$Var(y_{ij}|\eta_{ij}) = \sigma^{۲}.$$

یک تابع پیوند مناسب برای میانگین پاسخ، با توجه به این‌که $\mu_{ij} \in R$، تابع پیوند همانی[۱۲] است. یعنی $g(\mu_{ij}) = \mu_{ij}$ بنابراین $\mu_{ij} = \eta_{ij} \in R$.

از آن‌جایی که پذیره استقلال در بین مشاهدات غیرنرمال همیشه برقرار نیست، پس به سراغ مدل‌های آمیخته خطی تعمیم‌یافته (بریسلو و کلیتون، ۱۹۹۳) خواهیم رفت. این مدل‌ها ترکیبی طبیعی از دو رده $LMMs$ و $GLMs$ هستند. در واقع با افزودن اثرات تصادفی به پیشگوی خطی در یک GLM، یک خانواده از مدل‌های منعطف که در موقعیت‌های عملی مختلفی کاربرد دارند، تشکیل می‌شود.

متغیر پاسخ در یک $GLMM$ مستقل شرطی بوده و دارای توزیعی از خانواده نمایی با تابع چگالی احتمال (۱.۱) می‌باشد و به‌طور مشابه، میانگین و واریانس آن از رابطه (۲.۱) به‌دست خواهند آمد. اگر میانگین شرطی برابر با $E[Y|X, Z, b] = \mu$ باشد، آن‌گاه تابع پیوند روی این میانگین به عنوان ترکیبی خطی از هر دو عامل ثابت و تصادفی به مدل معرفی خواهد شد:

$$g(\mu) = X\beta + Zb.$$

[۱۱] Linear Predictor

[۱۲] Identity Link Function

یک استفاده معمول از $GLMMs$ در تحلیل داده‌های طولی[۱۳] است.

تعریف ۳.۲.۱. اگر یک داده متعادل[۱۴] را در نظر بگیرید، یعنی این‌که برای هر عضو i $(i = ۱, \ldots, m)$ تعداد یکسانی مشاهده (n) وجود داشته باشد، آن‌گاه مجموعه داده‌ها را داده طولی می‌نامیم هرگاه بتوان برای $i = ۱, \ldots, m$ و $j = ۱, \ldots, n$ بردار داده‌ها را روی کل m عضو به صورت

$$\mathbf{y} = \left\{\mathbf{y}_i\right\}_{i=۱}^{m} = \left\{\left\{y_{ij}\right\}_{j=۱}^{n}\right\}_{i=۱}^{m},$$

نوشت. برای عضو iام، مشاهده در زمان jام با y_{ij} نشان داده می‌شود. بنابراین بردار داده‌های عضو iام عبارتست از

$$\mathbf{y}_i = [y_{i۱} \quad y_{i۲} \ldots y_{ij} \ldots y_{in}]^T = \left\{y_{ij}\right\}_{j=۱}^{n}.$$

به عنوان مثال‌هایی از داده‌های طولی می‌توان به اندازه فشار خون گرفته‌شده از صد نفر با گروه‌های سنی مختلف در ده روز مشخص (که در آن متغیر پاسخ y_{ij} اندازه فشار خون گرفته‌شده از عضو iام، $i = ۱, \ldots, ۱۰۰$، در زمان jام، $j = ۱, \ldots, ۱۰$، می‌باشد) و اندازه‌گیری میزان مقاومت بتن در سه واحد زمانی مختلف، اشاره کرد.

فرض کنید در یک داده طولی، X_i ماتریس طرح $n_i \times p$ متناظر با بردار $\boldsymbol{\beta} = (\beta_۱, \ldots, \beta_p)^T$ و Z_i ماتریس طرح $n_i \times q$ متناظر با بردار $\mathbf{b}_i = (b_{i۱}, \ldots, b_{iq})^T$ باشند. بنابراین اثرات تصادفی روی هر عضو آن، i، مقدار متفاوتی را در بر خواهد داشت که وابستگی بین اعضا با شرطی شدن روی این اثرات از بین خواهد رفت. با این اوصاف، مدل آمیخته خطی تعمیم‌یافته برای یک داده طولی عبارتست از

$$g(E[y_{ij}|\mathbf{b}_i, x_{ij}, z_{ij}]) = x_{ij}^T\boldsymbol{\beta} + z_{ij}^T\mathbf{b}_i, \quad i = ۱, \ldots, m, \quad j = ۱, \ldots, n_i,$$

که در آن، m تعداد افراد و n_i تعداد تکرارهای پاسخ برای فرد iام است و معمولا اثرات تصادفی $\mathbf{b}_i$ دارای توزیع نرمال چندمتغیره با میانگین صفر می‌باشند. با این وجود، متغیرهای پاسخ، y_{ij}، مشروط به این اثرات، مستقل شرطی بوده و دارای توزیعی از خانواده توزیع‌های نمایی به شکل (۱.۱) خواهند بود که در آن، پارامتر متعارف[۱۵] γ_{ij} تابعی از پیشگوی خطی به صورت $\eta_{ij} = x_{ij}^T\boldsymbol{\beta} + z_{ij}^T\mathbf{b}_i$ است.

---

[۱۳] Longitudinal Data

[۱۴] Balanced Data

[۱۵] Canonical Parameter

۳.۱ دیدگاه‌های بسامدی و بیزی در استنباط آماری

روش‌های معمول استنباط بسامدی، روش‌های مبتنی بر درستنمایی هستند که برای محاسبه تابع درستنمایی در رده $GLMMs$، حل عددی انتگرال‌های با بعد بالا را شامل می‌شوند که دارای محاسبات زمان‌بر و پیچیده‌ای می‌باشند. تقریب این نوع انتگرال‌ها با دو رهیافت صورت می‌پذیرد. رهیافت اول، شامل روش‌هایی است که انتگرال را به‌صورت عددی تقریب می‌زنند، مانند روش‌های مربع‌بندی[۱۶] (پینیرو وبیتس، ۱۹۹۵). رهیافت دوم، روش‌هایی هستند که تابع زیر انتگرال را طوری تقریب می‌زنند که انتگرال آن صورت بسته داشته باشد، از جمله روش‌های تقریب لاپلاس[۱۷] (بریسلو و کلیتون، ۱۹۹۳؛ بریسلو و لین، ۱۹۹۵) و شبه‌درستنمایی تاوانیده[۱۸] (بریسلو و کلیتون، ۱۹۹۳).

هر یک از روش‌های ذکرشده، دارای معایب قابل توجهی هستند. مشکل اساسی روش‌های انتگرال‌گیری عددی در شناسایی نواحی مهم (نواحی چگال‌تر) برای تابعی که باید انتگرال‌گیری شود، می‌باشد. انتگرال‌گیری‌های عددی برای نواحی با بعد حداکثر ۶ خوب عمل می‌کنند و در غیر این‌صورت با خطای زیادی مواجه می‌شوند، زیرا با افزایش بعد انتگرال، با تجمیع خطاها مواجه می‌شویم. افزایش بعد انتگرال، افزایش محاسبات را نیز به همراه خواهد داشت (بریسلو و کلیتون، ۱۹۹۳).

با توجه به مشکلات موجود در دیدگاه استنباط بسامدی در این رده از مدل‌ها، دیدگاه بیزی، که استنباط در آن مبتنی بر توزیع پسین است، می‌تواند یک جایگزین مناسب باشد. توزیع پسین در یک مدل بیزی، به‌روز شده باور و دیدگاه اولیه در مورد پارامتر با توجه به مشاهدات است. نحوه تشکیل این توزیع برای یک نمونه معمول nتایی، به صورت زیر می‌باشد.

فرض کنید $Y_1, \ldots, Y_n$ یک نمونه تصادفی nتایی با تابع چگالی احتمال $f(\cdot|\theta)$ باشد، که در آن θ پارامتر مورد علاقه می‌باشد و مقادیر مختلف خود را در مجموعه فضای پارامتر Θ اختیار می‌کند. در یک مدل بیزی، فرض می‌شود مقدار واقعی پارامتر θ نیز تحققی از یک توزیع احتمالی به نام توزیع پیشین است. اگر این توزیع را با $\pi(\theta)$ و تابع درستنمایی را با

$$L(\theta|y) = f_{y|\theta}(y_1, \ldots, y_n|\theta),$$

نشان دهیم آن‌گاه بنابر قاعده بیز، توزیع پسین $\pi(\theta|y)$ معادل است با

$$\pi(\theta|y) = \frac{L(\theta|y)\pi(\theta)}{\int_\Theta L(\theta|y)\pi(\theta)\mathrm{d}\theta} \propto c(y)L(\theta|y)\pi(\theta),$$

[۱۶] Quadrature Methods

[۱۷] Laplace Approximation

[۱۸] Penalized Quasi Likelihood

که در آن $c(y)$ ثابت نرمال‌ساز نامیده می‌شود و برابر است با

$$c(y) = \frac{۱}{\int_\Theta L(\theta|y)\pi(\theta)\mathrm{d}\theta}.$$

رهیافت بیزی برای تقریب انتگرال‌های با بعد بالا از روش‌های مبتنی بر نمونه‌گیری (نمونه‌گیری از توزیع پسین) استفاده می‌کند. از جمله این روش‌ها می‌توان به روش‌های نمونه‌گیری نقاط مهم[۱۹] و مونت کارلوی زنجیر مارکوفی[۲۰] ($MCMC$) اشاره کرد. در این روش‌ها، قانون قوی اعداد بزرگ و قضیه ارگودیک[۲۱] (بیرخوف، ۱۹۴۲) ضامن همگرایی برآوردگرهای حاصل از نمونه‌ها به کمیت‌های مورد علاقه از توزیع پسین می‌باشند. روش نمونه‌گیری نقاط مهم برای انتگرال‌های با بعد کوچک مورد استفاده قرار می‌گیرد، در حالی‌که روش‌های $MCMC$ برای مسایل انتگرال‌گیری با بعد بالا نیز می‌توانند مفید باشند به‌طوری که دیگر نیازی به حل انتگرال‌های با بعد بالا، ماکسیمم‌سازی و مشتق‌گیری از یک تابع پیچیده نیست.

۴.۱ روش‌های نمونه‌گیری $MCMC$

هدف اصلی در روش‌های انتگرال‌گیری مونت کارلویی، ارایه تقریبی مناسب برای انتگرال‌های از نوع $\tau = \int g(x)\mathrm{d}x$ می‌باشد. این انتگرال را می‌توان به صورت

$$\tau = E_f[h(X)] = \int h(x)f(x)\mathrm{d}x,$$

نوشت که در آن، $f(\cdot)$ یک تابع چگالی و $h(\cdot)$ تابعی است که متناسب با $g(\cdot)$ تعریف می‌شود. حل مونت کارلویی این انتگرال، نیازمند تولید نمونه $X_۱, \ldots, X_m$ از $f(x)$ می‌باشد. با محاسبه میانگین نمونه‌ای $\bar{h}_m = \frac{۱}{m}\sum_{i=۱}^m h(X_i)$ و بنابر قانون قوی اعداد بزرگ، چنان‌چه $m \to \infty$، همگرایی به مقدار انتگرال، τ، تضمین می‌شود، یعنی

$$\bar{h}_m \xrightarrow{p} E_f[h(X)].$$

روش‌های نمونه‌گیری $MCMC$، شامل روش‌هایی است که برای انتگرال‌گیری به روش مونت کارلویی، اغلب در مسایل مربوط به استنباط بیزی، مورد استفاده قرار می‌گیرند. در یک مدل بیزی، با شرط مفروض بودن چگالی پیشین، $\pi(\theta)$ و چگالی توام $f(x,\theta)$، (x,θ)، مساله انتگرال‌گیری به امید ریاضی شرطی تابعی مانند $g(\theta)$ نسبت به توزیع پسین $\pi(\theta|x)$ تبدیل می‌شود، یعنی

$$E[g(\theta)|x] = \int_\Theta g(\theta)\pi(\theta|x)\mathrm{d}\theta.$$

[۱۹] Importance Sampling

[۲۰] Markov Chain Monte Carlo

[۲۱] Ergodic Theorem

روش‌های انتگرال‌گیری $MCMC$، به‌ویژه الگوریتم‌های متروپولیس هستینگس[۲۲] (متروپولیس و همکاران، ۱۹۵۳ و هستینگز، ۱۹۷۰) و نمونه‌گیر گیبز[۲۳] (گمان و گمان، ۱۹۸۴ و گلفاند و اسمیت، ۱۹۹۰) به عنوان ابزار بسیار مفید برای تحلیل مدل‌های آماری پیچیده پدید آمده‌اند.

روش‌های معمول نمونه‌گیری $MCMC$، از جمله متروپولیس–هستینگس دارای مشکلات متعددی نظیر تشخیص همگرایی زنجیر و زمان طولانی محاسبات می‌باشند. اما نمونه‌گیر گیبز این‌گونه مسایل را به دنباله‌ای از مسایل ساده‌تر تقسیم می‌کند و اجرای آن ساده‌تر است. البته ممکن است این دنباله‌ها زمان زیادی برای همگرایی نیاز داشته باشند ولی با این وجود، این الگوریتم مفید و جالب‌توجه است، زیرا این اجازه را می‌دهد که یک نمونه از توزیع پسین توام تولید کنیم. در ادامه دو الگوریتم متروپولیس–هستینگس و نمونه‌گیر گیبز را به‌طور مختصر شرح می‌دهیم. قبل از تشریح این الگوریتم‌ها، چند تعریف را بیان می‌کنیم.

تعریف ۱.۴.۱ (فرآیند تصادفی). اگر t متغیر شاخص (که معمولا بیانگر متغیر زمان است) و $X^{(t)}$ متغیر متناسب با t باشد، آن‌گاه یک فرآیند تصادفی، مجموعه‌ای از متغیرهای تصادفی $\{X^{(t)} : t \in T\}$ است که در آن $X^{(t)}$ به ازای هر $t \in T$ یک متغیر تصادفی است و T مجموعه اندیس‌گذار فرآیند نامیده می‌شود.

تعریف ۲.۴.۱ (زنجیر مارکوف). فرآیند تصادفی $\{X^{(t)}\}$ را یک زنجیر مارکوف می‌گوییم، هرگاه

$$\begin{aligned} P(X^{(i+1)} = a|X^{(i)}, X^{(i-1)}, \ldots, X^{(0)}) &= P(X^{(i+1)} = a|X^{(i)}) \\ &= K(X^{(i)}, X^{(i+1)}), \end{aligned}$$

که در آن $K(\cdot,\cdot)$ را هسته انتقال[۲۴] یا هسته مارکوف می‌نامند. بنابراین یک زنجیر مارکوف، دنباله‌ای از متغیرهای تصادفی است که با فرض معلوم بودن $X^{(i)}$، که حالت فرآیند در زمان فعلی iام است، متغیر $X^{(i+1)}$ که حالت بعدی فرآیند است مستقل از حالت‌های گذشته، یعنی $X^{(i-1)}, \ldots, X^{(0)}$ می‌باشد. از آن‌جا که نمونه بعدی در هر زنجیر فقط به نمونه قبلی آن وابسته است، این عدم وابستگی به گذشته اجازه می‌دهد تا زنجیرهای مارکوف برای ساده‌سازی مسایل پیچیده مورد استفاده قرار گیرند.

تعریف ۳.۴.۱ (زنجیر مارکوف مانا). زنجیر مارکوف $\{X^{(t)}\}$ یک زنجیر مارکوف مانا است، هرگاه

[۲۲] Metropolis-Hastings Algorithms

[۲۳] Gibbs Sampler

[۲۴] Transition Kernel

دارای یک توزیع مانای[۲۵] $f(\cdot)$ باشد به طوری‌که اگر $X^{(t)} \sim f(\cdot)$ آن‌گاه $X^{(t+۱)} \sim f(\cdot)$.

۵.۱ الگوریتم متروپولیس-هستینگس

الگوریتم متروپولیس-هستینگس مثالی از مجموعه روش‌های نمونه‌گیری $MCMC$ است. این الگوریتم، با شرط داشتن توزیع پسین توام $\pi(\theta|\mathbf{x})$، مجموعه‌ای از متغیرهای تصادفی $\{\theta^{(t)}\}$، برای $t = ۱, ۲, \ldots$، را به‌طور پی‌درپی تولید می‌کند. برای تولید این نمونه‌ها، این الگوریتم، ما را ملزم به تعیین چگالی پیشنهادی $p(\theta^{(t+۱)}, \theta^{(t)})$ می‌کند. این تابع، تابع چگالی احتمال θ در زمان $t+۱$، با توجه به مقدار آن در زمان t است. بنابراین عملکرد الگوریتم متروپولیس-هستینگس به صورت زیر بیان می‌شود:
با شرط داشتن $\theta^{(t)}$،

۱. نمونه $\theta^{(t+۱)} \sim p(\theta^{(t+۱)}, \theta^{(t)})$ را تولید کن.

۲. قرار بده

$$\boldsymbol{\theta}^{(t)} = \begin{cases} \theta^{(t+۱)} & \rho(\theta^{(t)}, \theta^{(t+۱)}), \\ \theta^{(t)} & ۱ - \rho(\theta^{(t)}, \theta^{(t+۱)}), \end{cases}$$

به طوری‌که

$$\rho(\theta^{(t)}, \theta^{(t+۱)}) = min\left\{۱, \frac{f(\mathbf{x}|\theta^{(t+۱)})\pi(\theta^{(t+۱)})}{f(\mathbf{x}|\theta^{(t)})\pi(\theta^{(t)})} \frac{p(\theta^{(t)}, \theta^{(t+۱)})}{p(\theta^{(t+۱)}, \theta^{(t)})}\right\},$$

که به آن احتمال پذیرش متروپولیس-هستینگس می‌گویند.

اجرای یک الگوریتم متروپولیس-هستینگس با مشکلاتی مواجه است. در مواردی ممکن است تعیین چگالی پیشنهادی دشوار باشد. یک روش جایگزین، کاوش در همسایگی مقدار جاری زنجیر می‌باشد. اجرای این ایده، شبیه‌سازی θ در زمان $t+۱$ از رابطه $\theta^{(t+۱)} = \theta^{(t)} + \varepsilon$ است، که در آن ε یک متغیر تصادفی مستقل از θ با میانگین صفر است. با شرط متقارن بودن توزیع ε، و در نتیجه توزیع پیشنهادی، یعنی $p(\theta^{(t+۱)}, \theta^{(t)}) = p(\theta^{(t)}, \theta^{(t+۱)})$، الگوریتم را الگوریتم متروپولیس-هستینگس قدم زدن تصادفی[۲۶] می‌نامند و احتمال پذیرش آن به صورت

[۲۵] Stationary Distribution

[۲۶] Random Walk Metropolis-Hastings Algorithm

$$\rho(\theta^{(t)}, \theta^{(t+۱)}) = min\left\{۱, \frac{f(\mathbf{x}|\theta^{(t+۱)})\pi(\theta^{(t+۱)})}{f(\mathbf{x}|\theta^{(t)})\pi(\theta^{(t)})}\right\},$$

تعریف خواهد شد. برای آگاهی بیشتر وعمیق‌تر از جزییات این الگوریتم و الگوریتم‌های مشابه به رابرت و کسلا (۲۰۰۴) مراجعه کنید.

۶.۱ نمونه‌گیر گیبز

بیشترین کاربرد نمونه‌گیر گیبز در مدل‌های بیزی، زمانی است که توزیع‌های شرطی کامل پارامترهای مورد علاقه وجود دارند. برای تولید نمونه از توزیع پسین توام $\pi(\boldsymbol{\theta}|\mathbf{x})$، که در آن $\boldsymbol{\theta} = (\theta_۱, \ldots, \theta_B)$ بردار پارامترهای مدل است، ابتدا باید چگالی‌های شرطی کامل را به‌دست آوریم. این چگالی‌های شرطی عبارتند از:

$$\begin{aligned}
&\pi(\theta_۱|\mathbf{x}, \theta_۲, \ldots, \theta_B),\\
&\pi(\theta_۲|\mathbf{x}, \theta_۱, \theta_۳, \ldots, \theta_B),\\
&\vdots\\
&\pi(\theta_B|\mathbf{x}, \theta_۱, \ldots, \theta_{B-۱}).
\end{aligned}$$

نمونه‌گیر گیبز با فرض داشتن $\boldsymbol{\theta}^{(t)} = (\theta_۱^{(t)}, \ldots, \theta_B^{(t)})$ در مرحله tام، $t = ۱, ۲, \ldots$، برای تولید یک زنجیر مارکوف از توزیع پسین توام $\pi(\boldsymbol{\theta}|\mathbf{x})$، دنباله‌ای از نمونه‌ها را به صورت زیر تولید می‌کند:

$$\begin{aligned}
&۱ : \theta_۱^{(t+۱)} \sim \pi(\theta_۱|\mathbf{x}, \theta_۲^{(t)}, \ldots, \theta_B^{(t)}),\\
&۲ : \theta_۲^{(t+۱)} \sim \pi(\theta_۲|\mathbf{x}, \theta_۱^{(t+۱)}, \theta_۳^{(t)}, \ldots, \theta_B^{(t)}),\\
&\vdots\\
&B : \theta_B^{(t+۱)} \sim \pi(\theta_B|\mathbf{x}, \theta_۱^{(t+۱)}, \ldots, \theta_{B-۱}^{(t+۱)}).
\end{aligned}$$

ثابت شده است (رابرت و کسلا، ۲۰۰۴) که نمونه تولیدشده توسط الگوریتم بالا، نمونه‌ای از توزیع پسین توام $\pi(\boldsymbol{\theta}|\mathbf{x})$ است.

مثال ۱.۶.۱. فرض کنید برای یک نمونه nتایی، مدل بیزی زیر در نظر گرفته شود:

$$X_i \sim N(\theta, \sigma^۲), \quad i = ۱, \dots, n,$$
$$\theta \sim N(\theta_۰, \tau^۲), \quad \sigma^۲ \sim IGamma(a, b),$$

که در آن $\theta_۰$، $\tau^۲$، a و b مقادیری معلوم هستند. بنابراین توزیع پسین $(\theta, \sigma^۲)$ عبارتست از

$$\pi(\theta, \sigma^۲|\mathbf{X}) \propto \left[\frac{۱}{(\sigma^۲)^{n/۲}} \mathrm{e}^{-\sum_i (x_i-\theta)^۲/(۲\sigma^۲)}\right] \times \left[\frac{۱}{\tau} \mathrm{e}^{-(\theta-\theta_۰)^۲/(۲\tau^۲)}\right] \times \left[\frac{۱}{(\sigma^۲)^{a+۱}} \mathrm{e}^{۱/b\sigma^۲}\right].$$

بنابراین توزیع‌های شرطی کامل به صورت

$$\begin{aligned}
\pi(\theta|\mathbf{X}, \sigma^۲) &\propto \mathrm{e}^{-\sum_i (x_i-\theta)^۲/(۲\sigma^۲)} \mathrm{e}^{-(\theta-\theta_۰)^۲/(۲\tau^۲\sigma^۲)} \\
&\propto \mathrm{e}^{\theta^۲(\frac{-۱}{۲\tau^۲} - \frac{n}{۲\sigma^۲}) + \theta(\frac{\theta_۰}{\tau^۲} + \frac{n\bar{x}}{\sigma^۲})} \\
&\propto \mathrm{e}^{\frac{-۱}{۲}(\frac{۱}{\tau^۲} + \frac{n}{\sigma^۲})\left(\theta^۲ - ۲\theta\left(\frac{\frac{\theta_۰}{۲\tau^۲} + \frac{n\bar{x}}{۲\sigma^۲}}{\frac{۱}{۲\tau^۲} + \frac{n}{۲\sigma^۲}}\right)\right)} \\
&\propto \mathrm{e}^{\frac{-۱}{۲}(\frac{\sigma^۲ + n\tau^۲}{\sigma^۲\tau^۲})\left(\theta^۲ - ۲\theta\left(\frac{\sigma^۲\theta_۰ + n\tau^۲\bar{x}}{\sigma^۲ + n\tau^۲}\right)\right)}, \\
\pi(\sigma^۲|\mathbf{X}, \theta) &\propto \left(\frac{۱}{\sigma^۲}\right)^{(n+۲a+۳)/۲} \mathrm{e}^{-\frac{۱}{۲\sigma^۲}(\sum_i (x_i-\theta)^۲ + (\theta-\theta_۰)^۲/\tau^۲ + ۲/b)},
\end{aligned}$$

قابل محاسبه هستند. با توجه به هسته چگالی‌های به‌دست آمده، به راحتی می‌توان نتیجه گرفت

$$\theta|\mathbf{X}, \sigma^۲ \sim N\left(\frac{\sigma^۲}{\sigma^۲ + n\tau^۲}\theta_۰ + \frac{n\tau^۲}{\sigma^۲ + n\tau^۲}\bar{x}, \frac{\sigma^۲\tau^۲}{\sigma^۲ + n\tau^۲}\right),$$
$$\sigma^۲|\mathbf{X}, \theta \sim IGamma\left(\frac{n}{۲} + a, \frac{۱}{۲}\sum_i (x_i - \theta)^۲ + b\right).$$

اکنون با داشتن این چگالی‌ها و به کمک نمونه‌گیر گیبز، می‌توان یک نمونه دلخواه از توزیع پسین توام $\pi(\theta, \sigma^۲|\mathbf{X})$ تولید کرد.

۷.۱ همگرایی یک زنجیر مارکوف

در کنار استفاده شایع از روش‌های نمونه‌گیری $MCMC$ در استنباط‌های بیزی، دو چالش مهم تشخیص همگرایی و ویژگی‌های آمیختگی[۲۷] زنجیر و زمان طولانی محاسبات مطرح هستند. برای شناسایی

[۲۷] Mixing Properties

همگرایی زنجیرهای $MCMC$، معیارهای مختلفی پیشنهاد شده‌اند. اما ابتدا همگرایی یک زنجیر مارکوف را تشریح می‌کنیم.

همان‌طور که قبلا اشاره شد، در روش‌های انتگرال‌گیری مونت کارلویی، مساله انتگرال‌گیری به مساله یافتن راهی برای تولید نمونه از $f(\cdot)$ تبدیل می‌شود. در بسیاری از موارد، تولید نمونه مستقیم و مستقل از چگالی هدف $f(\cdot)$ مشکل است. در این موارد می‌توان با استفاده از یک زنجیر مارکوف مانا که توزیع مانای آن $f(\cdot)$ است، به‌طور غیر مستقیم یک نمونه تولید کرد. اما از آن‌جایی که نمونه تولیدشده از زنجیر مارکوف وابسته است، کیفیت یک نمونه مستقل با حجمی مشابه را ندارد. کیفیت چنین نمونه‌ای را می‌توان از طریق کمیتی به نام حجم نمونه موثر[۲۸] (ESS) سنجید (گلمن و همکاران، ۱۹۹۸؛ رابرت و کسلا، ۲۰۰۴). هرچه مقدار ESS محاسبه‌شده برای هر پارامتر به تعداد نمونه‌های تولیدشده نزدیک‌تر باشد، نمونه‌های تولیدشده از کیفیت بالاتری برخوردار هستند.

تعریف ۱.۷.۱. اگر n تعداد نمونه‌های تولیدشده از زنجیر باشد، آن‌گاه حجم نمونه موثر برای پارامتر مورد نظر به صورت

$$ESS = \frac{n}{1 + 2\sum_{k=1}^{\infty} \rho_k},$$

محاسبه می‌شود که در آن ρ_k مقدار خودهمبستگی در گام k است. در واقع ESS کمیتی است که تعداد نمونه‌های مستقل به‌دست آمده از زنجیر را تخمین می‌زند.

در مباحث نظری زنجیرهای مارکوف، اگر زنجیر دارای یک توزیع مانای $f(\cdot)$ باشد، می‌توان نمونه‌ها را از آن تولید کرد. در نتیجه، در یک زنجیر، هسته مارکوف K و توزیع مانای $f(\cdot)$ در رابطه زیر صدق می‌کنند (رابرت و کسلا، ۲۰۰۴):

$$\int_{\chi} K(x, y) f(x) \mathrm{d}x = f(y).$$

وجود یک توزیع مانا، ویژگی نافروکاستنی[۲۹] را بر هسته K تحمیل می‌کند. یعنی، K، صرفنظر از مقدار اولیه $X^{(0)}$، اجازه حرکت بر روی تمام وضعیت‌های فضای حالت زنجیر را خواهد داشت. از دیگر نتایج وجود یک توزیع مانا در زنجیرهای مارکوف، بازگشتی[۳۰] بودن زنجیر است. بازگشتی بودن به این معنی است که زنجیر می‌تواند به هر مجموعه دلخواه، به تعداد نامتناهی بازگردد. در یک زنجیر بازگشتی، توزیع مانا یک توزیع حدی[۳۱] است. یعنی توزیع $X^{(t)}$، به ازای هر مقدار اولیه $X^{(0)}$، به توزیع مانای

[۲۸] Effective Sample Size

[۲۹] Irreduciblity

[۳۰] Recurrency

[۳۱] Limiting Distribution

$f(\cdot)$ همگرا می‌شود. این ویژگی، ارگودیک بودن نیز نامیده می‌شود. بنابراین، بنابر قضیه ارگودیک، همگرایی به صورت زیر حاصل می‌شود:

$$\frac{1}{T}\sum_{t=1}^{T} h(X^{(t)}) \xrightarrow{p} E_f[h(X)].$$

با توجه به مطالب فوق می‌توان گفت، رهیافت $MCMC$ برای نمونه‌گیری از چگالی هدف $f(\cdot)$، یک زنجیر مارکوف $X^{(t)}$، با مقدار اولیه $X^{(0)}$، را با استفاده از یک هسته مارکوف K با توزیع مانای $f(\cdot)$ تولید می‌کند و سپس برای یک مقدار به اندازه کافی بزرگ T_0، نمونه‌های $X^{(T_0)}, X^{(T_0+1)}, \ldots$ را تولید می‌کند که از توزیع $f(\cdot)$ گرفته شده‌اند. بنابراین، بررسی همگرایی زنجیر به توزیع مانای $f(\cdot)$، یکی از مباحث مهمی است که در روش‌های نمونه‌گیری $MCMC$ مطرح می‌شود.

شبیه‌سازی از یک زنجیر $MCMC$ به دو بخش تقسیم می‌شود: پیش از همگرایی زنجیر و بعد از همگرایی. بخش پیش از همگرایی که به دوره داغیدن[۳۲] مشهور است، به عنوان بخشی از زنجیر است که در آن نمونه‌های تولیدشده را نمی‌توان به عنوان یک معرف خوب از توزیع پسین در نظر گرفت. به عبارت ساده‌تر، نمونه‌های $X^{(0)}, X^{(1)}, \ldots, X^{(T_0-1)}$ باید نادیده گرفته شوند و $X^{(T_0)}$ به عنوان تحققی از توزیع $f(\cdot)$ در نظر گرفته می‌شود. پیدا کردن دوره داغیدن و حذف این نمونه‌ها، باعث صرفه‌جویی در زمان (هزینه محاسباتی) خواهد شد. در بخش بعد از همگرایی، نمونه‌های (وابسته) $X^{(T_0)}, X^{(T_0+1)}, \ldots$ تولیدشده از $f(\cdot)$ برای استنباط در مورد کمیت‌های مورد علاقه استفاده می‌شوند. این امر، نظارت بر همگرایی را با مقایسه استنباط‌های ساخته‌شده از چند دنباله تولیدشده مستقل با نقاط شروع متفاوت، پیشنهاد می‌کند (گلمن و روبین، ۱۹۹۲).

حال بحثی که در این‌جا مطرح می‌شود این است که چگونه تشخیص دهیم یک زنجیر همگراست. زیرا انتظار داریم زنجیر تولیدشده در نهایت به توزیع مانایی که مد نظر است، همگرا شود. کاربران $MCMC$، همگرایی را با پیاده کردن روش‌های تشخیص همگرایی روی خروجی تولیدشده از نمونه‌گیری، بررسی می‌کنند. روش‌های متعددی برای بررسی همگرایی زنجیرهای $MCMC$ وجود دارند. از جمله این روش‌ها می‌توان به معیارهای همگرایی گلمن و روبین (۱۹۹۲)، جی‌وک (۱۹۹۲)، هیدل‌برگر و ولچ (۱۹۸۳) اشاره کرد. برای یک زنجیر مشخص، این معیارها توسط توابعی از بسته‌های $lattice$ (سرکار، ۲۰۰۸) و $coda$ (پلامر و همکاران، ۲۰۰۶) موجود در نرم‌افزار R قابل محاسبه هستند. ما در این‌جا دو مورد از معروف‌ترین معیارها را معرفی می‌کنیم.

[۳۲] Burn-in

۱.۷.۱ معیار همگرایی گلمن-روبین

گلمن و روبین در سال ۱۹۹۲ روشی را جهت تشخیص همگرایی توسط شبیه‌سازی دنباله‌های متعدد با نقاط شروع مختلف پیشنهاد کردند. با داشتن چگالی هدف $f(\theta)$، این معیار در هفت مرحله شکل می‌گیرد.

- مرحله اول، شامل تولید $m \geq ۲$ دنباله با طول $۲n$ و با مقادیر آغازی متفاوت می‌باشد. در این روش، n تکرار اول هر دنباله حذف شده و تمرکز روی n تکرار آخر خواهد بود.

- مرحله دوم، برای هر پارامتر مورد نظر، مقادیر زیر محاسبه می‌شوند:

$$B = \frac{n}{m-۱}\sum_{i=۱}^{m}(\bar{\theta}_{i.} - \bar{\theta}_{..})^{۲},$$
$$W = \frac{۱}{m}\sum_{i=۱}^{m} s_i^{۲},$$

که در آن $\bar{\theta}_{i.}$ میانگین هر دنباله، $\bar{\theta}_{..}$ میانگین تمام m دنباله و $s_i^{۲} = \frac{۱}{n-۱}\sum_{j=۱}^{n}(\theta_{ij} - \bar{\theta}_{i.})$ واریانس درون هر دنباله می‌باشند. توجه داشته باشید که $\frac{B}{n}$ واریانس بین میانگین‌های m دنباله و W، متوسط m واریانس درون دنباله‌ای است.

- مرحله سوم، میانگین $\hat{\mu}$ را برابر میانگین نمونه‌ای mn مقادیر شبیه‌سازی‌شده θ، یعنی $\hat{\mu} = \bar{\theta}_{..}$ قرار می‌دهیم.

- مرحله چهارم، واریانس توزیع هدف به عنوان متوسط وزنی W و B محاسبه می‌شود. یعنی $\hat{\sigma}^{۲} = \frac{n-۱}{n}W + \frac{۱}{n}B$. از این رو، اگر $n \to \infty$، آن‌گاه $E(W) = \sigma^{۲}$.

- مرحله پنجم، توزیع $N(\hat{\mu}, \hat{\sigma}^{۲})$ به عنوان برآورد توزیع هدف در نظر گرفته می‌شود که نتیجه آن تقریب توزیع t با میانگین $\hat{\mu}$، مقیاس $\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^{۲} + B/mn}$ و درجه آزادی $df = ۲\hat{V}^{۲}/\hat{var}(\hat{V})$ برای θ خواهد بود.

- مرحله ششم، مرحله نظارت بر همگرایی توسط برآورد عاملی است که عامل کاهش مقیاس بالقوه[۳۳]

[۳۳] Potential Scale Reduction Factor

$(PSRF)$ نام دارد و به صورت زیر برآورد می‌شود:

$$\hat{R} = \sqrt{\frac{\hat{V}}{W}\frac{df}{df - ۲}}$$
$$= \sqrt{(\frac{n-۱}{n} + \frac{m+۱}{mn}\frac{B}{W})\frac{df}{df-۲}}.$$

این عامل زمانی که $n \to \infty$، به سمت یک کاهش می‌یابد.

- در مرحله آخر می‌توان گفت اگر $\hat{R}$ عددی نزدیک به یک را نشان دهد، همگرایی به توزیع مانا حاصل شده است. در غیر این‌صورت، نیازمند اجرای زنجیرهای طولانی‌تر، یا به عبارتی تولید نمونه‌های بیشتر خواهیم بود.

۲.۷.۱ معیار همگرایی جی‌وک

جی‌وک در سال ۱۹۹۲ یک آماره آزمون تشخیص همگرایی که برابری میانگین‌های بخش اول و آخر زنجیرهای مارکوف را اندازه‌گیری می‌کند، ارایه داد. این دو بخش، به‌طور معمول، ۱۰ درصد اول و ۵۰ درصد آخر زنجیر در نظر گرفته می‌شوند. آزمون تشخیص همگرایی برای تعیین این‌که چه تعداد از تکرارهای آغازی باید نادیده گرفته شوند، مورد استفاده قرار می‌گیرد. به عبارتی این آزمون به عنوان یک برآوردگر برای دوره داغیدن شناخته می‌شود.

اگر n تعداد تکرارها و $n_۰$ برابر با اولین تکراری باشد که بخواهیم با انجام این آزمون بررسی کنیم که زنجیر همگرا می‌شود یا نه، باید آماره آزمون جی‌وک که یک Z-امتیاز[۳۴] استاندارد است، محاسبه شود. فرض کنید $\theta(X)$ کمیت مورد دلخواه باشد و $\theta^t = \theta(X^{(t+n_۰)})$. اگر n_A تکرار اول، تعداد تکرارهای بخش A و n_B تکرار آخر، تعداد تکرارهای بخش B، که در آن n^* مقداری است که $n - n^* + ۱ = n_B$، باشند آن‌گاه بخش‌های A و B و میانگین‌های آن‌ها به صورت زیر تعریف می‌شوند:

$$A = \{t : ۱ \le t \le n_A\},$$

$$B = \{t : n^* \le t \le n\},$$

$$۱ < n_A < n^* < n, \quad \frac{n_A + n_B}{n} < ۱,$$

$$\bar{\theta}_A = \frac{۱}{n_A}\sum_{t \in A}\theta^t,$$

[۳۴] Z-score

$$\bar{\theta}_B = \frac{۱}{n - n^* + ۱} \sum_{t \in B} \theta^t.$$

آزمون جی‌وک بیان‌گر این مطلب است که اگر زنجیر در $n_۰$ همگرا شود، آن‌گاه نمونه‌های تولیدشده از توزیع مانای زنجیر گرفته شده‌اند و در نتیجه میانگین‌های محاسبه‌شده از اولین و آخرین بخش زنجیر، معادل هستند. بنابراین آماره جی‌وک که به‌طور مجانبی دارای توزیع نرمال استاندارد است، عبارتست از

$$Z_n = \frac{\bar{\theta}_A - \bar{\theta}_B}{\sqrt{\frac{\hat{S}_\theta^A}{n_A} + \frac{\hat{S}_\theta^B}{n - n^* + ۱}}} \to N(۰, ۱) \quad n \to \infty,$$

که در آن $\frac{\hat{S}_\theta^A}{n_A}$ و $\frac{\hat{S}_\theta^B}{n-n^*+۱}$ واریانس‌های مجانبی $\bar{\theta}_A$ و $\bar{\theta}_B$ هستند. حال با توجه به مطالب ذکرشده می‌توان فرضیه صفر، فرضیه برابری میانگین‌ها، را روی فضای حالت θ آزمون کرد. در این آزمون، اگر مقدار $|Z_n|$ بزرگ باشد، آن‌گاه فرضیه صفر رد می‌شود و این یعنی زنجیر هنوز همگرا نشده است.

۸.۱ تحلیل حساسیت

همان‌طور که می‌دانیم در مواردی که یک مدل شامل پارامترهای نامعلوم و متعدد است، استفاده از روش‌های بیزی معمول است. با این وجود یکی از بزرگترین مسایل در استنباط بیزی، انتخاب و استفاده از توزیع پیشین پارامترهاست. توزیع پیشین، دیدگاه و آگاهی اولیه در مورد پارامتر است و در مواردی که هیچ‌گونه اطلاعی در مورد ماهیت پارامترها در دسترس نیست، انتخاب توزیع پیشین مشکل خواهد بود.

در تحلیل‌های بیزی، نتایج استنباط پسین پارامترها برای مسایل تصمیم‌گیری مورد استفاده قرار می‌گیرد. بنابراین برای بررسی حساسیت این تصمیم‌گیری‌ها، تحلیل حساسیت بیزی مفید می‌باشد. بررسی حساسیت استنباط بیزی نسبت به توزیع پیشین پارامترها نقش مهمی را ایفا می‌کند که توزیع پسین را با استفاده از اجرای دوباره مدل با تغییر نوع توزیع‌های پیشین یا تغییر مقادیر پارامترهای این توزیع‌ها، مورد مطالعه قرار می‌دهد. بنابراین ایده اصلی تحلیل حساسیت بیزی، که استواری بیزی[۳۵] نیز نامیده می‌شود، پس از مشخص کردن یک رده از پیشین‌هایی که سازگار با اطلاعات موجود هستند، مطالعه و ارزیابی تاثیر تغییرات توزیع‌های پیشین پارامترها در استنباط نهایی است. تحلیل حساسیت نشان خواهد داد که استنباط‌های بیزی نسبت به انتخاب توزیع‌های پیشین مختلف استوار هستند یا

[۳۵] Bayesian Robustness

خیر. به عبارتی یک استنباط بیزی استوار است، اگر نتایج استنباط به انتخاب توزیع‌های پیشین برای پارامترهای مدل، چندان وابسته نباشند.

۹.۱ سایر تعاریف

تعریف ۱.۹.۱ (ماتریس معین مثبت[۳۶]). ماتریس متقارن A با بعد $n \times n$ را معین مثبت گوییم، هرگاه به ازای هر بردار ستونی nتایی $X \neq ۰$ داشته باشیم، $X^T AX > ۰$.

تعریف ۲.۹.۱ (معیار اطلاع انحراف). معیار اطلاع انحراف[۳۷] (DIC) در مسایل مربوط به انتخاب مدل بیزی که در آن توزیع‌های پسین توسط شبیه‌سازی $MCMC$ به دست آمده‌اند، مناسب می‌باشد. این معیار به عنوان تعمیمی از AIC برای انتخاب مدل در چارچوب بیزی در نظر گرفته می‌شود.

اگر انحراف را به صورت $D(\theta) = -۲ln[L(y|\theta)]$ و تعداد پارامترهای موثر مدل را به صورت $pD = \bar{D} - D(\bar{\theta})$ تعریف کنیم، که در آن $\bar{D} = E[D(\theta)]$، $D(\bar{\theta}) = -۲ln[L(y|\bar{\theta})]$ و $\bar{\theta} = E(\theta|y)$ باشند، آن‌گاه DIC برابر است با

$$\begin{aligned} DIC &= pD + \bar{D} \\ &= D(\bar{\theta}) + ۲pD. \end{aligned}$$

در بین چندین مدل بیزی نامزد، مدلی که کمترین DIC را داشته باشد، انتخاب می‌شود. برای اطلاع در مورد ویژگی‌های این ملاک انتخاب مدل و چگونگی تعریف آن به اسپیگل‌هالتر و همکاران (۲۰۰۲) مراجعه کنید.

[۳۶] Positive Definite Matrix

[۳۷] Deviance Information Criterion

فصل ۲

مدل رگرسیون آمیخته بتا

در این فصل به دنبال معرفی یک مدل رگرسیونی مناسب برای شرایطی که متغیر پاسخ وابسته با مقیاس پیوسته و محدود به فاصله $(۰,۱)$ می‌باشد، هستیم. متغیر پاسخ در این مدل از یک توزیع بتا پیروی می‌کند که پارامترهای آن بر حسب میانگین و پارامتر دقت تعریف می‌شوند. پارامتر دقت به دو صورت ثابت (مدل اول) و وابسته به متغیرهای تبیینی (مدل دوم) در نظر گرفته می‌شود. با توجه به وجود وابستگی در متغیرهای پاسخ، از یک مدل رگرسیونی آمیخته استفاده می‌کنیم. از آن‌جایی که این مدل در رده مدل‌های آمیخته خطی تعمیم‌یافته قرار می‌گیرد، استفاده از یک رهیافت بیزی برای استنباط، اولین پیشنهاد می‌باشد. بنابراین، در این فصل، ابتدا به معرفی مدل رگرسیون آمیخته بتا می‌پردازیم و سپس برازش آن را با استفاده از روش‌های نمونه‌گیری $MCMC$ نشان می‌دهیم.

۱.۲ مدل رگرسیون بتا

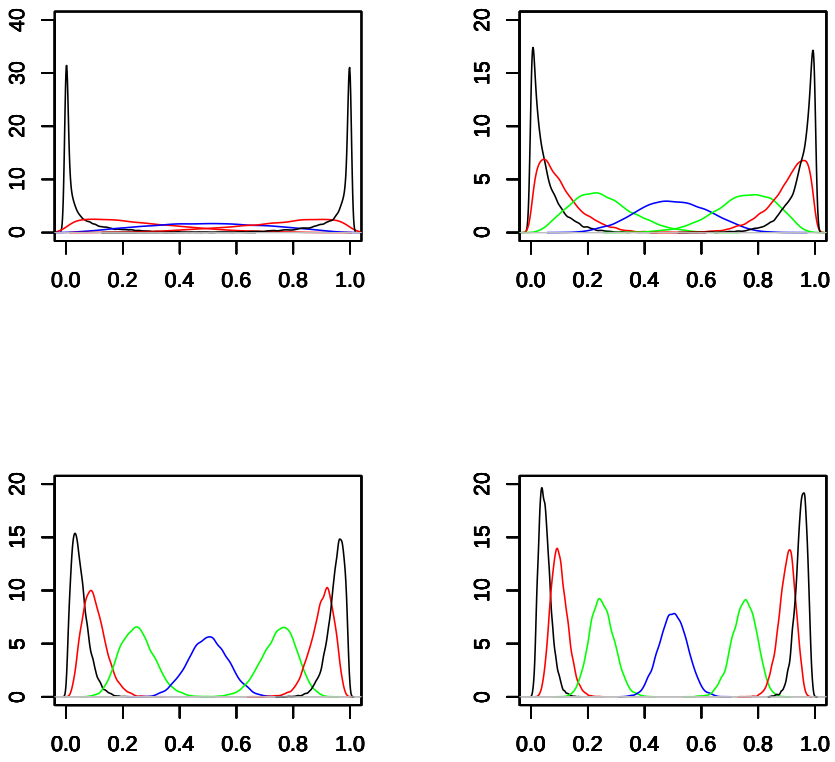
اگر تابع چگالی احتمال متغیر تصادفی Y به صورت

$$f(y;\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}y^{\alpha-۱}(۱-y)^{\beta-۱},\quad ۰<y<۱,\quad \alpha,\beta>۰,$$

باشد، می‌گوییم Y دارای توزیع بتا است و می‌نویسیم $Y\sim Beta(\alpha,\beta)$. متغیر تصادفی بتا پیوسته و محدود به فاصله $(۰,۱)$ است. با این حال هر متغیر محدود به فاصله (a,b) را نیز می‌توان با استفاده از تبدیل $y^*=\frac{y-a}{y-b}$، به فاصله $(۰,۱)$ محدود کرد و توزیع بتا را برای آن به‌کار برد.

توزیع بتا یک توزیع انعطاف‌پذیر است، زیرا تابع چگالی احتمال آن به ازای تغییر مقادیر پارامترهای توزیع می‌تواند شکل‌های متفاوتی داشته باشد (شکل ۱.۲). این انعطاف‌پذیری در طیف گسترده‌ای از

موقعیت‌های کاربردی مورد استفاده قرار می‌گیرد. بنابراین می‌توان گفت توزیع بتا یک انتخاب مناسب برای مدل‌سازی داده‌های پیوسته و محدود به فاصله $(۰, ۱)$ است. کاربردهای مختلفی از توزیع بتا توسط باری (۱۹۹۹) و جانسون و همکاران (۱۹۹۵) معرفی شده‌اند. اولین بار فراری و سریباری-نتو (۲۰۰۴) مدلی را پیشنهاد کردند که در آن، متغیر پاسخ با توزیع بتا در چارچوب مدل‌های خطی تعمیم‌یافته با استفاده از میانگین پاسخ، μ، و یک پارامتر مثبت به نام پارامتر دقت، ϕ، مدل‌سازی شده است.



شکل ۱.۲: نمودارهای توابع چگالی بتا به ازای مقادیر مختلفی از α و β

در چارچوب مدل‌های خطی تعمیم‌یافته و بر مبنای توزیع بتا، پارامترهای جدید به صورت $\mu = \frac{\alpha}{\alpha+\beta}$ و $\phi = \alpha + \beta$ تعریف می‌شوند. در نتیجه خواهیم داشت $\alpha = \mu\phi$ و $\beta = (۱ - \mu)\phi$. بنابراین می‌توان نوشت

$$E(Y) = \mu,$$
$$Var(Y) = \frac{\mu(۱ - \mu)}{۱ + \phi}.$$

با توجه به رابطه واریانس، برای μ ثابت، هرگاه ϕ افزایش یابد، واریانس متغیر پاسخ نیز کاهش می‌یابد. از این رو ϕ را می‌توان به عنوان یک پارامتر ارزیابی دقت مدل تعبیر کرد. حال تابع چگالی متغیر پاسخ y را می‌توان به صورت زیر بازنویسی کرد:

$$f(y|\mu,\phi)=\frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)}y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1},\quad 0<y<1, \qquad (1.2)$$

که در آن $\phi>0$، $0<\mu<1$ و با نماد $Y\sim Beta(\mu\phi,(1-\mu)\phi)$ نشان داده می‌شود.

اگر $Y_1,\ldots,Y_n$، n متغیر تصادفی مستقل با میانگین $\mu_1,\ldots,\mu_n$ باشند، می‌نویسیم $Y_i\sim Beta(\mu_i\phi,(1-\mu_i)\phi)$. در نتیجه مدل رگرسیون بتا مستلزم یک تبدیل برای میانگین μ_i است که می‌توان آن را به صورت

$$g(\mu_i)=\sum_{j=1}^{p}x_{ij}\beta_j=\mathbf{x}_i^T\boldsymbol{\beta}=\eta_i,$$

نوشت که $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_p)^T$ بردار ضرایب رگرسیونی نامعلوم، $\mathbf{x}_i=(x_{i1},\ldots,x_{ip})^T$ بردار متغیرهای کمکی معلوم برای iامین عضو، η_i پیشگوی خطی و $g(.)$ یک تابع پیوند یکنوا و مشتق‌پذیر می‌باشند.

یک GLM نیازمند تعریف تابع پیوند برای میانگین پاسخ می‌باشد. یک تابع پیوند مناسب برای میانگین پاسخ که قادر باشد پیشگوی خطی را به بازه $(0,1)$ بنگارد، تابع پیوند لجیت خواهد بود. به عبارتی

$$logit(\mu_i)=\ln\frac{\mu_i}{1-\mu_i}=\mathbf{x}_i^T\boldsymbol{\beta}.$$

به‌طور مشابه، یک تابع پیوند مناسب برای پارامتر دقت، که یک پارامتر اکیدا مثبت است، تابع پیوند لگاریتمی، $\ln(\phi_i)$، می‌باشد. این پارامتر می‌تواند ثابت در نظر گرفته شود (فراری و سریباری-نتو، ۲۰۰۴) یا در یک ترکیب رگرسیونی مشابه میانگین، مانند $\ln(\phi_i)=w_i^T\boldsymbol{\delta}$ مدل‌سازی شود که در آن w_i بردار متغیرهای کمکی و $\boldsymbol{\delta}$ بردار ضرایب رگرسیونی نامعلوم هستند (اسمیتسون و ورکوی‌لن، ۲۰۰۶).

پذیره استقلال پاسخ‌ها یکی از پذیره‌های اساسی مدل معرفی‌شده بالا می‌باشد. اما همانطور که در فصل پیش مطرح کردیم، موقعیت‌های مختلفی وجود دارند که در آن‌ها پاسخ‌ها وابسته هستند. در چنین مواردی که مورد دلخواه این پایان‌نامه نیز می‌باشد، به سراغ رده $GLMMs$ خواهیم رفت.

۲.۲ تحلیل بیزی مدل‌های رگرسیون آمیخته بتا

به منظور تعمیم مدل‌های خطی تعمیم‌یافته بیزی (دی و همکاران، ۲۰۰۰) و رگرسیون بتا (برنسکام و همکاران، ۲۰۰۷)، دو مدل رگرسیون آمیخته بتا (فیگوئرا و همکاران، ۲۰۱۳) را معرفی می‌کنیم: مدل اول، مدلی است که در آن پارامتر دقت به ازای همه مشاهدات ثابت در نظر گرفته می‌شود؛ مدل دوم، علاوه بر میانگین یک پیشگوی خطی با اثرات آمیخته به پارامتر دقت نیز الصاق می‌شود.

۱.۲.۲ مدل اول

فرض کنید $\mathbf{y}_1, \ldots, \mathbf{y}_m$ با شرط وجود اثرات تصادفی، $\mathbf{b}_i$، بردارهای تصادفی پیوسته و مستقل باشند که هر $\mathbf{y}_i = (y_{i1}, \ldots, y_{in_i})^T$ بردار پاسخ مشاهده‌شده برای نمونه واحد i و هر مولفه آن، y_{ij}، محدود به فاصله $(0, 1)$ است. حال یک مدل رگرسیونی با ساختار زیر را در نظر بگیرید:

$$g(E(\mathbf{y}_i|\mathbf{b}_i)) = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i, \quad i = 1, \ldots, m. \tag{۲.۲}$$

در این مدل، $g(\cdot)$ یک تابع پیوند برای میانگین شرطی بردار پاسخ با پیشگوی آمیخته خطی $\boldsymbol{\eta_i}$ به صورت $\eta_i = X_i\boldsymbol{\beta} + Z_i\mathbf{b}_i$ است، X_i ماتریس طرح $n_i \times p$ متناظر با بردار ضرایب رگرسیونی $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ (اثرات ثابت) و Z_i ماتریس طرح $n_i \times q$ متناظر با بردار $\mathbf{b}_i = (b_{i1}, \ldots, b_{iq})^T$ (اثرات تصادفی) هستند.

با انتخاب یک تابع پیوند لجیت برای میانگین شرطی، jامین مولفه رابطه (۲.۲) به صورت

$$logit(\mu_{ij}) = \ln(\frac{\mu_{ij}}{1-\mu_{ij}}) = \eta_{ij} = x_{ij}^T\boldsymbol{\beta} + z_{ij}^T\mathbf{b}_i, \tag{۳.۲}$$

خواهد بود به‌طوری‌که $z_{ij} = (z_{ij1}, \ldots, z_{ijq})^T$، $x_{ij} = (x_{ij1}, \ldots, x_{ijp})^T$ و میانگین شرطی پاسخ، $\mu_{ij} = E(y_{ij}|\mathbf{b}_i)$، معادل است با

$$\mu_{ij} = \frac{\exp(\eta_{ij})}{1+\exp(\eta_{ij})} = \frac{\exp(x_{ij}^T\boldsymbol{\beta} + z_{ij}^T\mathbf{b}_i)}{1+\exp(x_{ij}^T\boldsymbol{\beta} + z_{ij}^T\mathbf{b}_i)}. \tag{۴.۲}$$

بنابراین برای هر $i = 1, \ldots, m$ و $j = 1, \ldots, n_i$ داریم

$$y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \phi \overset{ind.}{\sim} Beta(\mu_{ij}\phi, (1-\mu_{ij})\phi).$$

یعنی y_{ij}ها مشروط به $\mathbf{b}_i$، $\boldsymbol{\beta}$و ϕ مستقل و دارای تابع چگالی (۱.۲) هستند، البته با جایگزینی μ_{ij} به جای μ که از رابطه (۴.۲) به‌دست می‌آید.

پذیره معمول برای توزیع اثرات تصادفی، توزیع نرمال می‌باشد. یعنی $\mathbf{b}_i|\Sigma_b \sim N_q(۰, \Sigma_b)$ که در آن Σ_b یک ماتریس معین مثبت است. در بسیاری از مسایل کاربردی، به دلایلی از جمله وجود نقاط پرت، پذیره نرمال بودن توزیع اثرات تصادفی انتخاب مناسبی نخواهد بود. بنابراین بهتر است یک توزیع چندمتغیره با دم‌های سنگین‌تر از نرمال برای این اثرات در نظر گرفته شود. توزیع t چندمتغیره با درجه آزادی مثبت، $\nu_b > ۰$، بردار مکانی $\mu_b = ۰ \in R^q$ و ماتریس کوواریانس معین مثبت Σ_b، به صورت $\mathbf{b}_i|\nu_b, \Sigma_b \sim t_q(\nu_b, ۰, \Sigma_b)$ یک نامزد مناسب برای این موارد می‌باشد. البته توجه داشته باشید که به ازای مقادیر بزرگ ν_b، توزیع t چندمتغیره به سمت توزیع نرمال چندمتغیره میل خواهد کرد (کاتز و ناداروجا، ۲۰۰۴). در این پایان‌نامه از پذیره نرمال برای توزیع اثرات تصادفی استفاده می‌کنیم.

تحلیل مدل رگرسیون آمیخته بتا از دیدگاه بیزی، مستلزم تعیین توزیع پیشین برای پارامترهای مدل، $\boldsymbol{\beta}$، Σ_b و ϕ می‌باشد. برای این‌که در یک الگوریتم $MCMC$ همگرایی به توزیع مانا حاصل شود، نیازمند داشتن یک توزیع پسین سره هستیم که نتیجه مطمئن آن انتخاب پیشین‌های سره می‌باشد. به‌طور معمول برای ضرایب رگرسیونی، توزیع پیشین نرمال چندمتغیره در نظر گرفته می‌شود، یعنی $\boldsymbol{\beta} \sim N_p(\mu_\beta, \Sigma_\beta)$. مشابه اثرات تصادفی، یک جایگزین مناسب، توزیع t چندمتغیره می‌باشد. یعنی $\boldsymbol{\beta} \sim t_p(\nu_\beta, \mu_\beta, \Sigma_\beta)$، البته با تعیین یک مقدار مناسب برای درجه آزادی (ν_β).

دومین پارامتر، ماتریس کوواریانس اثرات تصادفی است که به یک توزیع پیشین مناسب نیاز دارد. خانواده ویشارت، خانواده‌ای از توزیع‌های احتمال روی ماتریس‌های متقارن و معین مثبت است. بنابراین یک توزیع پیشین برای این ماتریس، توزیع ویشارت معکوس خواهد بود (فونگ و همکاران، ۲۰۱۰). یعنی $\Sigma_b \sim IW_q(\psi, c)$ با تابع چگالی

$$\pi(\Sigma_b) = \frac{|\psi|^{c/۲}}{۲^{cq/۲}.\Gamma_q(c/۲)} |\Sigma_b|^{-\frac{c+q+۱}{۲}} \exp\left(\frac{-۱}{۲} tr(\psi \Sigma_b^{-۱})\right),$$

که در آن $c \geq q$ درجه آزادی توزیع، ψ ماتریس معین مثبت با بعد $q \times q$ و $\Gamma_q(\cdot)$ تابع گاما چندمتغیره (جیمز، ۱۹۶۴) می‌باشند.

چندین انتخاب برای توزیع پیشین پارامتر دقت وجود دارد. یک انتخاب معمول برای این پارامتر، با توجه به این‌که یک پارامتر مثبت است، توزیع گامای معکوس به صورت $\phi \sim IG(\varepsilon, \varepsilon)$، با یک مقدار مثبت و کوچک برای ε، خواهد بود. تابع چگالی این توزیع به شکل زیر است:

$$\pi(\phi) = \frac{\varepsilon^\varepsilon}{\Gamma(\varepsilon)} \phi^{-(\varepsilon+۱)} e^{\frac{-\varepsilon}{\phi}}.$$

این توزیع پیشین دارای میزان آگاهی‌بخشی کمی می‌باشد (فیگوئرا و همکاران، ۲۰۱۳). گلمن (۲۰۰۶) نشان داد که توزیع پیشین ϕ به صورت $\phi = u^۲, u \sim U(۰, a)$ با تابع چگالی $\pi(\phi) = \frac{۱}{۲a\phi^{۱/۲}}$، برای مقادیر بزرگ a، آگاهی‌بخشی کمتری نسبت به توزیع گامای معکوس دارد. فیگوئرا و همکاران (۲۰۱۳)

توزیع پیشین انعطاف‌پذیرتری را برای این پارامتر پیشنهاد کردند. این توزیع و تابع چگالی آن برای مقادیر مثبت a، عبارتند از

$$\phi = (aB)^۲, B \sim Beta(۱+\varepsilon, ۱+\varepsilon),$$
$$\pi(\phi) = \frac{\Gamma(۲+۲\varepsilon)}{۲\left(\Gamma(۱+\varepsilon)\right)^۲ a^{۱+۲\varepsilon}} \phi^{\frac{\varepsilon-۱}{۲}} \left(a-\phi^{۱/۲}\right)^{\varepsilon}.$$

توجه داشته باشید که برای مقادیر کوچک ε، توزیع پیشنهادی فیگوئرا و همکاران (۲۰۱۳) با توزیع پیشنهادی گلمن (۲۰۰۶) معادل خواهد شد.

۲.۲.۲ مدل دوم

این مدل، شامل پارامترهای دقت مختلف ϕ_{ij} برای هر y_{ij} و مدل رگرسیون آمیخته (۴.۲) برای پارامتر مکانی μ_{ij} می‌باشد. مشابه پارامتر میانگین، برای مدل‌بندی پارامتر دقت نیز می‌توان از یک پیشگوی خطی آمیخته استفاده کرد. یک تابع پیوند مناسب برای این پارامتر، تابع پیوند لگاریتمی است. بنابراین

$$\ln(\phi_{ij}) = \tau_{ij} = w_{ij}^T \boldsymbol{\delta} + h_{ij}^T \mathbf{d}_i, \qquad (۵.۲)$$

که در آن $w_{ij}^T = (w_{ij۱}, \ldots, w_{ijp^*})$ بردار متناظر با اثرات ثابت $\boldsymbol{\delta}$ (با بعد p^*) و $h_{ij}^T = (h_{ij۱}, \ldots, w_{ijq^*})$ بردار متناظر با اثرات تصادفی $\mathbf{d}_i$ (با بعد q^*) می‌باشند. توجه داشته باشید که نیازی نیست ماتریس‌های طرح $W_i = (w_{i۱}, \ldots, w_{in_i})^T$ و $H_i = (h_{i۱}, \ldots, h_{in_i})^T$ همان ماتریس‌های X_i و Z_i باشند. بنابراین متغیرهای پاسخ در این مدل، با شرط وجود اثرات تصادفی، $\mathbf{b}_i$ها و $\mathbf{d}_i$ها، مستقل و به صورت

$$y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \mathbf{d}_i, \boldsymbol{\delta} \overset{ind.}{\sim} Beta(\mu_{ij}\phi_{ij}, (۱-\mu_{ij})\phi_{ij}), \quad i = ۱, \ldots, m, \quad j = ۱, \ldots, n_i,$$

نمایش داده می‌شوند.

مشابه مدل اول، فرض کنید $\mathbf{d}_i|\Sigma_d \sim N_{q^*}(۰, \Sigma_d)$ که Σ_d یک ماتریس معین مثبت است. همچنین، توزیع‌های پیشین برای $\boldsymbol{\delta}$ و Σ_d مشابه توزیع‌های به‌کار رفته برای $\boldsymbol{\beta}$ و Σ_b خواهند بود. یعنی

$$\boldsymbol{\delta} \sim N_{p^*}(\mu_\delta, \Sigma_\delta), \quad \Sigma_d \sim IW_{q^*}(\psi^*, c^*).$$

۳.۲ برازش مدل با استفاده از نمونه‌گیری $MCMC$

تحت مدل اول، فرض کنید $\mathbf{y}^T = (\mathbf{y}_۱^T, \ldots, \mathbf{y}_m^T)$ و $\boldsymbol{\eta}^T = (\boldsymbol{\eta}_۱^T, \ldots, \boldsymbol{\eta}_m^T)$ که $\mathbf{y}_i^T = (y_{i۱}, \ldots, y_{in_i})$ و $\boldsymbol{\eta}_i^T = (\eta_{i۱}, \ldots, \eta_{in_i})$ وابسته هستند. بنابر مدل رگرسیونی (۳.۲)، $\boldsymbol{\eta}_i$ها به شرط مفروض بودن $\boldsymbol{\beta}$

و Σ_b، مستقل و دارای توزیعی با ویژگی‌های

$$f(\boldsymbol{\eta}_i|\boldsymbol{\beta},\Sigma_b) \propto f(\mathbf{b}_i|\boldsymbol{\beta},\Sigma_b),$$

$$\boldsymbol{\eta}_i|\boldsymbol{\beta},\Sigma_b \sim N_{n_i}(X_i\beta, Z_i\Sigma_b Z_i^T),$$

$$f(\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_m|\boldsymbol{\beta},\Sigma_b) = \prod_{i=1}^{m} f(\boldsymbol{\eta}_i|\boldsymbol{\beta},\Sigma_b),$$

خواهند بود. $\mathbf{y}_i$ها نیز به شرط مفروض بودن $\mathbf{b}_i$ها، $\boldsymbol{\beta}$ و ϕ مستقل می‌باشند. پس تابع چگالی آن‌ها به صورت

$$f(\mathbf{y}_1,\ldots,\mathbf{y}_m|\boldsymbol{\eta}_i,\phi) = \prod_{i=1}^{m} f(\mathbf{y}_i|\boldsymbol{\eta}_i,\phi) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} f(y_{ij}|\eta_{ij},\phi),$$

تعریف می‌شود. بنابراین با توجه به روابط قبل، تابع درستنمایی حاشیه‌ای مدل برابر است با

$$\begin{aligned}
L(\phi,\boldsymbol{\beta},\Sigma_b|\mathbf{y}) &= f(\mathbf{y}_1,\ldots,\mathbf{y}_m|\phi,\boldsymbol{\beta},\Sigma_b)\\
&= \int_{R^m} f(\mathbf{y}_1,\ldots,\mathbf{y}_m|\boldsymbol{\eta}_i,\phi,\boldsymbol{\beta},\Sigma_b)\mathrm{d}\boldsymbol{\eta}_i\\
&= \int_{R^m} f(\mathbf{y}_1,\ldots,\mathbf{y}_m|\boldsymbol{\eta}_i,\phi) f(\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_m|\boldsymbol{\beta},\Sigma_b)\mathrm{d}\boldsymbol{\eta}_i\\
&= \int_{R^m} \prod_{i=1}^{m}\prod_{j=1}^{n_i} f(y_{ij}|\eta_{ij},\phi)\prod_{i=1}^{m} f(\boldsymbol{\eta}_i|\boldsymbol{\beta},\Sigma_b)\mathrm{d}\boldsymbol{\eta}_i\\
&= \int_{R^m} \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{\Gamma(\phi)}{\Gamma(\mu_{ij}\phi)\Gamma((1-\mu_{ij})\phi)} y_{ij}^{\mu_{ij}\phi-1}(1-y_{ij})^{(1-\mu_{ij})\phi-1}\\
&\quad \prod_{i=1}^{m} \frac{1}{(2\pi)^{n_i/2}|Z_i\Sigma_b Z_i^T|^{1/2}} e^{\frac{-(\eta_i-X_i\beta)^T(Z_i\Sigma_b Z_i^T)^{-1}(\eta_i-X_i\beta)}{2}}\mathrm{d}\boldsymbol{\eta}_i.
\end{aligned}$$

به راحتی می‌توان دریافت که در این مدل، به تعداد مشاهدات، m، اثرات تصادفی موجود است. به عنوان مثال، در داده‌های مقاومت بتن، $m = 108$ است. بنابراین، با توجه به بعد بالای این انتگرال و صورت پیچیده تابع زیر آن، حل عددی این انتگرال بسیار دشوار و شامل خطای قابل توجهی است. در نتیجه رهیافت پیش رو برازش مدل از دیدگاه استنباط بیزی و روش‌های نمونه‌گیری $MCMC$ می‌باشد.

تحت این فرض که پارامترهای $\boldsymbol{\beta}$، Σ_b و ϕ مستقل هستند، توزیع پسین توام (با در نظر گرفتن توزیع

پیشین گامای معکوس برای پارامتر ϕ) به صورت

$$\begin{aligned}\pi(\boldsymbol{\beta},\Sigma_b,\phi,\boldsymbol{\eta}|\mathbf{y}) &\propto \left(\prod_{i=1}^{m}\prod_{j=1}^{n_i} f(y_{ij}|\eta_{ij},\phi)\right)\left(\prod_{i=1}^{m} f(\boldsymbol{\eta}_i|\boldsymbol{\beta},\Sigma_b)\right)\pi(\boldsymbol{\beta})\pi(\phi)\pi(\Sigma_b)\\ &\propto \exp\Bigg[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln\Big(\frac{\Gamma(\phi)}{\Gamma(\mu_{ij}\phi)\Gamma((1-\mu_{ij})\phi)}y_{ij}^{\mu_{ij}\phi-1}(1-y_{ij})^{(1-\mu_{ij})\phi-1}\Big)\\ &\quad+\sum_{i=1}^{m}\ln\Big(\frac{1}{(2\pi)^{n_i/2}|Z_i\Sigma_b Z_i^T|^{1/2}}e^{\frac{-(\eta_i-X_i\beta)^T(Z_i\Sigma_b Z_i^T)^{-1}(\eta_i-X_i\beta)}{2}}\Big)\\ &\quad+\ln\Big(\frac{1}{(2\pi)^{p/2}|\Sigma_\beta|^{1/2}}e^{\frac{-(\beta-\mu_\beta)^T\Sigma_\beta^{-1}(\beta-\mu_\beta)}{2}}\Big)\\ &\quad+\ln\Big(\frac{|\psi|^{c/2}}{2^{cq/2}\Gamma_q(c/2)}|\Sigma_b|^{-\frac{c+q+1}{2}}e^{\frac{-1}{2}tr(\psi\Sigma_b^{-1})}\Big)\\ &\quad+\ln\Big(\frac{\varepsilon^\varepsilon}{\Gamma(\varepsilon)}\phi^{-(\varepsilon+1)}e^{\frac{-\varepsilon}{\phi}}\Big)\Bigg],\end{aligned}$$

خواهد بود. نمونه‌گیر گیبز با شرط مفروض بودن توزیع‌های شرطی کامل، یک نمونه مونت کارلویی از این توزیع پسین توام تولید می‌کند. توزیع‌های شرطی کامل عبارتند از

$$\begin{aligned}\pi(\boldsymbol{\beta}|\Sigma_b,\phi,\boldsymbol{\eta},\mathbf{y}) &\propto \exp\Bigg[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\Big(-\ln\Gamma(g^{-1}(\eta_{ij}(\beta))\phi)\Gamma((1-g^{-1}(\eta_{ij}(\beta)))\phi)\\ &\quad+g^{-1}(\eta_{ij}(\beta))\phi\ln\frac{y_{ij}}{1-y_{ij}}\Big)+\sum_{i=1}^{m}\ln\Big(e^{\frac{-(\eta_i-X_i\beta)^T(Z_i\Sigma_b Z_i^T)^{-1}(\eta_i-X_i\beta)}{2}}\Big)\\ &\quad+\ln\Big(\frac{1}{(2\pi)^{p/2}|\Sigma_\beta|^{1/2}}e^{\frac{-(\beta-\mu_\beta)^T\Sigma_\beta^{-1}(\beta-\mu_\beta)}{2}}\Big)\Bigg]\\ &\propto \prod_{i=1}^{m}\prod_{j=1}^{n_i}\frac{(\frac{y_{ij}}{1-y_{ij}})^{g^{-1}(\eta_{ij}(\beta))\phi}}{\Gamma(g^{-1}(\eta_{ij}(\beta))\phi)\Gamma((1-g^{-1}(\eta_{ij}(\beta)))\phi)}\\ &\quad\times\frac{e^{-\left((\beta-\mu_\beta)^T\Sigma_\beta^{-1}(\beta-\mu_\beta)+\sum_{i=1}^{m}(\eta_i-X_i\beta)^T(Z_i\Sigma_b Z_i^T)^{-1}(\eta_i-X_i\beta)\right)/2}}{(2\pi)^{p/2}|\Sigma_\beta|^{1/2}},\end{aligned}$$

$$\pi(\Sigma_b|\boldsymbol{\beta},\phi,\boldsymbol{\eta},\mathbf{y}) \propto \exp\Bigg[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\Big(-\ln\Gamma(g^{-1}(\eta_{ij}(b_i))\phi)\Gamma((1-g^{-1}(\eta_{ij}(b_i)))\phi) + g^{-1}(\eta_{ij}(b_i))\phi\ln\frac{y_{ij}}{1-y_{ij}}\Big) + \sum_{i=1}^{m}\ln\Big(e^{\frac{-(\eta_i-X_i\beta)^T(Z_i\Sigma_bZ_i^T)^{-1}(\eta_i-X_i\beta)}{2}}\Big) + \ln\Big(\frac{|\psi|^{c/2}}{2^{cq/2}\Gamma_q(c/2)}|\Sigma_b|^{-\frac{c+q+1}{2}}e^{\frac{-1}{2}tr(\psi\Sigma_b^{-1})}\Big)\Bigg]$$
$$\propto \prod_{i=1}^{m}\prod_{j=1}^{n_i}\frac{(\frac{y_{ij}}{1-y_{ij}})^{g^{-1}(\eta_{ij}(b_i))\phi}}{\Gamma(g^{-1}(\eta_{ij}(b_i))\phi)\Gamma((1-g^{-1}(\eta_{ij}(b_i)))\phi)} \times \frac{|\psi|^{c/2}}{2^{cq/2}\Gamma_q(c/2)}|\Sigma_b|^{-\frac{c+q+1}{2}}e^{\frac{-1}{2}\left(tr(\psi\Sigma_b^{-1})+\sum_{i=1}^{m}(\eta_i-X_i\beta)^T(Z_i\Sigma_bZ_i^T)^{-1}(\eta_i-X_i\beta)\right)},$$

$$\pi(\phi|\boldsymbol{\beta},\Sigma_b,\boldsymbol{\eta},\mathbf{y}) \propto \exp\Bigg[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\Big(\ln\frac{\Gamma(\phi)}{\Gamma(\mu_{ij}\phi)\Gamma((1-\mu_{ij})\phi)} + \mu_{ij}\phi\ln\frac{y_{ij}}{1-y_{ij}}\Big) + \ln\Big(\frac{\varepsilon^{\varepsilon}}{\Gamma(\varepsilon)}\phi^{-(\varepsilon+1)}e^{\frac{-\varepsilon}{\phi}}\Big)\Bigg]$$
$$\propto \prod_{i=1}^{m}\prod_{j=1}^{n_i}\frac{\Gamma(\phi)}{\Gamma(\mu_{ij}\phi)\Gamma((1-\mu_{ij})\phi)}(\frac{y_{ij}}{1-y_{ij}})^{\mu_{ij}\phi} \times \frac{\varepsilon^{\varepsilon}}{\Gamma(\varepsilon)}\phi^{-(\varepsilon+1)}e^{\frac{-\varepsilon}{\phi}},$$

و برای $i,k=1,\ldots,m, i\neq k$،

$$\pi(\boldsymbol{\eta}_i|\boldsymbol{\eta}_k,\mathbf{y}_i,\boldsymbol{\beta},\Sigma_b,\phi) \propto \exp\Bigg[\sum_{i=1}^{m}\sum_{j=1}^{n_i}\Big(-\ln\Gamma(g^{-1}(\eta_{ij}(\beta))\phi)\Gamma((1-g^{-1}(\eta_{ij}(\beta)))\phi) + g^{-1}(\eta_{ij}(\beta))\phi\ln\frac{y_{ij}}{1-y_{ij}}\Big) + \sum_{i=1}^{m}\ln\Big(e^{\frac{-(\eta_i-X_i\beta)^T(Z_i\Sigma_bZ_i^T)^{-1}(\eta_i-X_i\beta)}{2}}\Big)\Bigg]$$
$$\propto \prod_{i=1}^{m}\prod_{j=1}^{n_i}\frac{(\frac{y_{ij}}{1-y_{ij}})^{g^{-1}(\eta_{ij})\phi}}{\Gamma(g^{-1}(\eta_{ij})\phi)\Gamma((1-g^{-1}(\eta_{ij}))\phi)} \times e^{\frac{-1}{2}\sum_{i=1}^{m}(\eta_i-X_i\beta)^T(Z_i\Sigma_bZ_i^T)^{-1}(\eta_i-X_i\beta)},$$

به‌طوری که $\eta(\beta)$ و $\eta(b)$ به منظور تاکید وابستگی پیشگوی خطی به ضرایب ثابت و تصادفی β و b استفاده شده‌اند. همان‌طور که مشاهده می‌کنید هیچ یک از این توزیع‌های شرطی کامل، شکل مشخص و شناخته‌شده‌ای ندارند. محاسبه این توزیع‌ها، کاری دشوار و در مواردی ناممکن است. اما نرم‌افزار

$JAGS$[۱] این توزیع‌ها را به صورت خودکار محاسبه می‌کند و تولید نمونه توسط آن به آسانی صورت می‌پذیرد (پلامر، ۲۰۰۳). تولید نمونه از هر کدام از این توزیع‌های شرطی کامل به کمک الگوریتم‌هایی مانند متروپولیس-هستینگس صورت می‌پذیرد که به الگوریتم حاصل، الگوریتم متروپولیس-درون-گیبز[۲] می‌گویند. سپس، استنباط‌های پسین بر پایه پارامترهای $\boldsymbol{\beta}$، Σ_b، ϕ و μ_i در $JAGS$ محاسبه می‌شوند.

تحت مدل دوم، فرض کنید $\boldsymbol{\tau}^T = (\boldsymbol{\tau}_1^T, \ldots, \boldsymbol{\tau}_m^T)$، که $\boldsymbol{\tau}_i^T = (\tau_{i1}, \ldots, \tau_{in_i})$. آن‌گاه بنابر مدل رگرسیونی (۵.۲)، $\boldsymbol{\tau}_i$ها مشروط به $\boldsymbol{\delta}$ و Σ_d مستقل بوده و تابع چگالی آن‌ها به صورت

$$f(\boldsymbol{\tau}_i|\boldsymbol{\delta}, \Sigma_d) \propto f(\mathbf{d}_i|\boldsymbol{\delta}, \Sigma_d),$$

$$\boldsymbol{\tau}_i|\boldsymbol{\delta}, \Sigma_d \sim N_{n_i}(W_i\delta, H_i\Sigma_d H_i^T),$$

$$f(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_m|\boldsymbol{\delta}, \Sigma_d) = \prod_{i=1}^{m} f(\boldsymbol{\tau}_i|\boldsymbol{\delta}, \Sigma_d),$$

بیان می‌شود. بنابراین با توجه به روابط موجود، تابع درستنمایی حاشیه‌ای را می‌توان به صورت زیر تعریف کرد:

$$\begin{aligned}
L(\boldsymbol{\beta}, \Sigma_b, \boldsymbol{\delta}, \Sigma_d|\mathbf{y}) &= f(\mathbf{y}_1, \ldots, \mathbf{y}_m|\boldsymbol{\beta}, \Sigma_b, \boldsymbol{\delta}, \Sigma_d) \\
&= \int_{R^m}\int_{R^m} \prod_{i=1}^{m}\prod_{j=1}^{n_i} f(y_{ij}|\eta_{ij}, \tau_{ij}) \prod_{i=1}^{m} f(\boldsymbol{\eta}_i|\boldsymbol{\beta}, \Sigma_b) \prod_{i=1}^{m} f(\boldsymbol{\tau}_i|\boldsymbol{\delta}, \Sigma_d)\mathrm{d}\boldsymbol{\eta}_i\mathrm{d}\boldsymbol{\tau}_i \\
&= \int_{R^m}\int_{R^m} \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{\Gamma(\phi_{ij})}{\Gamma(\mu_{ij}\phi_{ij})\Gamma((1-\mu_{ij})\phi_{ij})} y_{ij}^{\mu_{ij}\phi_{ij}-1}(1-y_{ij})^{(1-\mu_{ij})\phi_{ij}-1} \\
&\quad \times \prod_{i=1}^{m} \frac{1}{(2\pi)^{n_i/2}|Z_i\Sigma_b Z_i^T|^{1/2}} e^{\frac{-(\eta_i - X_i\beta)^T(Z_i\Sigma_b Z_i^T)^{-1}(\eta_i - X_i\beta)}{2}} \\
&\quad \times \prod_{i=1}^{m} \frac{1}{(2\pi)^{n_i/2}|H_i\Sigma_d H_i^T|^{1/2}} e^{\frac{-(\tau_i - W_i\delta)^T(H_i\Sigma_d H_i^T)^{-1}(\tau_i - W_i\delta)}{2}} \mathrm{d}\boldsymbol{\eta}_i\mathrm{d}\boldsymbol{\tau}_i,
\end{aligned}$$

که حل عددی آن بسیار دشوار است.

با فرض استقلال میان $\boldsymbol{\beta}$، Σ_b، $\boldsymbol{\delta}$ و Σ_d، توزیع پسین توام (با در نظر گرفتن توزیع پیشین گامای

[۱] Just Another Gibbs Sampler

[۲] Metropolis-within-Gibbs

معکوس برای پارامتر ϕ) از رابطه

$$\begin{aligned}
\pi(\boldsymbol{\beta}, \Sigma_b, \boldsymbol{\eta}, \boldsymbol{\tau}, \boldsymbol{\delta}, \Sigma_d | \mathbf{y}) \quad &\propto \quad \left(\prod_{i=1}^{m} \prod_{j=1}^{n_i} f(y_{ij} | \eta_{ij}, \tau_{ij}) \right) \left(\prod_{i=1}^{m} f(\boldsymbol{\eta}_i | \boldsymbol{\beta}, \Sigma_b) \right) \\
&\quad \left(\prod_{i=1}^{m} f(\boldsymbol{\tau}_i | \boldsymbol{\delta}, \Sigma_d) \right) \pi(\boldsymbol{\beta}) \pi(\Sigma_b) \pi(\boldsymbol{\delta}) \pi(\Sigma_d) \\
&\propto \quad \prod_{i=1}^{m} \prod_{j=1}^{n_i} \frac{\Gamma(\phi_{ij})}{\Gamma(\mu_{ij}\phi_{ij}) \Gamma((1-\mu_{ij})\phi_{ij})} y_{ij}^{\mu_{ij}\phi_{ij}-1} (1-y_{ij})^{(1-\mu_{ij})\phi_{ij}-1} \\
&\quad \prod_{i=1}^{m} \frac{1}{(2\pi)^{n_i/2} |Z_i \Sigma_b Z_i^T|^{1/2}} e^{\frac{-(\eta_i - X_i\beta)^T (Z_i \Sigma_b Z_i^T)^{-1} (\eta_i - X_i\beta)}{2}} \\
&\quad \prod_{i=1}^{m} \frac{1}{(2\pi)^{n_i/2} |H_i \Sigma_d H_i^T|^{1/2}} e^{\frac{-(\tau_i - W_i\delta)^T (H_i \Sigma_d H_i^T)^{-1} (\tau_i - W_i\delta)}{2}} \\
&\quad \times \frac{e^{-(\beta-\mu_\beta)^T \Sigma_\beta^{-1} (\beta-\mu_\beta)/2}}{(2\pi)^{p/2} |\Sigma_\beta|^{1/2}} \times \frac{e^{-(\delta-\mu_\delta)^T \Sigma_\delta^{-1} (\delta-\mu_\delta)/2}}{(2\pi)^{p^*/2} |\Sigma_\delta|^{1/2}} \times \frac{\varepsilon^\varepsilon}{\Gamma(\varepsilon)} \phi^{-(\varepsilon+1)} e^{\frac{-\varepsilon}{\phi}} \\
&\quad \times \frac{|\psi|^{c/2} |\Sigma_b|^{-\frac{c+q+1}{2}}}{2^{cq/2} \Gamma_q(c/2)} e^{\frac{-1}{2} tr(\psi \Sigma_b^{-1})} \times \frac{|\psi^*|^{c^*/2} |\Sigma_d|^{-\frac{c^*+q^*+1}{2}}}{2^{c^*q^*/2} \Gamma_{q^*}(c^*/2)} e^{\frac{-1}{2} tr(\psi^* \Sigma_d^{-1})},
\end{aligned}$$

به‌دست می‌آید. در این مدل، توزیع‌های شرطی کامل عبارتند از

$$\pi(\Sigma_b | \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\delta}, \Sigma_d, \mathbf{y}), \pi(\boldsymbol{\beta} | \Sigma_b, \boldsymbol{\eta}, \boldsymbol{\tau}, \boldsymbol{\delta}, \Sigma_d, \mathbf{y}),$$

$$\pi(\boldsymbol{\delta} | \Sigma_b, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\tau}, \Sigma_d, \mathbf{y}), \pi(\Sigma_d | \Sigma_b, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\delta}, \mathbf{y}),$$

$$\pi(\boldsymbol{\eta}_i | \boldsymbol{\eta}_k, \boldsymbol{\tau}_k, \mathbf{y}_i, \Sigma_b, \boldsymbol{\beta}, \boldsymbol{\delta}, \Sigma_d), \pi(\boldsymbol{\tau}_i | \boldsymbol{\tau}_k, \boldsymbol{\eta}_k, \mathbf{y}_i, \Sigma_b, \boldsymbol{\beta}, \boldsymbol{\delta}, \Sigma_d),$$

$$i, k = 1, \ldots, m, i \neq k.$$

که مشابه توزیع‌های شرطی کامل در مدل اول محاسبه می‌شوند و دارای صورت‌های شناخته‌شده‌ای نیستند. مشابه مدل اول، نمونه‌های مونت کارلویی از توزیع پسین توام را می‌توان در $JAGS$ تولید کرد و استنباط‌های پسین بر پایه μ_{ij} و ϕ_{ij} نیز در $JAGS$ محاسبه می‌شوند.

فصل ۳

مطالعه شبیه‌سازی

در این بخش با استفاده از شبیه‌سازی، عملکردهای دو مدل اول و دوم معرفی‌شده در فصل دوم، مورد ارزیابی قرار می‌گیرند. برای این منظور، ابتدا مدل‌های مورد نظر شبیه‌سازی و برازش داده می‌شوند. سپس در هر یک از آن‌ها، مدل مناسب براساس DIC انتخاب خواهد شد و در نهایت همگرایی برآوردگرهای مدل منتخب مورد بررسی قرار می‌گیرد. بررسی همگرایی مدل با روش‌های تشخیص همگرایی گلمن-روبین و جی‌وک انجام می‌شود. همچنین، بررسی رفتار آمیختگی زنجیر با رسم نمودارهای اثر[۱] و خودهمبستگی[۲] صورت می‌گیرد.

۱.۳ مدل با پارامتر دقت ثابت

در این شبیه‌سازی، به منظور تشریح مدل پیشنهادی اول، مدل با پارامتر دقت ثابت، و ارزیابی عملکرد تحلیل بیزی آن، مدل رگرسیون آمیخته بتا با شرایط زیر را برای $i = ۱, \ldots, ۱۰۰$ و $j = ۱, \ldots, ۵$ در نظر گرفتیم:

$$y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \phi \sim Beta(\mu_{ij}\phi, (۱ - \mu_{ij})\phi),$$

که در آن $\boldsymbol{\beta} = (\beta_۱, \beta_۲, \beta_۳)^T$، $\mathbf{b}_i = (b_{i۱}, b_{i۲})^T$ و

$$\ln(\frac{\mu_{ij}}{۱ - \mu_{ij}}) = \eta_{ij} = x_{ij}\boldsymbol{\beta} + z_{ij}\mathbf{b}_i. \qquad (۱.۳)$$

---

[۱] Trace Plots

[۲] Autocorrelation Plots

در این مطالعه اثرات تصادفی را دارای توزیع $N_۲(\mu_b, \Sigma_b)$ با $\mu_b = (۰, ۰)$ در نظر گرفتیم.

۱.۱.۳ شبیه‌سازی مدل با متغیرهای تبیینی پیوسته

در این بخش، مقادیر متغیرهای تبیینی را از توزیع یکنواخت استاندارد تولید کردیم، یعنی فرض کردیم $x_{ij}, z_{ij} \sim U(۰, ۱)$. مقادیر واقعی برای پارامترهای مدل، ϕ، $\boldsymbol{\beta}$ و Σ_b را به صورت

$$\phi = ۴۹, \quad \boldsymbol{\beta} = (-۲, ۱, ۲)^T, \quad \Sigma_b = \begin{pmatrix} ۱/۰۹ & -۰/۳۶ \\ -۰/۳۶ & ۰/۱۳ \end{pmatrix}$$

در نظر گرفتیم. همچنین برای این پارامترها، توزیع‌های پیشین $\Sigma_b \sim IW_۲(\Psi, c)$ و $\boldsymbol{\beta} \sim N_۳(\mu_\beta, \Sigma_\beta)$ با مقادیر

$$\Sigma_\beta = \begin{pmatrix} ۵ & ۰ & ۰ \\ ۰ & ۵ & ۰ \\ ۰ & ۰ & ۵ \end{pmatrix}, \Psi = \begin{pmatrix} ۵ & ۱۵ \\ ۱۵ & ۴۵ \end{pmatrix}, a = ۰/۰۰۱, c = ۵, \mu_\beta = (۰, ۰, ۰)^T,$$

را اختصاص دادیم.

با توجه به آن‌که توزیع‌های پیشین مختلفی برای پارامتر دقت در متون آماری پیشنهاد شده‌اند، بر آن شدیم تا اثر این پیشین‌ها را به عنوان یکی از پارامترهای مهم بر استنباط پسین بررسی کنیم. این توزیع‌های پیشین مختلف در جدول ۱.۳ آمده‌اند. برازش مدل با این پیشین‌ها را با در نظر گرفتن دو زنجیر با نقاط آغازی مختلف و ۱۰۰۰۰ تکرار مونت کارلویی، برای هر یک از مدل‌های جدول ۱.۳ انجام دادیم.

در جدول ۱.۳، برآوردهای میانگین پسین و انحراف معیار پارامتر دقت در هر برازش و مقادیر DIC گزارش شده‌اند. اولین نتیجه‌ای که از این جدول می‌توان گرفت این است که نوع توزیع پیشین پارامتر دقت، تاثیری بر استنباط پسین ندارد زیرا مقادیر برآوردشده پارامتر دقت در هر برازش نزدیک به هم هستند. نتیجه دوم، انتخاب مدل با توزیع پیشین $a \cdot ۴$ به عنوان مناسب‌ترین مدل است، زیرا براساس DIC، مدلی که کمترین مقدار DIC را داشته باشد از برازش بهتری نسبت به سایر مدل‌ها برخوردار است.

با توجه به انتخاب مدل با توزیع پیشین $a \cdot ۴$ براساس DIC، برازش این مدل انجام شد و نتایج در جدول ۲.۳ گزارش شده‌اند. در این جدول، مقادیر خلاصه‌شده‌ای از توزیع پسین از جمله میانگین پسین، انحراف معیار و فاصله اعتبار ۹۵٪ گزارش شده‌اند. نتایج نشان می‌دهند که مقادیر برآوردشده هر یک از پارامترها به مقدار واقعی آن‌ها نزدیک هستند و دقت استنباط‌ها قابل قبول است. البته عدم اعتماد

جدول ۱.۳: توزیع‌های پیشین مختلف برای ϕ و مقادیر DIC هر مدل

مدل	توزیع پیشین برای ϕ	میانگین	انحراف معیار	DIC
$a \cdot ۱$	$\phi \sim IG(۰/۰۰۱, ۰/۰۰۱)$	۵۲/۳۴۳۲	۴/۰۹۴۱	−۲۶۲۹
$a \cdot ۲$	$\phi = u^۲, u \sim U(۰, ۵۰)$	۵۲/۸۵۴۷	۳/۳۳۸۱	−۲۶۷۳
$a \cdot ۳$	$\phi = (۵۰B)^۲, B \sim beta(۱/۱, ۱/۱)$	۵۲/۳۰۸۸	۳/۳۲۶۴	−۲۶۸۵
$a \cdot ۴$	$\phi = (۵۰B)^۲, B \sim beta(۱/۵, ۱/۵)$	۵۲/۸۲۴۸	۳/۳۶۹	−۲۶۸۷
$a \cdot ۵$	$\ln(\phi) \sim t(۰, ۵, ۱۰)$	۵۲/۴۶۸۶	۴/۰۵۰۸	−۲۵۹۵

به استنباط پارامترهای وابستگی اثرات تصادفی، Σ_b، را نیز می‌توان نتیجه گرفت، زیرا برآوردهای این پارامترها دارای اریبی هستند و برآوردگر پارامتر $\Sigma_{b_{۲۲}}$، با توجه به آن‌که مقدار واقعی پارامتر در فاصله اعتبار ٪۹۵ قرار نگرفته است، نوعی ناسازگاری را نمایش می‌دهد.

جدول ۲.۳: نتایج برازش مدل با پارامتر دقت ثابت برای توزیع پیشین $a \cdot ۴$

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ٪۹۵
$\beta_۱$	−۲	−۲/۰۱۲۲	۰/۰۴۱۸	(−۲/۰۹۲۴, −۱/۹۲۷۷)
$\beta_۲$	۱	۱/۰۰۰۳	۰/۰۳۹۵	(۰/۹۲۱۵, ۱/۰۷۷۸)
$\beta_۳$	۲	۲/۰۱۵۹	۰/۰۳۹۸	(۱/۹۳۶, ۲/۰۹۶۸)
ϕ	۴۹	۵۲/۸۲۴۸	۳/۳۶۹	(۴۶/۱۷۵, ۵۹/۴۳۷)
$\Sigma_{b_{۱۱}}$	۱/۰۹	۱/۲۲۲	۰/۱۷۶۲	(۰/۹۲۱۸, ۱/۶۰۱۴)
$\Sigma_{b_{۱۲}}$	−۰/۳۶	−۰/۲۴۰۶	۰/۰۸۷۸	(−۰/۴۲۶۱, −۰/۰۸۱۸)
$\Sigma_{b_{۲۲}}$	۰/۱۳	۰/۵۸۳۸	۰/۰۸۴۶	(۰/۴۴۵۶, ۰/۷۶۶۱)

حال سوالی که این‌جا مطرح می‌شود این است که آیا توزیع زنجیرهای تولیدشده به توزیع مانای خود همگرا هستند؟ بررسی این مساله توسط نسخه چندمتغیره تشخیص همگرایی گلمن و روبین که در براکز و گلمن (۱۹۹۸) آمده است، صورت می‌گیرد که در آن $PSRF$ را برای یک ترکیب خطی از متغیرها

محاسبه می‌کند. بنابر این آزمون، یک زنجیر همگراست اگر عامل کاهش مقیاس چندمتغیره[۳] یعنی $MPSRF$ نزدیک یک (کمتر از ۱٫۲) باشد. مقدار این عامل برای مدل با توزیع پیشین $a \cdot ۴$ برابر ۱٫۰۱ به‌دست آمده است که از ۱٫۲ کمتر است. پس می‌توان گفت زنجیر تولیدشده همگراست و این یعنی نمونه‌های تولیدشده از توزیع مانا، یعنی توزیع پسین توام، آمده‌اند. کیفیت این نمونه‌های تولیدشده را با محاسبه ESS نیز می‌توان سنجید. جدول ۳.۳، مقادیر ESS را برای هر یک از پارامترها نشان می‌دهد. همان‌طور که مشاهده می‌کنید، این مقادیر، برای هر یک از پارامترها نزدیک به تعداد نمونه‌های تولیدشده از توزیع پسین به‌دست آمده است (با توجه به تعداد تکرارها و تعداد زنجیرها، تعداد نمونه‌های تولیدشده از توزیع پسین برابر با ۲۰۰۰ است). بنابراین نمونه‌های تولیدشده از کیفیت قابل قبولی برخوردارند.

جدول ۳.۳: مقادیر ESS برای پارامترهای مدل اول با توزیع پیشین $a \cdot ۴$

پارامترها	$\beta_۱$	$\beta_۲$	$\beta_۳$	ϕ	$\Sigma_{b_{۱۱}}$	$\Sigma_{b_{۱۲}}$	$\Sigma_{b_{۲۲}}$
ESS	۱۷۵۰٫۹۱	۲۰۰۰	۱۸۷۸٫۶۱	۱۸۷۴٫۶۹	۱۶۴۸٫۲۲	۱۶۷۰٫۵۸	۱۶۸۸٫۱۸

بررسی همگرایی زنجیر برای هر پارامتر نیز توسط آزمون‌های همگرایی گلمن–روبین و جی‌وک انجام شدند. نتایج این آزمون‌های همگرایی در جدول ۴.۳ گزارش شده‌اند. همان‌طور که مشاهده می‌کنید، در این جدول مقادیر معیار همگرایی گلمن–روبین و جی‌وک برای هر یک از پارامترها گزارش شده است. با بررسی نتایج به وضوح می‌توان گفت زنجیر برای هر پارامتر همگرا است. زیرا $PSRF$ برای هر پارامتر مقداری نزدیک به یک و قدرمطلق آماره آزمون جی‌وک، $|Z_n|$، برای هر پارامتر مقداری کوچک را نشان می‌دهند.

روش دوم ارزیابی همگرایی زنجیرها برای هر پارامتر با رسم نمودارهای جی‌وک، میانگین تجمعی، اثر و خودهمبستگی صورت می‌پذیرد. با رسم نمودارهای جی‌وک می‌توان همگرایی را در هر زنجیر و برای هر یک از پارامترها بررسی کرد. در این نمودار، نیمه اول زنجیر به ۱۹ بخش تقسیم شده (به‌طور پیش‌فرض زنجیر به ۲۰ بخش تقسیم می‌شود) و سپس آماره آزمون جی‌وک، Z_n، بارها و بارها محاسبه می‌شود. اولین Z_n با تمام تکرارهای زنجیر، دومین Z_n بعد از کنار گذاشتن بخش اول و سومین بعد از کنار گذاشتن دو بخش اول و به همین ترتیب محاسبه می‌شوند. در نهایت آخرین Z_n تنها با استفاده از نیمه دوم زنجیر محاسبه خواهد شد.

نمودارهایی که در شکل ۱.۳ گزارش شده‌اند، نمودارهای جی‌وک تحت زنجیر اول برای هر یک از پارامترها هستند. با توجه به این نمودارها مشاهده می‌کنید که تعداد زیادی از Z_nهای محاسبه‌شده در

[۳] Multivariate Proportional Scale Reduction Factor

جدول ۴.۳: نتایج آزمون‌های همگرایی برای مدل با پارامتر دقت ثابت با توزیع پیشین $a \cdot ۴$

پارامترها	آماره آزمون		
	گلمن-روبین	جی‌وک	
	برآورد آماره آزمون	زنجیر۱	زنجیر۲
$\beta_۱$	۱/۰۰	۱/۷۱۱	۰/۴۵۸
$\beta_۲$	۱/۰۱	−۰/۱۱۴	−۰/۴۴۳
$\beta_۳$	۱/۰۰	−۱/۲۸۲	۰/۴۴۹
ϕ	۱/۰۰	−۱/۹۸۹	−۱/۸۲۸
$\Sigma_{b_{۱۱}}$	۱/۰۰	−۰/۷۹۶	−۰/۴۲۳
$\Sigma_{b_{۱۲}}$	۱/۰۰	−۰/۳۵۵	−۱/۰۸۹
$\Sigma_{b_{۲۲}}$	۱/۰۰	۰/۳۰۴	۱/۵۰۳

فاصله مشخص‌شده قرار دارند، یا به عبارتی مقادیر کوچکی را شامل شده‌اند. بنابراین می‌توان گفت نمونه‌های تولیدشده از توزیع مانای زنجیر، گرفته شده‌اند.

نمودار میانگین تجمعی، میانگین پسین تجمعی را در مقابل تعداد تکرارها رسم می‌کند. شکل ۲.۳، این نمودارها را برای هر یک از پارامترها نمایش می‌دهد. با بررسی این نمودارها می‌توان همگرایی هر پارامتر به مقدار واقعی آن را مشاهده کرد.

نمودارهای موجود در شکل ۳.۳، نمودارهای اثر را تحت هر یک از پارامترها نشان می‌دهند. با رسم این نمودار می‌توان رفتار آمیختگی زنجیرها را در فضای پارامتر مشاهده کرد. با توجه به این نمودارها و نوسانات منظم زنجیر در هر یک از آن‌ها می‌توان نتیجه گرفت که زنجیر برای هر پارامتر از آمیختگی خوبی برخوردار است. این آمیختگی به این معنی است که نواحی چگال پسین به خوبی مورد کاوش قرار گرفته‌اند.

روشی دیگر برای ارزیابی همگرایی، ارزیابی خودهمبستگی‌های بین زنجیرهای مارکوف است. از آن‌جا که یک مقدار نمونه در زمان t وابسته به مقدار قبلی آن در زمان $t-۱$ است، نمونه‌ها همواره همبسته خواهند بود. ارزیابی همبستگی در نمونه‌های موجود در زنجیرها با نمودار خودهمبستگی در مقابل تاخیر صورت می‌پذیرد که انتظار داریم با افزایش تاخیر، همبستگی نیز به سرعت کاهش یابد. هدف از رسم این نمودار، مشاهده کاهش سریع همبستگی یک زنجیر به سمت صفر است که این نشان‌دهنده این مطلب

است که نمونه‌های تولیدشده در زنجیر، دارای کیفیت قابل قبولی هستند و به توزیع مانای خود همگرا می‌شوند.

نمودارهای موجود در شکل ۴.۳، نمودارهای خودهمبستگی را به ازای هر یک از پارامترها نشان می‌دهند. در هر یک از این نمودارها می‌توان مشاهده کرد که نمونه‌های تولیدشده از کیفیت بالایی برخوردارند و وابستگی آن‌ها به سرعت به سمت صفر میل می‌کند. به عبارت دیگر، زنجیر برای هر پارامتر ناهمبسته است.

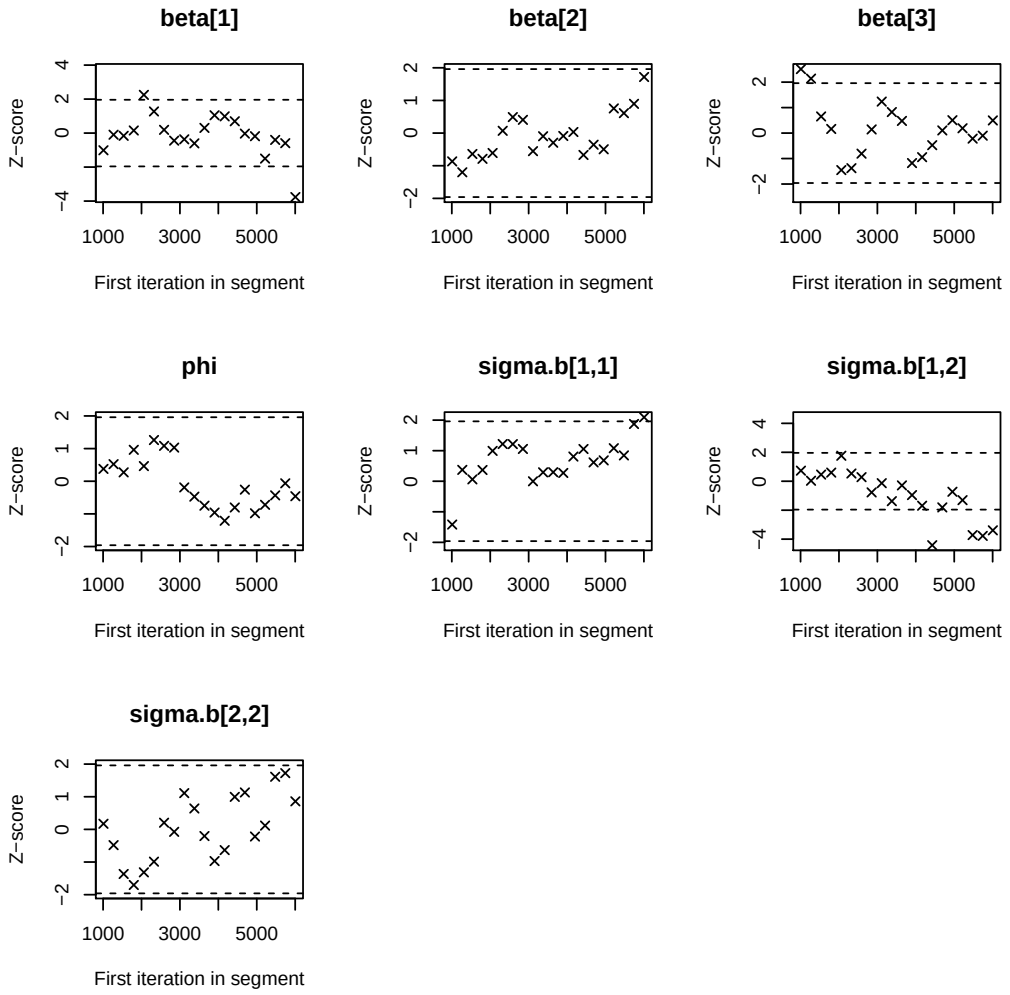
۲.۱.۳ شبیه‌سازی مدل با متغیرهای تبیینی گسسته

در این بخش از شبیه‌سازی، مدل اول با متغیرهای تبیینی گسسته برازش داده شده است. این متغیرها را از توزیع برنولی تولید کردیم؛ یعنی $x_{ij}, z_{ij} \sim Bernoulli(۰/۵)$. پارامترها نیز مانند بخش قبل دارای توزیع پیشین $\phi = (۵۰B)^۲, B \sim beta(۱/۵, ۱/۵)$، $\boldsymbol{\beta} \sim N_۳(\mu_\beta, \Sigma_\beta)$ و $\Sigma_b \sim IW_۲(\Psi, c)$ با مجموعه مقادیر مشابه هستند. برازش مدل را با در نظر گرفتن دو زنجیر با نقاط آغازی مختلف و ۱۰۰۰۰ تکرار مونت کارلویی انجام دادیم.

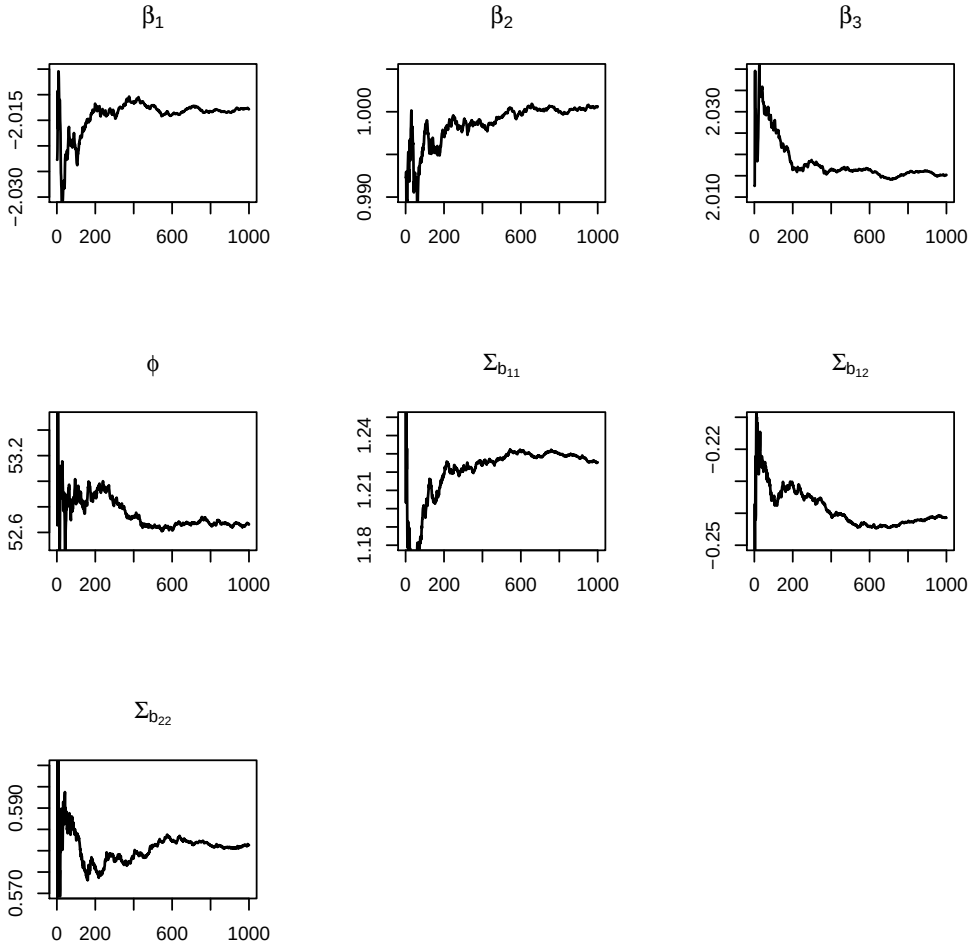
نتایج برازش این مدل در جدول ۵.۳ آمده‌اند که شامل مقدار واقعی هر پارامتر، برآوردهای میانگین پسین، انحراف معیار و فاصله اعتبار ٪۹۵ می‌باشند. مشاهده می‌کنیم که مقدار برآوردهای بیزی هر پارامتر به مقدار واقعی آن نزدیک است. اما پارامترهای وابستگی اثرات تصادفی، Σ_b، ناسازگار و دارای اریبی بالایی هستند که بیانگوی عملکرد نه چندان خوب روش برازش در برآورد پارامترهای وابستگی مدل است. البته در بسیاری از موارد در موقعیت‌های کاربردی آن‌چه که اهمیت دارد تحلیل و استنباط مناسب برای پارامترهای رگرسیونی، یعنی بردار β، است و اثرات تصادفی برای لحاظ کردن وابستگی بین پاسخ‌ها در مدل در نظر گرفته می‌شوند.

آزمون همگرایی گلمن-روبین برای این مدل انجام شد و مقدار $MPSRF$ برابر با $۱/۰۱ < ۱/۲$ به‌دست آمد که اثبات‌کننده همگرایی زنجیرها به توزیع مانا می‌باشد.

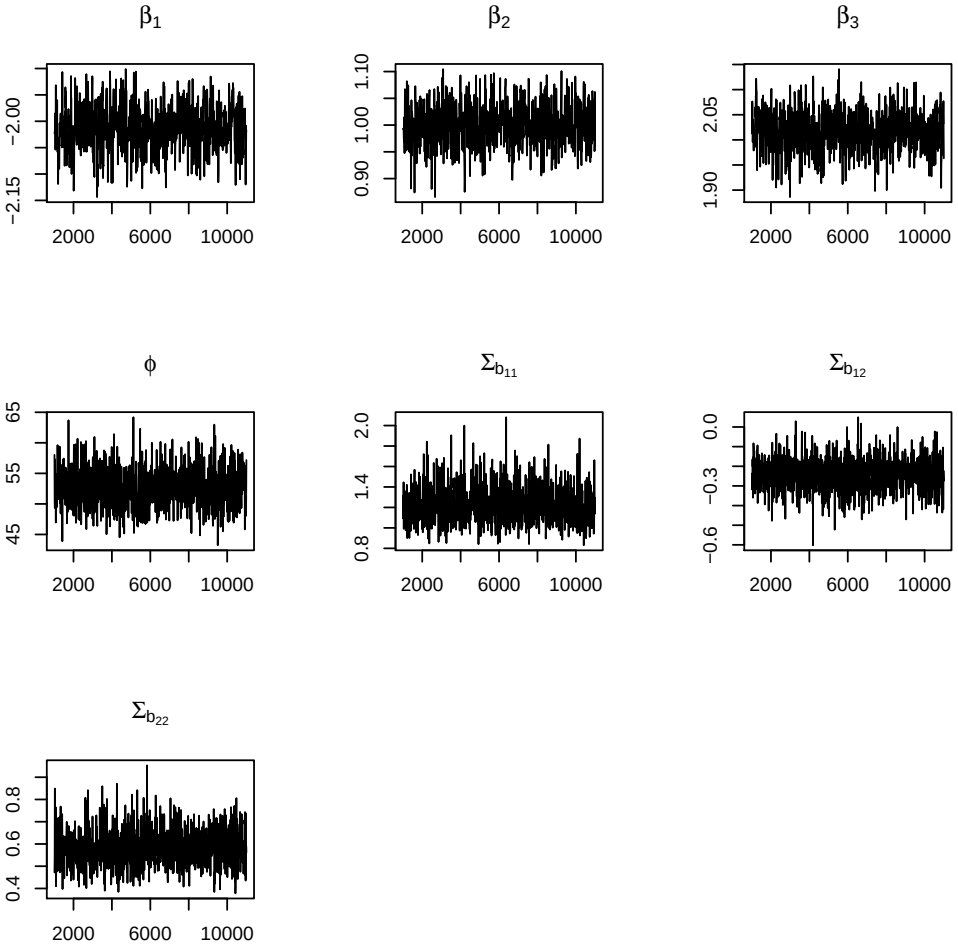
بقیه نتایج تحلیل بیزی برای این متغیرهای تبیینی مشابه متغیرهای تبیینی پیوسته به‌دست آمدند که از گزارش آن‌ها صرفنظر کردیم.



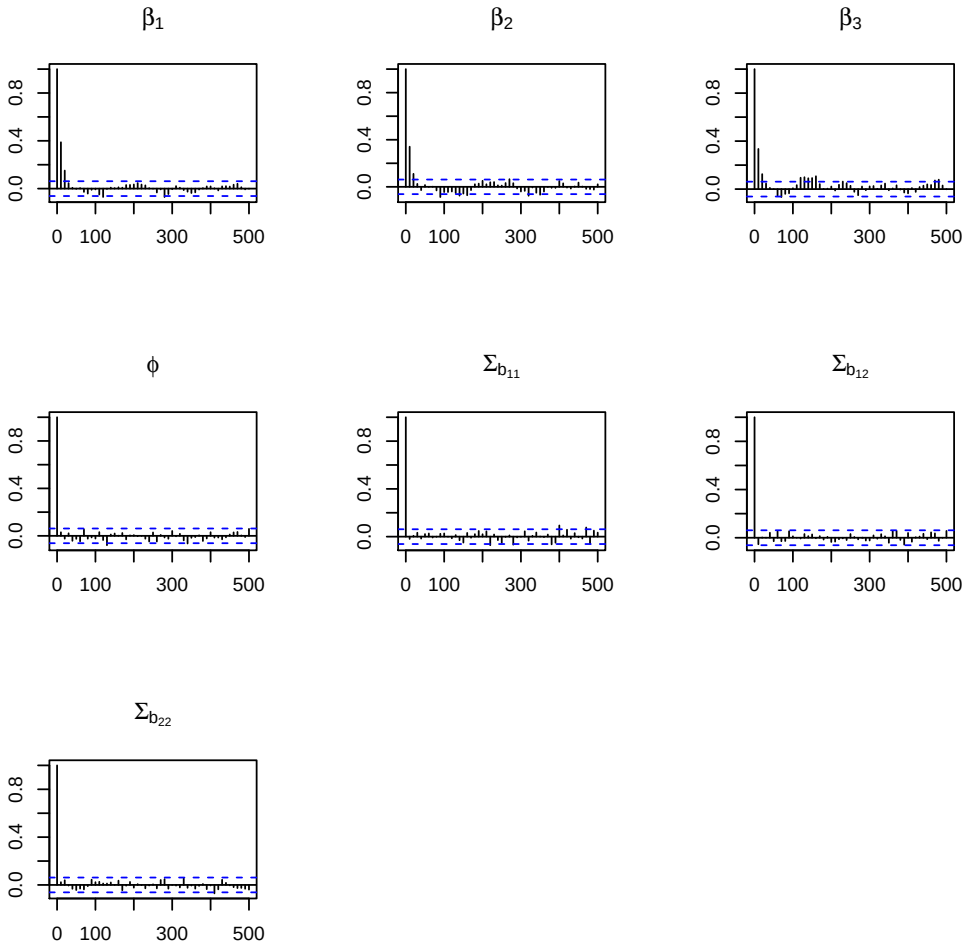
شکل ۱.۳: نمودارهای جی‌وک پارامترهای مدل اول با توزیع پیشین $a \cdot ۴$ تحت زنجیر اول



شکل ۲.۳: نمودارهای میانگین تجمعی پارامترهای مدل اول با توزیع پیشین $a \cdot ۴$



شکل ۳.۳: نمودارهای اثر پارامترهای مدل اول با توزیع پیشین $a \cdot ۴$



شکل ۴.۳: نمودارهای خودهمبستگی پارامترهای مدل اول با توزیع پیشین $a \cdot ۴$ تحت زنجیر اول

جدول ۵.۳: نتایج برازش مدل اول با متغیرهای تبیینی گسسته

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
β_1	-2	$-1/978$	$0/0009$	$(-2/055, -1/899)$
β_2	1	$1/003$	$0/0008$	$(0/935, 1/072)$
β_3	2	$2/064$	$0/0009$	$(1/981, 2/143)$
ϕ	49	$46/999$	$0/081$	$(40/238, 54/61)$
$\Sigma_{b_{11}}$	$1/09$	$1/069$	$0/0039$	$(0/769, 1/441)$
$\Sigma_{b_{12}}$	$-0/36$	$-0/09$	$0/0022$	$(-0/29, 0/096)$
$\Sigma_{b_{22}}$	$0/13$	$0/675$	$0/0022$	$(0/505, 0/887)$

۲.۳ مدل با پارامتر دقت متغیر

در این بخش به برازش مدل دوم که یک مدل رگرسیون آمیخته بتا با پارامتر دقت متغیر است، یعنی

$$y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \mathbf{d}_i, \boldsymbol{\delta} \sim Beta(\mu_{ij}\phi_{ij}, (1-\mu_{ij})\phi_{ij}),$$

خواهیم پرداخت. برای شبیه‌سازی این مدل، از همان مجموعه داده‌های شبیه‌سازی‌شده در مدل اول استفاده کردیم. یعنی،

$$x_{ij}, z_{ij} \sim U(0,1), \boldsymbol{\beta} = (-2, 1, 2)^T, \Sigma_b = \begin{pmatrix} 1/09 & -0/36 \\ -0/36 & 0/13 \end{pmatrix}$$

و $\boldsymbol{\delta} = (3/55, 0/41, -2/01)^T$. همچنین توزیع پیشین پارامترهای Σ_b و $\boldsymbol{\beta}$ را همان توزیع‌های پیشین پیشنهادشده در مدل اول در نظر گرفتیم. میانگین پاسخ نیز در این مدل همان ساختار (۱.۳) را دارد و برای پارامتر دقت که دارای تابع پیوند لگاریتمی است، یک ساختار رگرسیونی به صورت

$$\ln(\phi_{ij}) = (\delta_1 + d_{i1}) + (\delta_2 + d_{i2})x_{ij2} + \delta_3 x_{ij3}, \qquad (۲.۳)$$

در نظر گرفتیم که در آن ضرایب رگرسیونی پارامتر دقت، $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3)^T$، دارای توزیع $N_3(0, \Sigma_\beta)$ و اثرات تصادفی پارامتر دقت، $\mathbf{d}_i = (d_{i1}, d_{i2})^T$، دارای توزیعی یکسان با اثرات تصادفی پارامتر مکانی، $\mathbf{b}_i$، یعنی $N_2(0, \Sigma_b)$ می‌باشند.

برازش این مدل را با در نظر گرفتن دو زنجیر با نقاط آغازی مختلف و ۱۰۰۰۰ تکرار مونت کارلویی انجام دادیم. نتیجه این برازش در جدول ۷.۳ گزارش شده است. در این جدول اطلاعاتی نظیر مقدار واقعی پارامتر، برآوردهای میانگین پسین، انحراف معیار و فاصله اعتبار ٪۹۵ گزارش شده‌اند. با توجه به جدول مشاهده می‌کنیم که مقادیر میانگین پسین، یا همان برآوردهای بیزی هر پارامتر نزدیک به مقادیر واقعی آن‌ها به‌دست آمده‌اند. همچنین ناسازگاری و اریبی در پارامتر $\Sigma_{b_{۲۲}}$ نیز مشاهده می‌شود. در این برازش، $MPSRF = ۱/۰۱ < ۱/۲$ می‌باشد که بیان‌کننده همگرایی زنجیر به توزیع مانا خواهد بود. همچنین کیفیت این نمونه‌های تولیدشده را می‌توان در جدول ۶.۳ مشاهده کرد. با توجه به مقادیر ESS به‌دست آمده برای هر یک از پارامترها، می‌توان گفت نمونه‌های تولیدشده از کیفیت قابل قبولی برخوردارند.

جدول ۶.۳: مقادیر ESS برای پارامترهای مدل دوم

پارامتر	$\beta_۱$	$\beta_۲$	$\beta_۳$	$\delta_۱$	$\delta_۲$	$\delta_۳$	$\Sigma_{b_{۱۱}}$	$\Sigma_{b_{۱۲}}$	$\Sigma_{b_{۲۲}}$
ESS	۲۰۰۰	۲۰۰۰	۱۸۶۳/۹	۱۷۵۶/۵	۱۹۳۹	۱۶۵۷	۱۸۴۲/۵	۱۹۰۲/۹	۱۸۲۵/۹

آزمون‌های همگرایی گلمن-روبین و جی‌وک برای مدل دوم انجام شده و نتایج آن‌ها در جدول ۸.۳ گزارش شده‌اند. در آزمون گلمن-روبین، $PSRF$ به ازای هر پارامتر عددی نزدیک به یک به‌دست آمده و در آزمون جی‌وک، قدرمطلق Z_nها به ازای هر پارامتر در هر زنجیر مقداری کوچک را نشان می‌دهد. بنابراین با توجه به این نتایج می‌توان گفت زنجیر برای هر پارامتر همگرا است. همچنین همگرایی و آمیختگی زنجیرها برای هر پارامتر توسط نمودارهای میانگین تجمعی، جی‌وک، اثر و خودهمبستگی در شکل‌های ۵.۳ تا ۸.۳ نمایش داده شده‌اند که همگی این شکل‌ها تاییدکننده همگرایی، آمیختگی خوب زنجیرهای تولیدشده و کیفیت قابل قبول نمونه‌های تولیدشده در زنجیرها هستند.

۳.۳ تحلیل حساسیت برای هر دو مدل

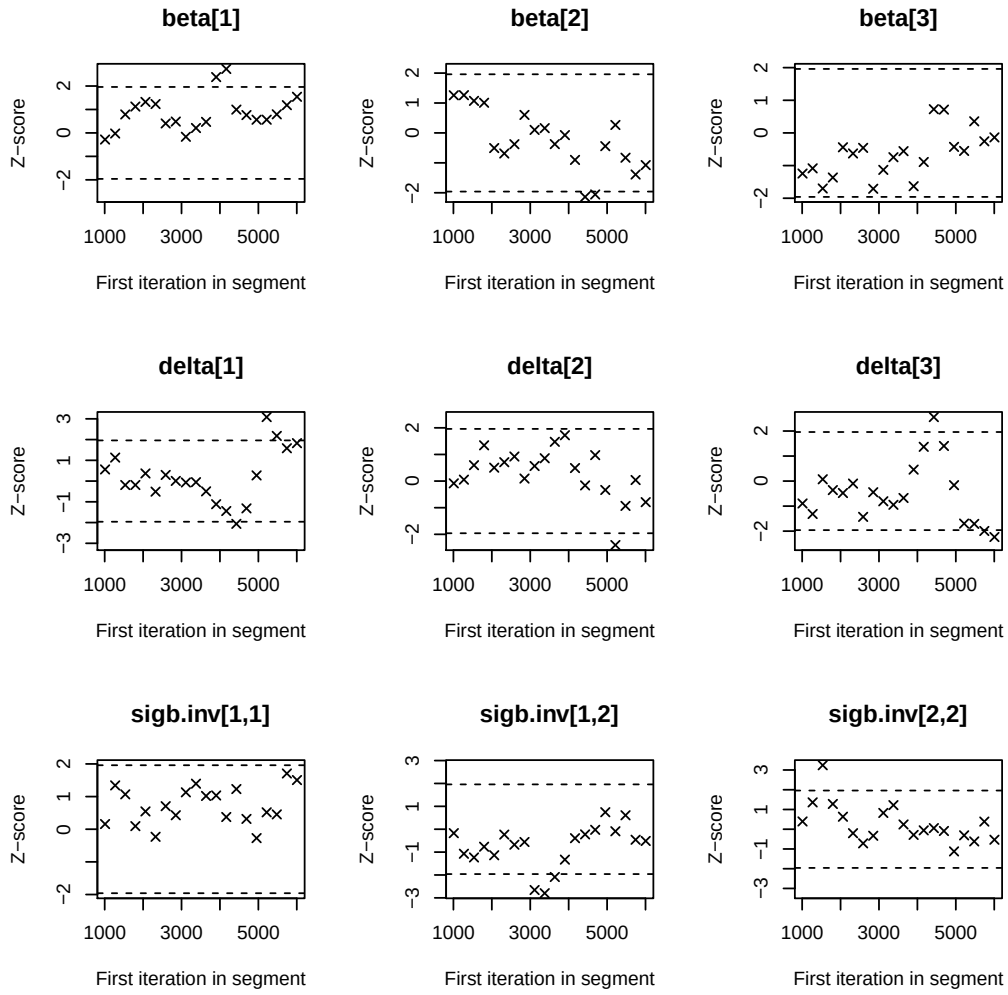
به منظور بررسی تاثیر توزیع پیشین پارامترها بر نتایج دو مدل پیشنهادی، توزیع‌های پیشین پارامترها را در هر دو مدل تغییر دادیم که نتایج حاصل در این بخش گزارش شده‌اند.

جدول ۷.۳: نتایج برازش مدل دوم

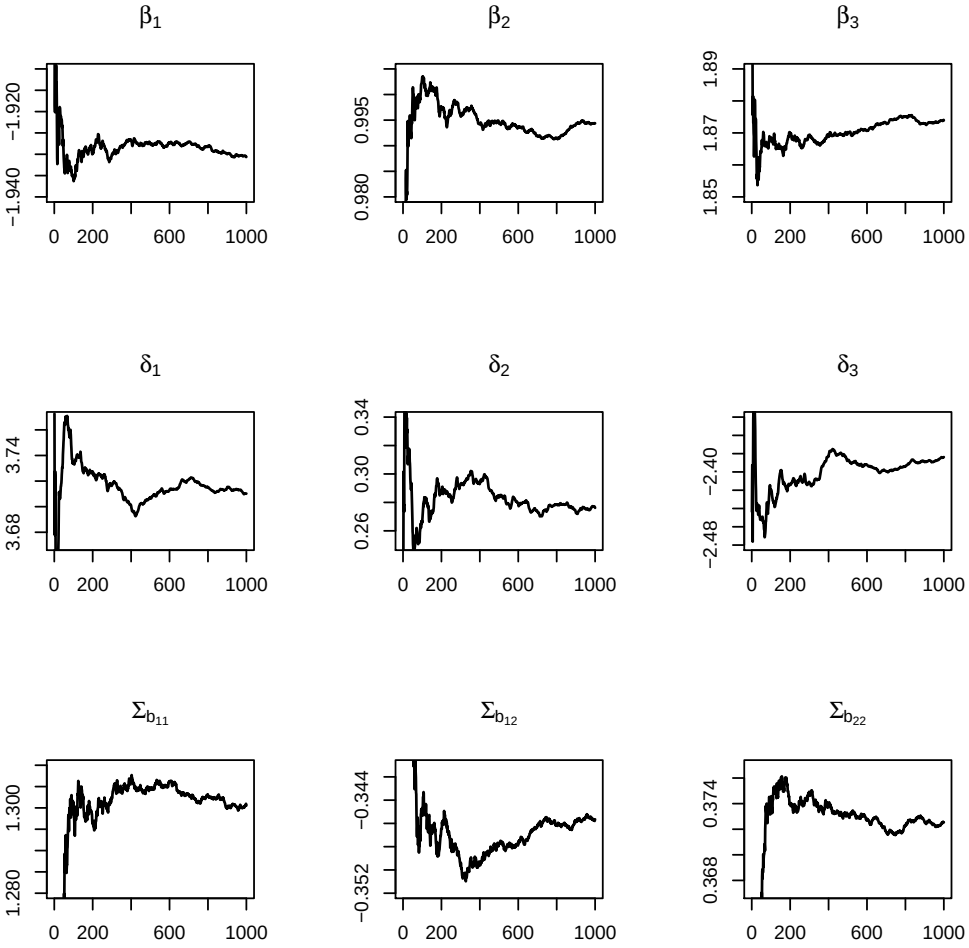
پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
β_1	-2	$-1/9356$	۰/۰۰۱۳	$(-2/0531, -1/8186)$
β_2	۱	۱/۰۰۰۲	۰/۰۰۱۳	$(0/8849, 1/1083)$
β_3	۲	۱/۸۶۸۶	۰/۰۰۱۵	$(1/7345, 2/0957)$
δ_1	۳/۵۵	۳/۷۱۰۰	۰/۰۰۳۵	$(3/394, 4/02)$
δ_2	۰/۴۱	۰/۳۰۳۵	۰/۰۰۴۷	$(-0/099, 0/738)$
δ_3	$-2/01$	$-2/407$	۰/۰۰۴۴	$(-2/783, -2/014)$
$\Sigma_{b_{11}}$	۱/۰۹	۱/۲۹۹۹	۰/۰۰۲۹	$(1/0674, 1/5673)$
$\Sigma_{b_{12}}$	$-0/36$	$-0/3425$	۰/۰۰۱۲	$(-0/4565, -0/246)$
$\Sigma_{b_{22}}$	۰/۱۳	۰/۳۷۱۸	۰/۰۰۰۸	$(0/3048, 0/4498)$

جدول ۸.۳: نتایج آزمون‌های همگرایی برای مدل دوم

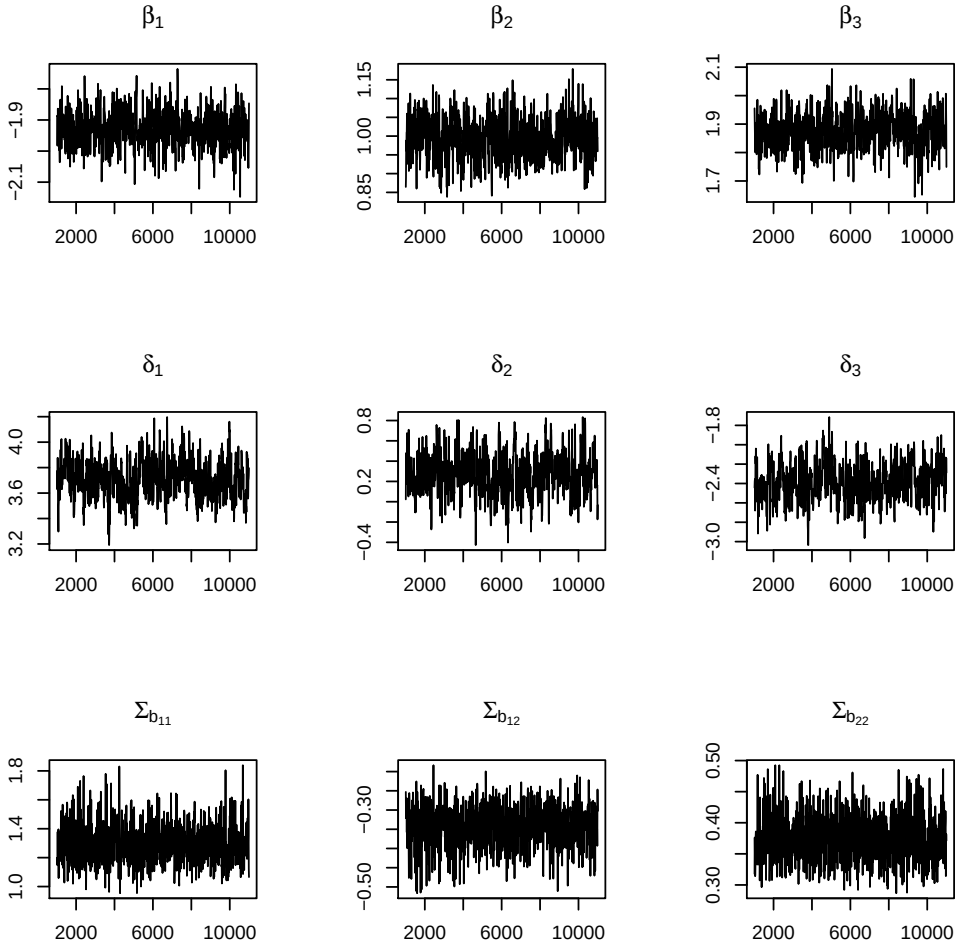
پارامترها	آزمون گلمن-روبین	آزمون جی‌وک	
	برآورد آماره آزمون	زنجیر۱	زنجیر۲
β_1	۱/۰۰۲	$-0/243$	$-0/244$
β_2	۱/۰۰۰	$-0/114$	۰/۹۶۷
β_3	۰/۹۹۹	۰/۳۳۴	۱/۸۵۱
δ_1	۱/۰۰۹	$-0/065$	$-0/404$
δ_2	۱/۰۰۰	۱/۵۹۱	$-2/139$
δ_3	۱/۰۰۰	۰/۲۳۶	۰/۵۰۵
$\Sigma_{b_{11}}$	۱/۰۰۴	۱/۰۴۶	$-0/785$
$\Sigma_{b_{12}}$	۱/۰۰۱	$-0/201$	۰/۸۹۹
$\Sigma_{b_{22}}$	۱/۰۰۱	۰/۵۰۶	$-0/798$



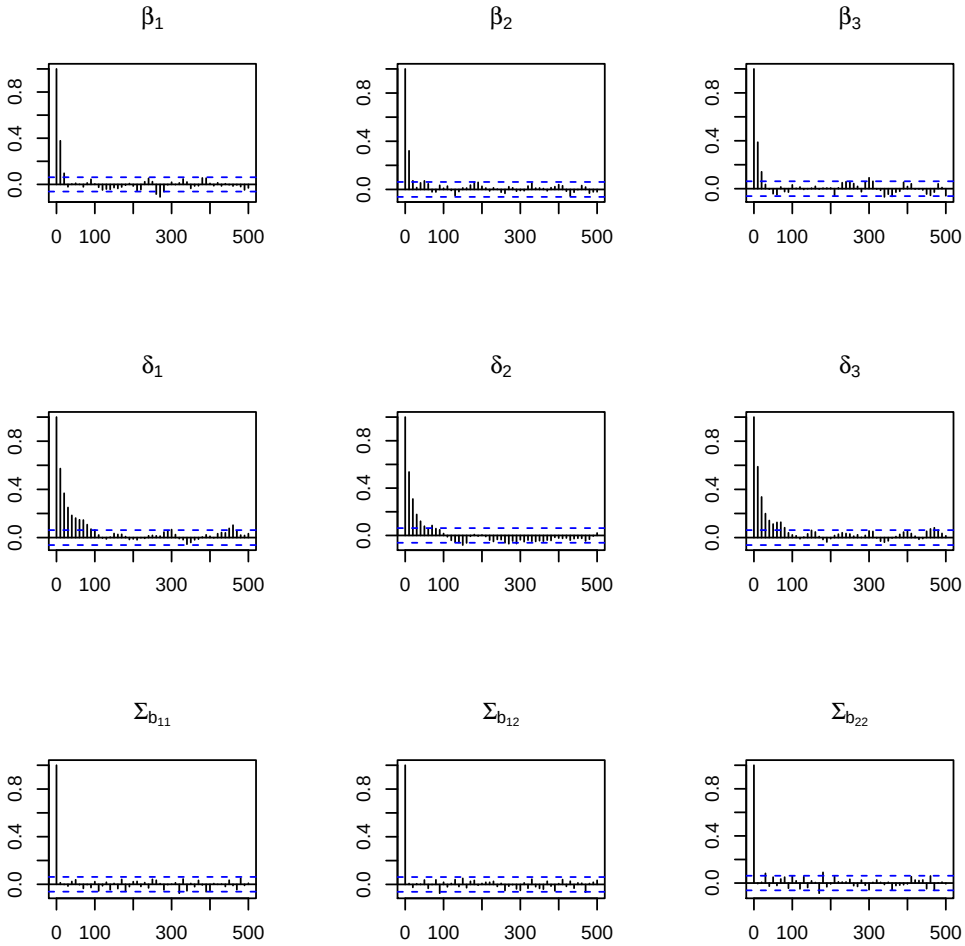
شکل ۵.۳: نمودارهای جی‌وک پارامترهای مدل دوم تحت زنجیر اول



شکل ۶.۳: نمودارهای میانگین تجمعی پارامترهای مدل دوم



شکل ۷.۳: نمودارهای اثر پارامترهای مدل دوم



شکل ۸.۳: نمودارهای خودهمبستگی پارامترهای مدل دوم تحت زنجیر اول

۱.۳.۳ مدل اول

با در نظر گرفتن همان شرایط ذکرشده در برازش مدل اول و نیز رابطه (۱.۳) برای میانگین، فرض کنید پارامترهای ϕ، $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^T$، و Σ_b دارای توزیع‌های

$$\boldsymbol{\beta} \sim N_3(\mu, \Sigma), \quad \Sigma_b \sim IW_2(\Psi, c), \quad \phi = u^2, u \sim U(0, 50), \tag{۳.۳}$$

با مقادیر

$$\Sigma = \begin{pmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{pmatrix}, \Psi = \begin{pmatrix} 10 & 25 \\ 25 & 50 \end{pmatrix} \mu = (100, 100, 100)^T, c = 10,$$

باشند.

مدل اول با شرایط ذکرشده برازش داده شد و نتایج آن در جدول ۹.۳ نمایش داده شده‌اند. همچنین، این مدل را با توزیع پیشین دیگری برای $\boldsymbol{\beta}$ به صورت

$$\boldsymbol{\beta} \sim t_3(\mu, \Sigma, \nu), \mu = (0, 0, 0)^T, \nu = 20, \Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \tag{۴.۳}$$

و با همان توزیع‌های پیشین ذکرشده در (۳.۳) برای ϕ و Σ_b، برازش دادیم که نتایج آن در جدول ۱۰.۳ آمده‌اند. از یک طرف، این نتایج نشان می‌دهند که مقادیر برآوردشده هر یک از پارامترها، در هر دو جدول، نزدیک به مقدار واقعی آن‌ها هستند. از طرف دیگر، با مقایسه نتایج این دو جدول با جدول ۲.۳، مشاهده می‌کنیم که برآوردهای حاصل از به‌کارگیری هر سه مجموعه از توزیع‌های پیشین بسیار نزدیک به هم هستند. بنابراین، می‌توان نتیجه گرفت که انتخاب توزیع پیشین پارامترها بر نتایج استنباط تاثیر چندانی ندارد و استنباط‌های بیزی نسبت به انتخاب توزیع‌های پیشین استوار هستند. به عبارت دیگر، مدل نسبت به نوع پیشینی که برای پارامترها انتخاب شده، حساس نیست.

۲.۳.۳ مدل دوم

مشابه زیربخش قبل، به برازش مدل دوم می‌پردازیم. روابط (۱.۳) و (۲.۳) برای میانگین و پارامتر دقت نیز برقرار هستند. همچنین برای پارامترهای $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^T$، $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3)^T$ و Σ_b توزیع‌های پیشین

$$\boldsymbol{\beta} \sim N_3(\mu, \Sigma), \quad \boldsymbol{\delta} \sim N_3(\mu, \Sigma), \quad \Sigma_b \sim IW_2(\Psi, c), \tag{۵.۳}$$

جدول ۹.۳: نتایج برازش مدل اول با مجموعه پیشین‌های (۳.۳)

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
$\beta_۱$	-۲	$-۱/۹۹۴$	۰/۰۴۷۲	$(-۲/۰۸۶۷, -۱/۹۰۱۳)$
$\beta_۲$	۱	۱/۰۱۳	۰/۰۴۰۶	(۰/۹۳۶, ۱/۰۹۲۹)
$\beta_۳$	۲	۱/۹۹۳	۰/۰۴۴	(۱/۹۰۲۸, ۲/۰۷۹۷)
ϕ	۴۹	۴۹/۷۴۳	۳/۰۳۴۷	(۴۳/۷۷۱, ۵۵/۷۳۶)
$\Sigma_{b_{۱۱}}$	۱/۰۹	۱/۱۳۳	۰/۱۵۷	(۰/۸۵۷۳, ۱/۴۷۶۹)
$\Sigma_{b_{۱۲}}$	$-۰/۳۶$	$-۰/۱۱۴$	۰/۰۸۰۸	$(-۰/۲۷۹۹, ۰/۰۳۹۵)$
$\Sigma_{b_{۲۲}}$	۰/۱۳	۰/۵۹۵	۰/۰۸۱۵	(۰/۴۵۲۳, ۰/۷۷۶۵)

جدول ۱۰.۳: نتایج برازش مدل اول با مجموعه پیشین‌های (۴.۳)

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
$\beta_۱$	-۲	$-۱/۹۹۷$	۰/۰۴۶۷	$(-۲/۰۸۹۷, -۱/۹۰۳۲)$
$\beta_۲$	۱	۱/۰۱۵	۰/۰۳۹۴	(۰/۹۳۹, ۱/۰۸۸۹)
$\beta_۳$	۲	۱/۹۹۴	۰/۰۴۵۴	(۱/۹۰۵۵, ۲/۰۸۲۷)
ϕ	۴۹	۴۹/۷۵۷	۳/۱۶۵۲	(۴۲/۷۶۹۳, ۵۶/۰۴۴۹)
$\Sigma_{b_{۱۱}}$	۱/۰۹	۱/۱۲۹	۰/۱۵۲۷	(۰/۸۶۹۶, ۱/۴۵۸۱)
$\Sigma_{b_{۱۲}}$	$-۰/۳۶$	$-۰/۱۱۵$	۰/۰۷۹۵	$(-۰/۳۶۶۴, ۰/۰۳۲۶)$
$\Sigma_{b_{۲۲}}$	۰/۱۳	۰/۵۹۶	۰/۰۸۳۲	(۰/۲۵۳۲, ۰/۷۷۴۲)

با مقادیر

$$\Sigma = \begin{pmatrix} ۱۵ & ۰ & ۰ \\ ۰ & ۱۵ & ۰ \\ ۰ & ۰ & ۱۵ \end{pmatrix}, \Psi = \begin{pmatrix} ۱۰ & ۲۵ \\ ۲۵ & ۵۰ \end{pmatrix} \mu = (۱۰۰, ۱۰۰, ۱۰۰)^T, c = ۱۰,$$

و توزیع‌های پیشین

$$\boldsymbol{\beta} \sim t_۳(\mu, \Sigma, \nu), \quad \boldsymbol{\delta} \sim t_۳(\mu, \Sigma, \nu), \tag{۶.۳}$$

$$\mu = (۰, ۰, ۰)^T, \nu = ۲۰, \Sigma = \begin{pmatrix} ۵ & ۰ & ۰ \\ ۰ & ۵ & ۰ \\ ۰ & ۰ & ۵ \end{pmatrix},$$

با همان پیشین ذکرشده در (۵.۳) برای Σ_b را در نظر گرفتیم.

نتایج برازش مدل دوم با دو مجموعه پیشین‌های (۵.۳) و (۶.۳) در جدول‌های ۱۱.۳ و ۱۲.۳ گزارش شده‌اند. مشاهده می‌کنیم که مقادیر برآوردشده هر یک از پارامترها، در هر دو جدول، بسیار نزدیک به مقدار واقعی آن‌ها هستند. همچنین، مقایسه نتایج این دو جدول با جدول ۷.۳ بیانگر این مطلب است که مدل دوم نیز نسبت به انتخاب پیشین پارامترها حساس نیست و استنباط‌های بیزی آن نسبت به انتخاب توزیع‌های پیشین استوار هستند.

البته برای بررسی بیشتر و دقیق‌تر، می‌توان از سایر توزیع‌های پیشین نیز استفاده کرد که به دلیل پرهیز از حجیم شدن پایان‌نامه از این کار خودداری کردیم.

جدول ۱۱.۳: نتایج برازش مدل دوم با مجموعه پیشین‌های (۵.۳)

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
β_1	-2	$-1/9329$	۰/۰۶۱۳	$(-2/056, -1/814)$
β_2	۱	۰/۹۹۷	۰/۰۵۷۹	$(0/883, 1/109)$
β_3	۲	۱/۸۷	۰/۰۶۹۹	$(1/742, 2/008)$
δ_1	۳/۵۵	۳/۷۰۸۳	۰/۱۵۹	$(3/382, 4/028)$
δ_2	۰/۴۱	۰/۳۰۹۴	۰/۲۰۷	$(-0/083, 0/715)$
δ_3	$-2/01$	$-2/4101$	۰/۱۹۵	$(-2/825, -1/994)$
$\Sigma_{b_{11}}$	۱/۰۹	۱/۲۸۴۳	۰/۱۲۸	$(1/0604, 1/552)$
$\Sigma_{b_{12}}$	$-0/36$	$-0/2864$	۰/۰۵۳۱	$(-0/388, -0/189)$
$\Sigma_{b_{22}}$	۰/۱۳	۰/۳۸۵۷	۰/۰۳۸۲	$(0/319, 0/468)$

جدول ۱۲.۳: نتایج برازش مدل دوم با مجموعه پیشین‌های (۶.۳)

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
β_1	-2	$-1/93345$	۰/۰۵۹۵	$(-2/0536, -1/821)$
β_2	۱	۰/۹۹۳	۰/۰۵۶۹	$(0/8854, 1/1112)$
β_3	۲	۱/۸۷۰۹	۰/۰۷۰۷	$(1/7364, 2/01)$
δ_1	۳/۵۵	۳/۷۴۵۱	۰/۱۵۸۹	$(3/4255, 4/0502)$
δ_2	۰/۴۱	۰/۲۸۰۶	۰/۲۰۵۷	$(-0/1169, 0/6858)$
δ_3	$-2/01$	$-2/4429$	۰/۲۰۲	$(-2/8321, -2/0098)$
$\Sigma_{b_{11}}$	۱/۰۹	۱/۲۹۱۷	۰/۱۲۵۹	$(1/0617, 1/5575)$
$\Sigma_{b_{12}}$	$-0/36$	$-0/288$	۰/۰۵۴۵	$(-0/4069, -0/1863)$
$\Sigma_{b_{22}}$	۰/۱۳	۰/۳۷۶۴	۰/۰۳۷۳	$(0/2211, 0/4672)$

فصل ۴

مثال‌های کاربردی

در این فصل، نحوه به‌کارگیری و برازش بیزی مدل‌های رگرسیونی آمیخته بتا را در موقعیت‌های کاربردی تشریح می‌کنیم. این موقعیت‌ها دو مثال واقعی شامل داده‌های نسبت بنزین تبدیل‌شده از نفت خام و درصد مقاومت بتن را شامل می‌شوند.

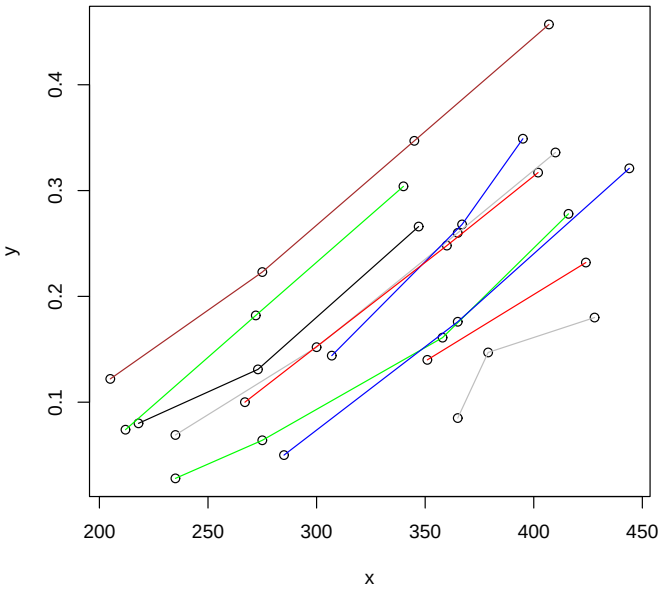
۱.۴ داده‌های نسبت بنزین تبدیل‌شده از نفت

مجموعه داده‌های بنزین اولین بار توسط پراتر (۱۹۵۶) گزارش شده است. در این داده‌ها، متغیر پاسخ y نسبت نفت خام تبدیل‌شده به بنزین بعد از فرآیند تقطیر و شکنش است و متغیرهای تبیینی، x_i، $i = ۱, \ldots, ۱۰$، درجه‌های حرارت تبخیر بنزین می‌باشند. بعد از مرتب کردن داده‌ها مشاهده می‌کنیم که مدل دارای ۱۰ خوشه است و هر خوشه شامل ۲، ۳ یا ۴ مشاهده می‌باشند. همچنین در این مجموعه داده، تعداد مشاهدات برابر با ۳۲ می‌باشد.

برای تحلیل این داده‌ها، فراری و سریباری (۲۰۰۴) یک مدل رگرسیون بتا را با پارامتر دقت ثابت برازش کردند که در آن دسته‌های نفت خام را به عنوان یک عامل ثابت با ۱۰ سطح و یک شیب ثابت مورد بررسی قرار دادند. در مقابل، ونابلز (۲۰۰۰) دسته‌ها را به عنوان یک عامل تصادفی در نظر گرفت. نمودار پراکنش[۱] داده‌ها، نمودار ۱.۴، نشان می‌دهد که یک مدل رگرسیونی برای میانگین با عرض از مبدا تصادفی و یک شیب ثابت برای داده‌ها می‌تواند مناسب باشد. بنابراین یک مدل رگرسیون آمیخته

[۱] Scatter Plot

بتا برای این داده‌ها می‌تواند مناسب باشد. ما در این بخش از هر دو مدل با پارامتر دقت ثابت و متغیر، برای تحلیل آن‌ها استفاده می‌کنیم.



شکل ۱.۴: نمودار پراکنش نسبت نفت خام تبدیل‌شده به بنزین در مقابل درجه حرارت‌های مختلف تبخیر

۱.۱.۴ مدل با پارامتر دقت ثابت

با توجه به نمودار پراکنش داده‌ها، اگرچه به نظر می‌رسد برای همه درجه‌های حرارت، شیب‌ها ثابت هستند، اما ما در این بخش مدل رگرسیونی برای میانگین پاسخ را با عرض از مبدا و شیب تصادفی به صورت

$$logit(\mu_{ij}) = (\beta_۱ + b_{i۱}) + (\beta_۲ + b_{i۲})x_{ij} \qquad (۱.۴)$$

برای $i = ۱, \ldots, ۱۰$، $j = ۱, \ldots, n_i$، $n_i = (۴, ۳, ۳, ۴, ۳, ۳, ۴, ۳, ۲, ۳)$، در نظر گرفتیم. پارامترهای Σ_b و $\boldsymbol{\beta}$ دارای توزیع‌های پیشین

$$\boldsymbol{\beta} \sim N_۲(\mu_\beta, \Sigma_\beta), \quad \Sigma_b \sim IW_۲(\Psi, c),$$

با مقادیر

$$\Sigma_\beta = \begin{pmatrix} ۵ & ۰ \\ ۰ & ۵ \end{pmatrix}, \Psi = \begin{pmatrix} ۱۰۰۰ & ۰ \\ ۰ & ۱۰۰۰ \end{pmatrix}, \mu_\beta = (۰, ۰)^T, c = ۴,$$

هستند. همانند فصل قبل توزیع‌های پیشین مختلفی را برای پارامتر دقت در نظر گرفتیم. این توزیع‌های پیشین در جدول ۱.۴ آمده‌اند. در نهایت مدل اول را برای هر یک از این توزیع‌ها و با در نظر گرفتن دو زنجیر با نقاط آغازی مختلف و ۲۰۰۰۰۰ تکرار مونت کارلویی برازش دادیم.

با توجه به مقادیر DIC محاسبه‌شده برای هر یک از مدل‌های موجود در جدول ۱.۴، نتیجه می‌گیریم که مدل با پیشین $b\cdot۴$ با کمترین DIC، از برازش بهتری نسبت به سایر مدل‌ها برخوردار است. بنابراین این پیشین را برای ϕ انتخاب کرده و مدل (۱.۴) را برازش دادیم. نتایج برازش این مدل در جدول ۲.۴ آمده‌اند که شامل میانگین پسین، انحراف معیار و فاصله اعتبار ٪۹۵ برای هر پارامتر می‌باشند. پس از انجام آزمون‌های همگرایی، گلمن-روبین و جی‌وک، که نتایج آن در جدول ۳.۴ گزارش شده‌اند، درمی‌یابیم که زنجیر تولیدشده برای هر پارامتر، جز $\beta_۱$، همگرا است، زیرا $PSRF$ برای هر پارامتر نزدیک به یک و $|Z_n|$ برای هر یک مقداری کوچک را نشان می‌دهند.

جدول ۱.۴: توزیع‌های پیشین ϕ و مقادیر DIC در مدل (۱.۴) برای داده‌های بنزین

مدل	توزیع پیشین برای ϕ	DIC
$b\cdot۱$	$\phi \sim IG(۰/۰۰۱, ۰/۰۰۱)$	$-۲۹۵/۴$
$b\cdot۲$	$\phi = u^۲, u \sim U(۰, ۵۰)$	$-۲۹۶/۳$
$b\cdot۳$	$\phi = (۵۰B)^۲, B \sim beta(۱/۱, ۱/۱)$	$-۲۹۶/۸$
$b\cdot۴$	$\phi = (۵۰B)^۲, B \sim beta(۱/۵, ۱/۵)$	$-۲۹۸/۲$

مقادیر ESS محاسبه‌شده برای هر یک از پارامترها در جدول ۴.۴ گزارش شده‌اند. این مقادیر، برای هر یک از پارامترها نزدیک به تعداد نمونه‌های تولیدشده از توزیع پسین به‌دست آمده است (با توجه به تعداد تکرارها و تعداد زنجیرها، تعداد نمونه‌های تولیدشده از توزیع پسین برابر با ۴۰۰۰ است). بنابراین می‌توان نتیجه گرفت نمونه‌های تولیدشده از کیفیت قابل قبولی برخوردارند.

همچنین همگرایی زنجیر به توزیع مانا برای هر یک از پارامترها را با رسم نمودارهای میانگین تجمعی، جی‌وک، خودهمبستگی و اثر بررسی کردیم. با استفاده از نمودار ۲.۴، نمودار میانگین تجمعی، می‌توان همگرایی هر پارامتر، جز $\beta_۱$، به مقدار واقعی آن را مشاهده کرد. با توجه به نمودارهای جی‌وک، شکل

جدول ۲.۴: نتایج برازش مدل (۱.۴) با پیشین $b \cdot ۴$ برای ϕ برای داده‌های بنزین

پارامترها	استنباط پسین		
	میانگین	انحراف معیار	فاصله اعتبار ٪۹۵
$\beta_۱$	$-۰/۶۶۲۵$	$۰/۵۹۳۶$	$(-۱/۲۷۱۹, -۰/۰۵۴۴)$
$\beta_۲$	$۰/۰۱۰۹۸$	$۰/۰۰۰۴۶$	$(۰/۰۱۰۱, ۰/۰۱۱۹)$
ϕ	$۳۰۶/۷۶۹۲$	$۹۰/۷۳$	$(۱۵۳/۵۳۴, ۵۰۵/۶)$
$\Sigma_{b_{۱۱}}$	$۱۱۰/۴۷۹۷$	$۵۳/۴۳$	$(۵۰/۳۸۶۲, ۲۵۰/۰۹۱)$
$\Sigma_{b_{۱۲}}$	$-۷/۶۲۵۸$	$۹۹/۲۷$	$(-۱۸۸/۸۶۱, ۱۴۹/۰۷۱)$
$\Sigma_{b_{۲۲}}$	$۵۰۳/۳۱۵۲$	$۶۶۶/۳$	$(۹۴/۲۳۴۲, ۱۸۰۹)$

جدول ۳.۴: نتایج آزمون‌های همگرایی برای مدل (۱.۴) با پیشین $b \cdot ۴$ برای ϕ برای داده‌های بنزین

آماره آزمون	پارامترهای مدل					
	$\beta_۱$	$\beta_۲$	ϕ	$\Sigma_{b_{۱۱}}$	$\Sigma_{b_{۱۲}}$	$\Sigma_{b_{۲۲}}$
گلمن-روبین	$۲۰۹/۰۱$	$۱/۰۲$	$۱/۰۰$	$۱/۰۱$	$۱/۰۳$	$۱/۰۸$
جی‌وک						
زنجیر ۱	$۳/۳۲۶$	$۰/۰۲۶۱$	$۰/۲۶۱$	$۰/۴۵۵$	$-۰/۹۰۴$	$-۰/۸۲۷$
زنجیر ۲	$-۸/۹۳۶$	$-۰/۲۸۴$	$۱/۳۳۲$	$۱/۴۲۲$	$-۲/۳۶۸$	$-۰/۰۸۰۲$

جدول ۴.۴: مقادیر ESS برای پارامترهای مدل (۱.۴) با توزیع پیشین $b \cdot ۴$ برای داده‌های بنزین

پارامترها	$\beta_۱$	$\beta_۲$	ϕ	$\Sigma_{b_{۱۱}}$	$\Sigma_{b_{۱۲}}$	$\Sigma_{b_{۲۲}}$
ESS	$۱۸/۹۲۳$	$۳۷۵۶/۵۷۳$	$۳۸۵۷/۷۸۱$	۴۰۰۰	۴۰۰۰	$۳۶۸۳/۸۰۳$

۳.۴، مشاهده می‌کنیم که تعداد زیادی از Z_nهای محاسبه‌شده در فاصله مشخص‌شده قرار دارند. بنابراین می‌توان گفت نمونه‌های تولیدشده از توزیع مانای زنجیر، گرفته شده‌اند. نمودارهای خودهمبستگی، شکل ۴.۴، نشان می‌دهند که زنجیر برای هر پارامتر ناهمبسته است. با مشاهده نمودار ۵.۴ می‌توان گفت زنجیر برای همه پارامترها از آمیختگی خوبی برخوردار است، یعنی نواحی چگال پسین به خوبی مورد کاوش قرار گرفته‌اند. شکل ۶.۴، نمودارهای چگالی پسین حاشیه‌ای برای هر پارامتر را نشان می‌دهد. البته در این نمودارها نیز می‌توان به وضوح مشاهده کرد که پارامتر β_1 آمیختگی مناسبی نسبت به سایر پارامترها ندارد. به عبارت دیگر، فضای حالت زنجیر به خوبی مورد کاوش قرار نگرفته است.

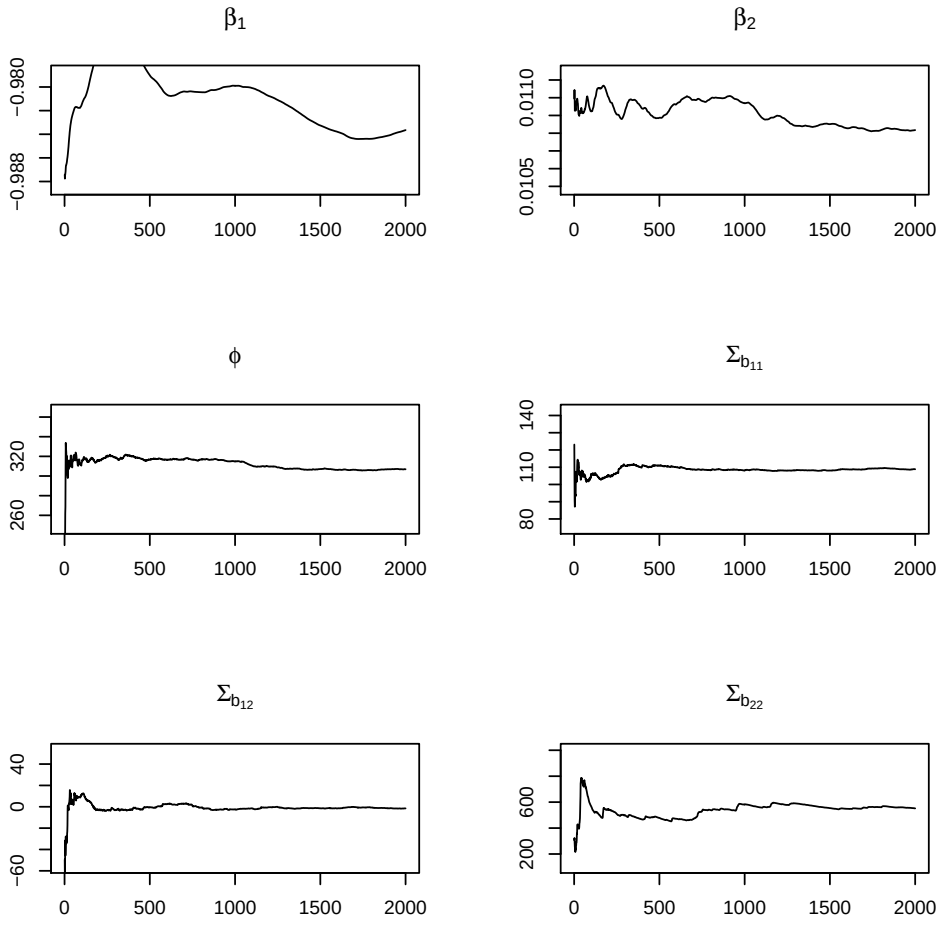
پارامتر β_1 که می‌توان آن را عرض از مبدا مشترک برای همه سطوح اثرات تصادفی معرفی کرد، در مدل‌بندی‌های کاربردی دارای اهمیت نیست. بنابراین می‌توان گفت اگرچه مدل برازش‌شده نتایج قابل قبولی برای β_1 ندارد، اما با توجه به نقش آن، این موضوع چندان ناخوشایند نیست. البته شاید این نتایج به دلیل غیر قابل شناسایی بودن پارامتر β_1 نیز باشد که نیاز به بررسی‌های بیشتری دارد که از حوصله این پایان‌نامه خارج است.

۲.۱.۴ مدل با پارامتر دقت متغیر

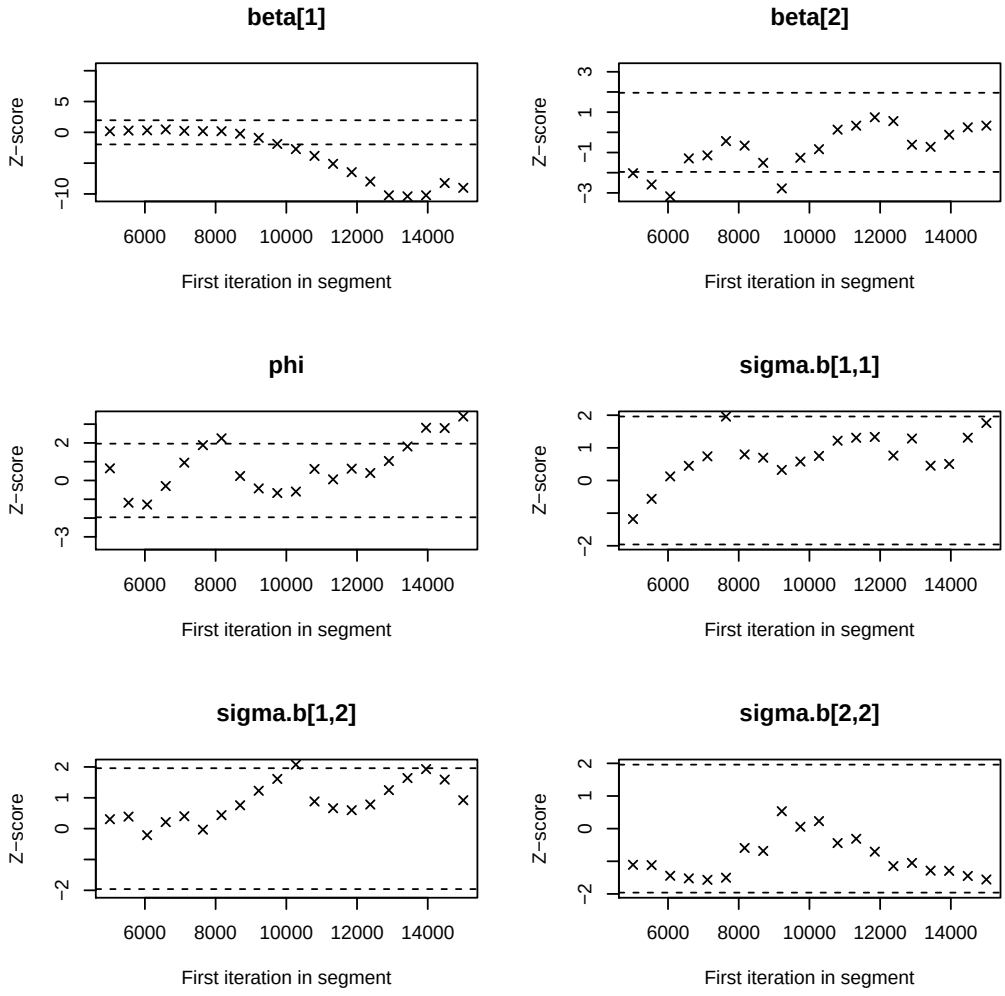
در این بخش، فرض می‌کنیم که پارامتر دقت به متغیرهای تبیینی وابسته است. مدل‌های رگرسیونی پارامتر دقت با تابع پیوند لگاریتمی، براساس ثابت یا تصادفی بودن عرض از مبدا و شیب، در جدول ۵.۴ آمده‌اند. توزیع پیشین پارامترهای $\boldsymbol{\beta}$ و Σ_b نیز مانند بخش قبل است. همچنین مدل رگرسیونی میانگین پاسخ دارای همان ساختار (۱.۴) است.

مدل دوم را به ازای هر یک از مدل‌های مختلف ϕ برازش دادیم و مقدار DIC نیز برای هر یک محاسبه شد. این مقادیر در جدول ۵.۴ آمده‌اند. با مشاهده مقادیر DIC، مدل $c\cdot 5$، که یک مدل با عرض از مبدا تصادفی و شیب ثابت برای داده‌ها است، انتخاب می‌شود. نتایج برازش مدل $c\cdot 5$ نیز در جدول ۶.۴ گزارش شده‌اند. با بررسی نتایج آزمون‌های همگرایی گلمن-روبین و جی‌وک، جدول ۷.۴، نتیجه می‌گیریم که زنجیر برای هر پارامتر، جز β_1، همگرا است. همچنین کیفیت این نمونه‌های تولیدشده را می‌توان با محاسبه ESS، که نتایج آن در جدول ۸.۴ آمده است، مشاهده کرد. با توجه به این مقادیر، می‌توان گفت نمونه‌های تولیدشده در زنجیرها برای هر یک از پارامترها، جز β_1، از کیفیت قابل قبولی برخوردارند.

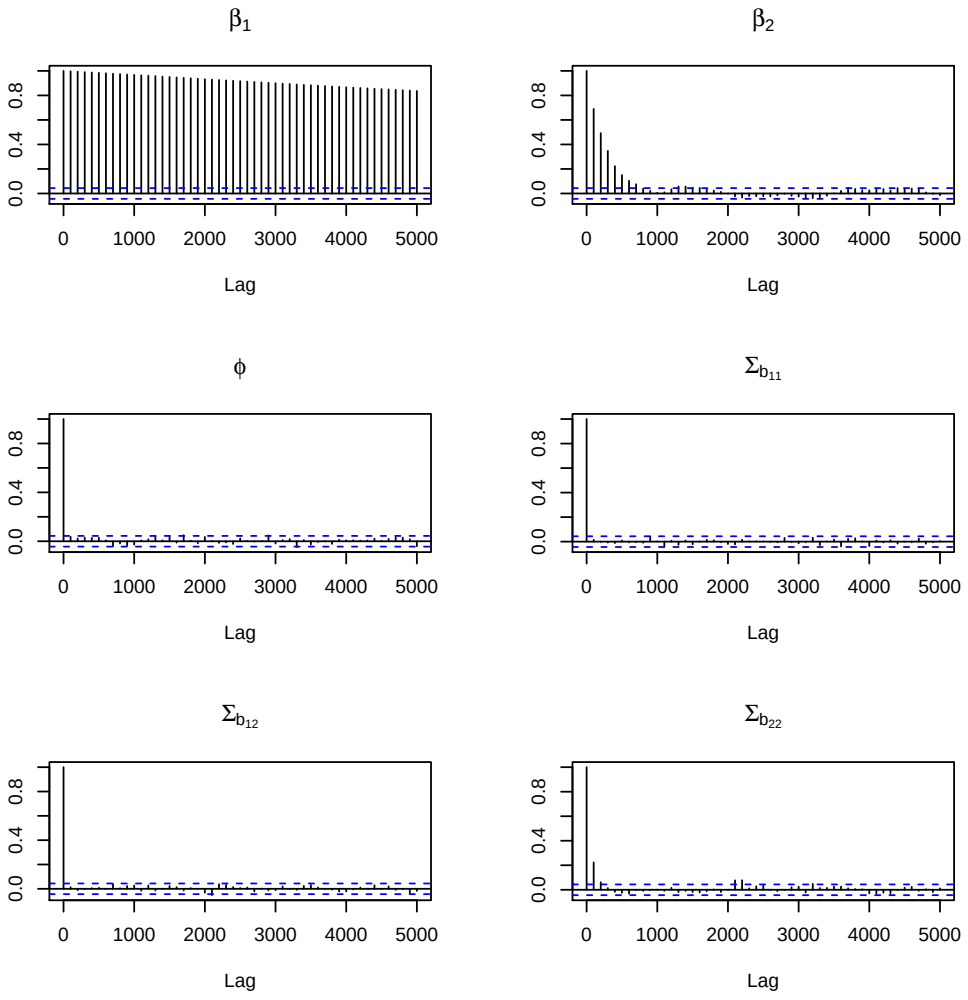
با توجه به نتایج حاصل از هر دو مدل با پارامتر دقت ثابت و متغیر، می‌توان گفت که افزایش درجه حرارت تبخیر، به‌طور متوسط، بر افزایش نسبت نفت خام تبدیل‌شده به بنزین تاثیر مستقیم دارد. این



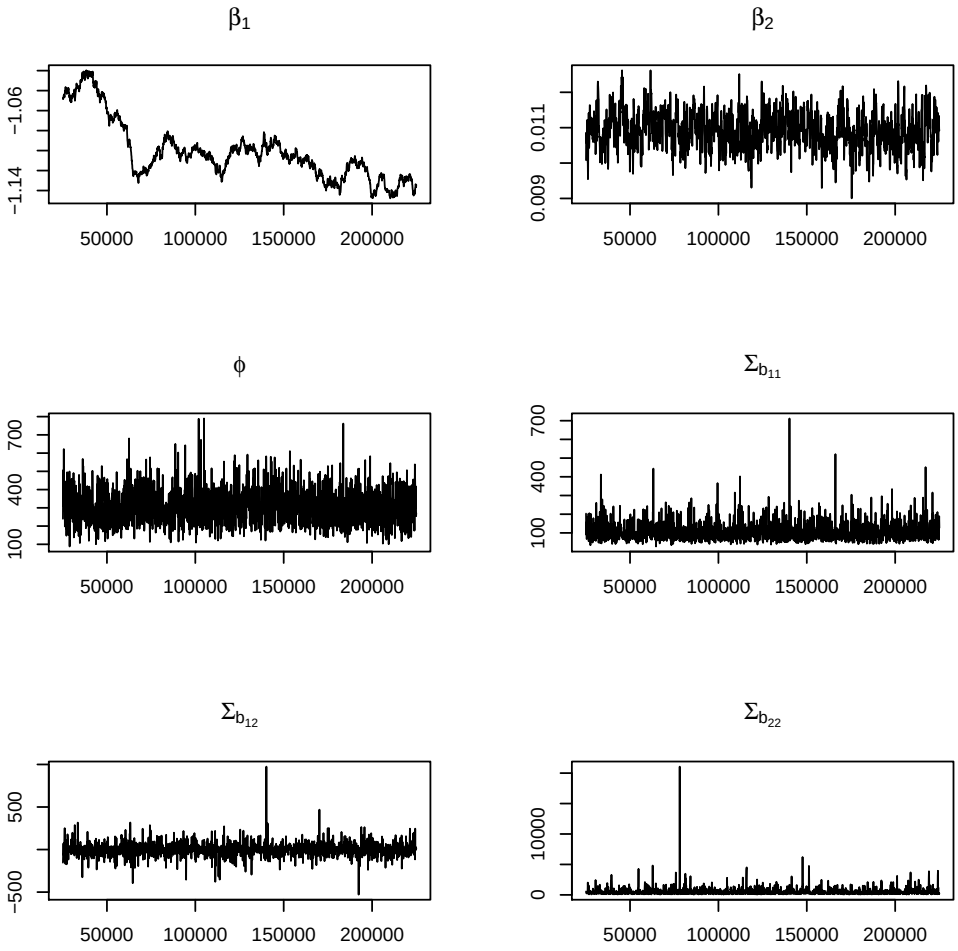
شکل ۲.۴: نمودارهای میانگین تجمعی پارامترهای مدل (۱.۴) با پیشین $b \cdot ۴$ برای ϕ برای داده‌های بنزین



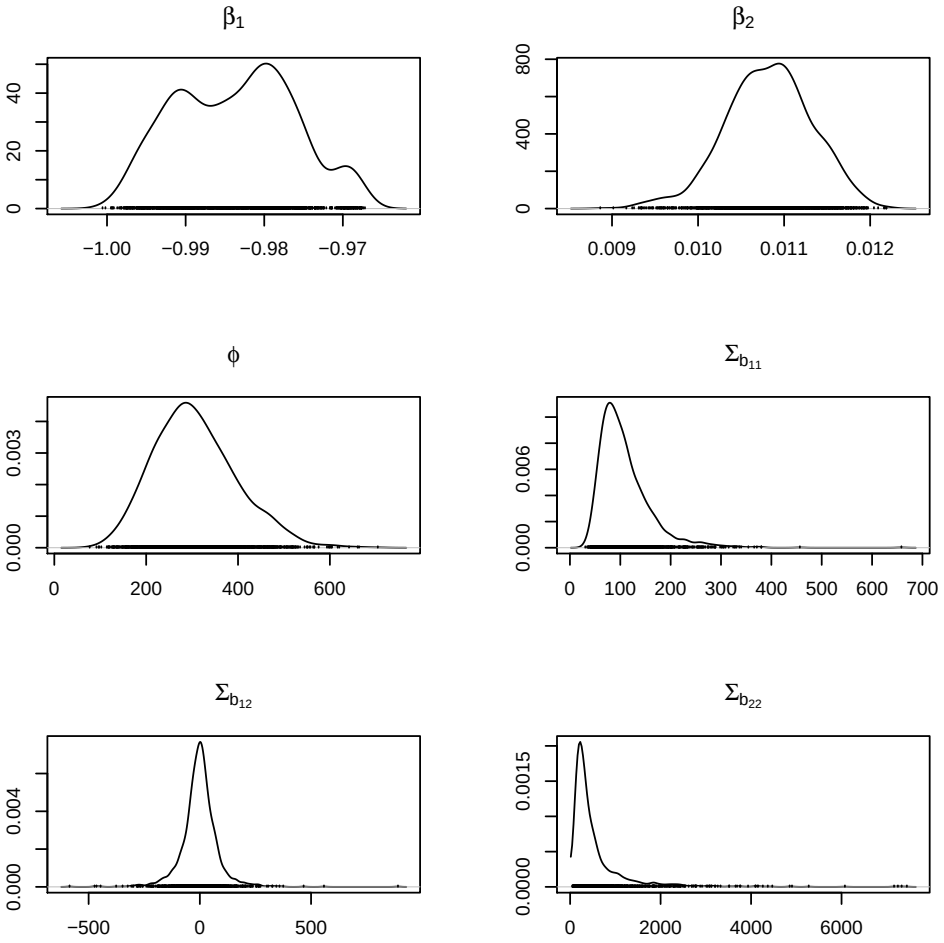
شکل ۳.۴: نمودارهای جی‌وک پارامترهای مدل (۱.۴) با پیشین $b \cdot 4$ برای ϕ برای داده‌های بنزین



شکل ۴.۴: نمودارهای خودهمبستگی پارامترهای مدل (۱.۴) با پیشین $b \cdot ۴$ برای ϕ برای داده‌های بنزین



شکل ۵.۴: نمودارهای اثر پارامترهای مدل (۱.۴) با پیشین $b \cdot ۴$ برای ϕ برای داده‌های بنزین



شکل ۶.۴: نمودارهای چگالی پسین حاشیه‌ای پارامترهای مدل (۱.۴) با پیشین $b \cdot ۴$ برای ϕ برای داده‌های بنزین

جدول ۵.۴: مدل‌های مختلف ϕ و مقادیر DIC برای هر یک برای داده‌های بنزین

مدل	$\ln(\phi_{ij})$	DIC
$c\cdot ۱$	$\delta_۱$	$-۲۹۲/۷$
$c\cdot ۲$	$\delta_۱ + d_{i۱}$	$-۲۰۰/۹۳$
$c\cdot ۳$	$\delta_۲ x_{ij}$	-۳۰۰
$c\cdot ۴$	$\delta_۱ + \delta_۲ x_{ij}$	$-۲۹۷/۲$
$c\cdot ۵$	$d_{i۱} + \delta_۲ x_{ij}$	-۳۱۵
$c\cdot ۶$	$(\delta_۱ + d_{i۱}) + \delta_۲ x_{ij}$	$-۳۰۶/۸$

نتیجه با علامت مثبت ضریب برآوردشده قابل حصول است. همچنین با توجه به مثبت بودن ضریب $\delta_۲$ در مدل $c\cdot ۵$، می‌توان گفت میزان دقت این تاثیر نیز با افزایش میزان درجه حرارت تبخیر، افزایش می‌یابد.

۲.۴ داده‌های مقاومت بتن

یکی از مهم‌ترین و متداول‌ترین مصالح ساختمانی، بتن است که به علت دارا بودن خواصی از جمله مقاومت خوب در برابر آتش‌سوزی، دسترسی آسان به مصالح و مقاومت فشاری خوب، استفاده از آن را با مقبولیت عمومی روبرو کرده است.

بتن مصالحی شبیه به سنگ است که از مخلوط کردن مقدار متناسبی از سیمان، ماسه، آب و افزودنی‌های دیگر به‌دست می‌آید. توده اصلی بتن، سنگ‌دانه‌های ریز و درشت (شن و ماسه) است و فعل و انفعالات شیمیایی بین آب و سیمان که به صورت شیره‌ای اطراف سنگدانه‌ها را پوشانده است، باعث یکپارچه شدن و چسبیدن سنگدانه‌ها به یکدیگر می‌شود. این سنگدانه‌ها اسکلت اصلی بتن را تشکیل داده و نیروی وارد بر بتن را تحمل می‌کنند. آب نیز در این مخلوط موجب واکنش شیمیایی در سیمان می‌شود که سخت شدن مخلوط بتن را پس از رسیدن به مقاومت نهایی بتن به همراه دارد.

یکی از معایب مهم ساختمان‌های بتنی وزن بسیار زیاد ساختمان می‌باشد که با میزان تخریب ساختمان بر اثر زلزله نسبت مستقیم دارد. اگر بتوانیم تیغه‌های جداکننده و پانل‌ها را از بتن سبک بسازیم، وزن ساختمان و در نتیجه تخریب ساختمان توسط زلزله مقدار زیادی کاهش می‌یابد. ولی کم بودن مقاومت

جدول ۶.۴: نتایج برازش مدل $c\cdot ۵$ برای ϕ برای داده‌های بنزین

پارامترها	استنباط پسین		
	میانگین	انحراف معیار	فاصله اعتبار ٪۹۵
$\beta_۱$	$-۰/۴۸۷۳$	$۰/۴۹۲۷$	$(-۰/۹۸, ۰/۰۱۲۴)$
$\beta_۲$	$۰/۰۰۹۵$	$۰/۰۰۱۲$	$(۰/۰۰۸۴, ۰/۰۱۱۶)$
$\delta_۲$	$۰/۰۱۸۱$	$۰/۰۰۹۳$	$(۰/۰۰۰۳, ۰/۰۳۴۹)$
$\Sigma_{b_{۱۱}}$	$۶۳/۵۹$	$۲۱/۳۹$	$(۱۴/۱۷, ۳۱۴/۶)$
$\Sigma_{b_{۱۲}}$	$-۶/۱۰۸۸$	$۴۴/۲۴$	$(-۸۶/۳۳۱, ۸۷/۷)$
$\Sigma_{b_{۲۲}}$	$۴۹۷/۴۱۶$	۵۹۴	$(۹۸/۲۱۳, ۲۰۳۴)$

جدول ۷.۴: نتایج آزمون‌های همگرایی برای مدل $c\cdot ۵$ برای ϕ برای داده‌های بنزین

آماره آزمون	پارامترهای مدل					
	$\beta_۱$	$\beta_۲$	$\delta_۲$	$\Sigma_{b_{۱۱}}$	$\Sigma_{b_{۱۲}}$	$\Sigma_{b_{۲۲}}$
گلمن-روبین	۵۷۸/۲۵	۱/۰۲	۱/۰۱	۱/۰۲	۱/۰۵	۱/۰۸
جی‌وک						
زنجیر ۱	$-۴/۱۹۱$	$-۰/۵۴۳$	$-۰/۳۵۲$	۱/۵۶۷	۱/۹۳۳	۱/۲۸۹
زنجیر ۲	$-۲/۲۸۶$	$-۰/۵۲۱$	۲/۵۸۲	$-۰/۱۱۶$	$-۰/۱۰۴$	$-۲/۰۸۶$

جدول ۸.۴: مقادیر ESS برای پارامترهای مدل $c\cdot ۵$ برای ϕ برای داده‌های بنزین

پارامترها	$\beta_۱$	$\beta_۲$	$\delta_۲$	$\Sigma_{b_{۱۱}}$	$\Sigma_{b_{۱۲}}$	$\Sigma_{b_{۲۲}}$
ESS	۹/۴۷۸	۴۰۰۰	۳۷۶۳/۴۰۲	۳۹۲۲/۶۵۱	۳۵۳۹/۶۶	۳۷۷۷/۹۰۲

بتن سبک عامل مهمی در محدود نمودن دامنه کاربرد این نوع بتن و بهره‌گیری از امتیازات آن بوده است. استفاده از میکروسیلیس در ساخت بتن سبک سبب شده است که مقاومت بتن سبک بالا رود و این محدودیت نیز کاهش یابد.

مجموعه داده‌های مقاومت بتن توسط رضایی و همکاران (۱۳۹۱) گزارش شده است. با توجه به این داده‌ها، متغیرهای تبیینی مورد بررسی عبارتند از ماسه[۲]، آب[۳]، سیمان[۴]و میکروسیلیس[۵]. متغیر پاسخ میزان مقاومت بتن[۶] در سه زمان مختلف (میزان مقاومت بتن پس از ۷، ۲۸ و ۹۰ روز) می‌باشد. بنابراین یک مجموعه داده طولی داریم که برای هر مورد (آزمایش) سه تکرار در طول زمان اندازه‌گیری شده است. همچنین داده‌ها شامل ۱۰۸ مشاهده در هر زمان می‌باشند. همانند مثال اول، این مجموعه داده را نیز برای دو حالت پارامتر دقت ثابت و متغیر تحلیل می‌کنیم. با رسم نمودار پراکنش هر یک از متغیرهای تبیینی در مقابل متغیر پاسخ (در هر سه زمان)، نمودارهای ۷.۴ تا ۱۰.۴، می‌توان مشاهده کرد که اثرات هر کدام از متغیرهای تبیینی در زمان‌های مختلف می‌توانند متفاوت باشند. بنابراین، برای مدل‌بندی میانگین پاسخ از یک مدل با پارامترهای تصادفی برای همه متغیرهای تبیینی استفاده می‌کنیم.

۱.۲.۴ مدل با پارامتر دقت ثابت

در این مدل فرض می‌کنیم که پارامتر دقت ثابت است و دارای توزیع‌های پیشین مختلف گزارش‌شده در جدول ۹.۴ می‌باشد. همچنین، این مدل شامل یک مدل رگرسیونی برای میانگین با ساختار

$$logit(\mu_{ij}) = (\beta_۱ + b_{i۱}) + (\beta_۲ + b_{i۲})S_{ij} + (\beta_۳ + b_{i۳})W_{ij} + (\beta_۴ + b_{i۴})SF_{ij} + (\beta_۵ + b_{i۵})C_{ij},$$

برای $j = ۱, ۲, ۳$ و $i = ۱, \ldots, ۱۰۸$ می‌باشد، که در آن $\mathbf{b}_i = (b_{i۱}, \ldots, b_{i۵})^T \sim N_۵(۰, \Sigma_b)$. پارامترهای $\boldsymbol{\beta} = (\beta_۱, \ldots, \beta_۵)^T$ و Σ_b دارای توزیع‌های پیشین

$$\boldsymbol{\beta} \sim N_۵(۰, \Sigma_\beta), \quad \Sigma_b \sim IW_۵(\Psi, c),$$

---

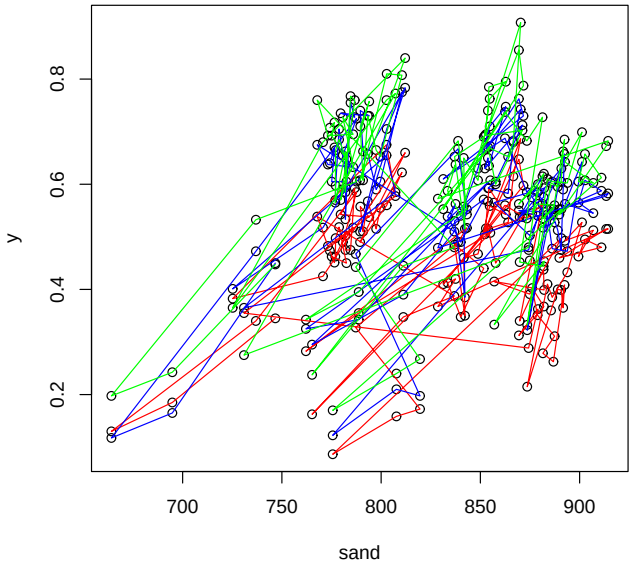
[۲] Sand

[۳] Water

[۴] Cement

[۵] Silicafume

[۶] Stre



شکل ۷.۴: نمودار پراکنش میزان مقاومت بتن در هر سه زمان در مقابل متغیر ماسه

با مقادیر

$$\Sigma_\beta = ۵I_۵, \Psi = ۱۰۰۰I_۵, c = ۲۵,$$

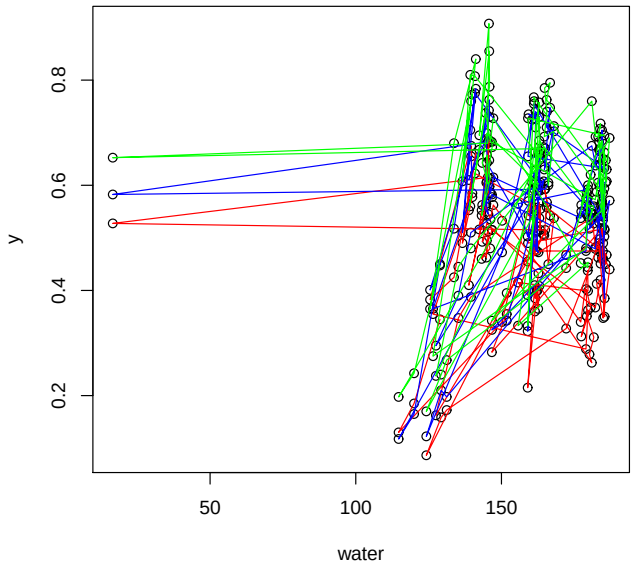
هستند، که در آن I_ℓ ماتریس همانی با بعد ℓ می‌باشد.

با مشاهده مقادیر DIC محاسبه‌شده برای هر یک از مدل‌ها می‌توان گفت مدل با توزیع پیشین $d\cdot ۳$ نسبت به سایر مدل‌ها از برازش بهتری برخوردار است. جدول ۱۰.۴ نتایج برازش این مدل را نشان می‌دهد. بنابر آزمون‌های همگرایی گلمن-روبین و جی‌وک، جدول ۱۱.۴، می‌توان گفت زنجیر برای هر پارامتر، جز $\beta_۱$، همگرا است.

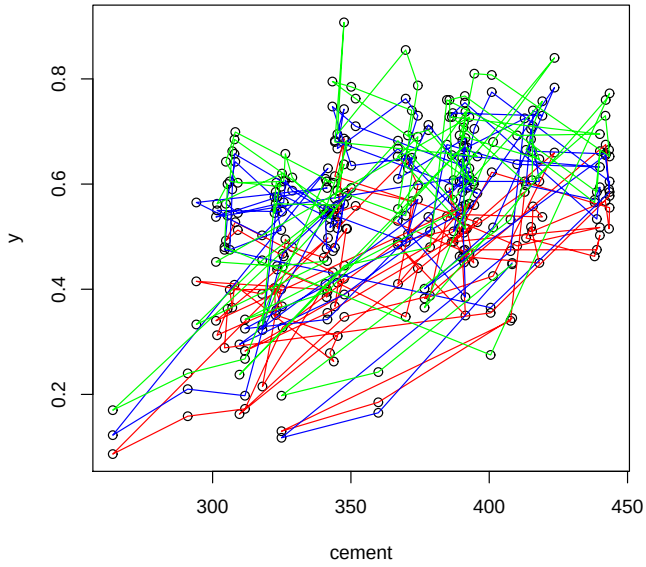
به دلیل حجیم نشدن پایان‌نامه، از گزارش روش‌های تشخیص نموداری خودداری کردیم. با این حال، تمامی نمودارها، تاییدکننده ویژگی‌های خوب آمیختگی و همگرایی زنجیر پارامترها هستند.

۲.۲.۴ مدل با پارامتر دقت متغیر

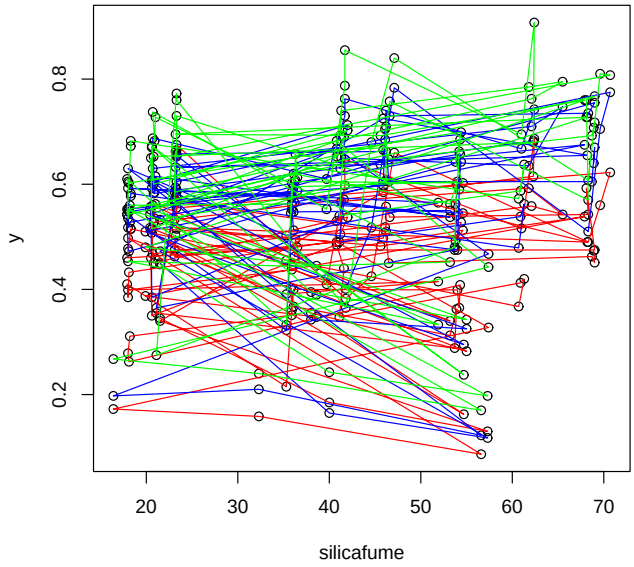
در این قسمت، مدل را بر اساس پارامتر دقت متغیر برازش دادیم. برای درک آن‌که کدام متغیرهای تبیینی می‌توانند بر پارامتر دقت تاثیر داشته باشند، از تحلیل اکتشافی استفاده کردیم. در تحلیل اکتشافی، پژوهشگر به دنبال بررسی داده‌های تجربی به منظور کشف و شناسایی شاخص‌ها و نیز روابط بین آن‌ها



شکل ۸.۴: نمودار پراکنش میزان مقاومت بتن در هر سه زمان در مقابل متغیر آب



شکل ۹.۴: نمودار پراکنش میزان مقاومت بتن در هر سه زمان در مقابل متغیر سیمان



شکل ۱۰.۴: نمودار پراکنش میزان مقاومت بتن در هر سه زمان در مقابل متغیر میکروسیلیس

است و این کار را بدون تحلیل هرگونه مدل معینی انجام می‌دهد. به بیان دیگر تحلیل اکتشافی علاوه بر آن‌که ارزش تجسسی یا پیشنهادی داشته باشد، می‌تواند ساختارساز، مدل‌ساز یا فرضیه‌ساز باشد و زمانی به‌کار می‌رود که پژوهشگر شواهد کافی و قبلی و پیش‌تجربی برای تشکیل فرضیه درباره تعداد عامل‌های زیربنایی داده‌ها نداشته و برای تعیین عامل‌هایی که هم‌پراشی بین متغیرها را توجیه می‌کنند، داده‌ها را می‌کاود. بنابراین تحلیل اکتشافی بیشتر به‌عنوان یک روش تدوین و تولید مدل و نه یک روش آزمون مدل در نظر گرفته می‌شود. پس مطلوب آن است که فرضیه‌ای که از طریق روش‌های تحلیل اکتشافی تدوین می‌شود، از طریق قرار گرفتن در معرض روش‌های آماری دقیق‌تر تایید یا رد شود. در تحلیل اکتشافی، متغیرهای معلوم باید با دقت زیادی انتخاب شوند تا حتی‌الامکان درباره متغیرهای نامعلومی که استخراج می‌شوند، اطلاعات بیشتری فراهم آید. همچنین تحلیل اکتشافی نیازمند نمونه‌هایی با حجم بسیار زیاد می‌باشد.

برای شناسایی متغیرهای تبیینی که بر پارامتر دقت اثر می‌گذارند، نمودارهای جعبه‌ای متغیر پاسخ را به همراه نمودار چگالی آن‌ها برای دسته‌های مختلف هر متغیر تبیینی رسم کردیم. در این نمودارها داده‌ها بر طبق مشاهدات نزدیک به هم گروه‌بندی و رسم می‌شوند. با استفاده از این نمودار می‌توان مرکزیت، پراکندگی و چولگی داده‌ها را در گروه‌های مختلف تفسیر نمود. در مقایسه نمودار جعبه‌ای پاسخ در دو دسته از مشاهدات متغیرهای تبیینی، می‌توان پراکندگی آن‌ها را با توجه به طول مستطیل‌های نمودار با

جدول ۹.۴: توزیع‌های پیشین ϕ و مقادیر DIC برای هر یک برای داده‌های مقاومت بتن

مدل	توزیع پیشین ϕ	DIC
$d\cdot ۱$	$\phi \sim IG(۰/۰۰۱, ۰/۰۰۱)$	$-۱۶۸۲/۸$
$d\cdot ۲$	$\phi = u^۲, u \sim U(۰, ۵۰)$	$-۱۶۷۳/۸$
$d\cdot ۳$	$\phi = (۵۰B)^۲, B \sim beta(۱/۱, ۱/۱)$	$-۱۷۰۲/۷$
$d\cdot ۴$	$\phi = (۵۰B)^۲, B \sim beta(۱/۵, ۱/۵)$	$-۱۶۸۶/۴$

یکدیگر مقایسه نمود. مستطیلی که طول بزرگتری دارد دارای پراکندگی بیشتری نیز می‌باشد.

نمودارهای ۱۱.۴ تا ۱۴.۴، نمودارهای جعبه‌ای و چگالی پاسخ برای دسته‌های مختلف هر یک از متغیرهای تبیینی در زمان اول را نمایش می‌دهند. با مشاهده این نمودارها می‌توان به این نتیجه رسید که متغیرهای ماسه، آب و سیمان می‌توانند بر واریانس (دقت) پاسخ اثرگذار باشند. از طرفی در گروه‌های مختلف متغیر میکروسیلیس، واریانس پاسخ تقریبا یک روند مشابه دارد. با توجه به این‌که متغیر پاسخ در سه زمان مختلف اندازه‌گیری شده است، نمودارهای جعبه‌ای و چگالی مذکور در زمان‌های دوم و سوم نیز ترسیم شدند. نمودارهای ۱۵.۴ و ۱۶.۴ مربوط به متغیر میکروسیلیس می‌باشند. حال می‌توان این‌طور نتیجه گرفت که متغیر میکروسیلیس در هر سه زمان تاثیری بر واریانس یا دقت پاسخ ندارد. بنابراین مدل نهایی برای پارامتر دقت شامل متغیرهای ماسه، آب و سیمان خواهد بود. در جدول ۱۲.۴ ساختارهای رگرسیونی مختلف با تابع پیوند لگاریتمی برای ϕ، براساس ثابت یا تصادفی بودن عرض از مبدا و شیب، آمده‌اند. همچنین، مشابه مدل اول، این مدل نیز شامل یک ساختار رگرسیونی با تابع پیوند لجیت برای میانگین پاسخ با عرض از مبدا و شیب تصادفی به صورت

$$logit(\mu_{ij}) = (\beta_۱ + b_{i۱}) + (\beta_۲ + b_{i۲})S_{ij} + (\beta_۳ + b_{i۳})W_{ij} + (\beta_۴ + b_{i۴})SF_{ij} + (\beta_۵ + b_{i۵})C_{ij},$$

می‌باشد.

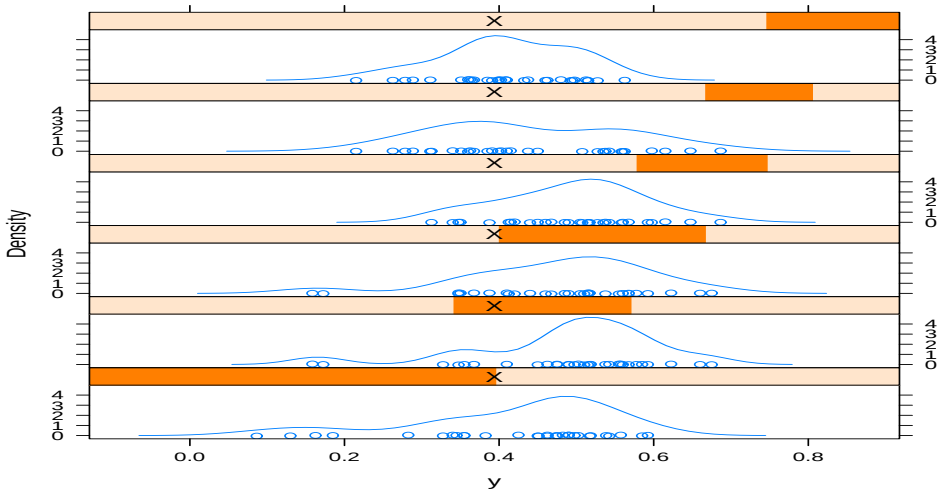
مدل دوم را برای هر یک از مدل‌های رگرسیونی پارامتر دقت برازش دادیم و مقادیر DIC را نیز برای هر یک به‌دست آوردیم. این مقادیر در جدول ۱۲.۴ گزارش شده‌اند. طبق این نتایج، مدل $e\cdot ۲$، مدلی با عرض از مبدا تصادفی و شیب ثابت، انتخاب می‌شود. نتایج برازش این مدل و آزمون‌های همگرایی گلمن-روبین و جی‌وک برای این مدل در جدول‌های ۱۳.۴ و ۱۴.۴ آمده‌اند. از آن‌جا که مقادیر $PSRF$ برای هر پارامتر نزدیک به یک و قدرمطلق Z_nها مقداری کوچک را نشان می‌دهند، می‌توان گفت زنجیر

جدول ۱۰.۴: نتایج برازش مدل با پارامتر دقت ثابت با توزیع پیشین $d \cdot ۳$ برای داده‌های مقاومت بتن

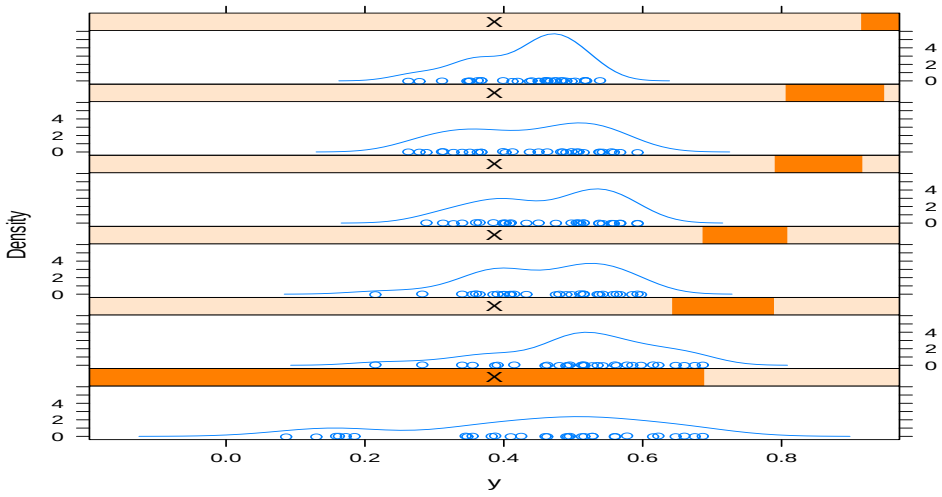
پارامترها	استنباط پسین		
	میانگین	انحراف معیار	فاصله اعتبار ٪۹۵
$\beta_۱$	$-۱۱/۴۹۸۱$	۰/۵۰۳۰	$(-۱۳/۳۷۸, -۸/۰۰۷)$
$\beta_۲$	۸/۷۰۹۹	۰/۰۰۰۲۳۶۷	(۷/۹۵۹, ۹/۵۲۲۱)
$\beta_۳$	$-۰/۴۰۱۳$	۰/۰۰۰۷۱۹۰	$(-۱/۶۳۷, ۰/۶۹۷۲)$
$\beta_۴$	۱۶/۱۲۰۵	۰/۰۰۰۶۸۵۶	(۱۴/۰۵۲, ۱۸/۱۷۸۱)
$\beta_۵$	$-۱۲/۷۱۴۳$	۰/۰۰۰۳۳۶۷	$(-۱۳/۷۲۸, -۱۱/۷۷)$
ϕ	۸۳/۵۷۶۵	۴/۸۲۰	(۵۲/۳۰۸, ۱۰۷/۱۶)
$\Sigma_{b_{۱۱}}$	۴۵/۶۴۵۱	۱/۳۵۲۳	(۴۱/۳۸۰۵, ۱۳۰/۹)
$\Sigma_{b_{۱۲}}$	۰/۳۲۲۱	۰/۳۴۴۵	$(-۳۹/۱۹۹۶, ۲۸/۶۰)$
$\Sigma_{b_{۱۳}}$	$-۰/۰۳۴۸$	۰/۱۳۱۱	$(-۴۱/۱۸۰۶, ۲۷/۹۴)$
$\Sigma_{b_{۱۴}}$	۰/۲۹۸۶	۰/۴۶۱	$(-۲۱/۸۰۴۶, ۴۰/۶۲)$
$\Sigma_{b_{۱۵}}$	۱۰/۱۷۰۳	۰/۷۶۰۹	$(-۲۰/۵۶۴۱, ۴۲/۵۷)$
$\Sigma_{b_{۲۲}}$	۵۲/۱۶۰۵	۱/۵۶۸	(۲۹/۶۵۲۳, ۹۹/۴۴)
$\Sigma_{b_{۲۳}}$	$-۰/۱۹۵۷$	۰/۳۹۰۵	$(-۲۶/۵۸۰۸, ۲۴/۷۶)$
$\Sigma_{b_{۲۴}}$	۰/۹۱۴۹	۰/۳۱۵۳	$(-۲۰/۸۰۵۹, ۲۸/۴۶)$
$\Sigma_{b_{۲۵}}$	۰/۰۷۸۲	۰/۵۸۱۲	$(-۲۸/۲۹۰۹, ۲۱/۵۹)$
$\Sigma_{b_{۳۳}}$	۵۱/۳۲۵۵	۱/۸۲۳۳	(۲۸/۰۶۷۸, ۹۶/۴۷)
$\Sigma_{b_{۳۴}}$	$-۰/۰۴۶۵$	۰/۲۴	$(-۲۶/۳۷۱۹, ۲۲/۱۶)$
$\Sigma_{b_{۳۵}}$	۰/۱۱۴	۰/۱۱۹۹	$(-۲۹/۸۶۷, ۲۳/۷۰)$
$\Sigma_{b_{۴۴}}$	۵۱/۷۷۸۸	۱/۱۴۰۴	(۲۸/۱۳۷۴, ۹۵/۱۶)
$\Sigma_{b_{۴۵}}$	$-۰/۲۱۴۷$	۰/۰۳۰۷	$(-۲۶/۲۶۴۴, ۲۶/۹۹)$
$\Sigma_{b_{۵۵}}$	۵۲/۳۷۷۵	۱/۰۱۴۴	(۲۸/۹۷۳, ۱۰۴/۶)

جدول ۱۱.۴: نتایج آزمون‌های همگرایی برای مدل با پارامتر دقت ثابت با توزیع پیشین $d \cdot ۳$ برای داده‌های مقاومت بتن

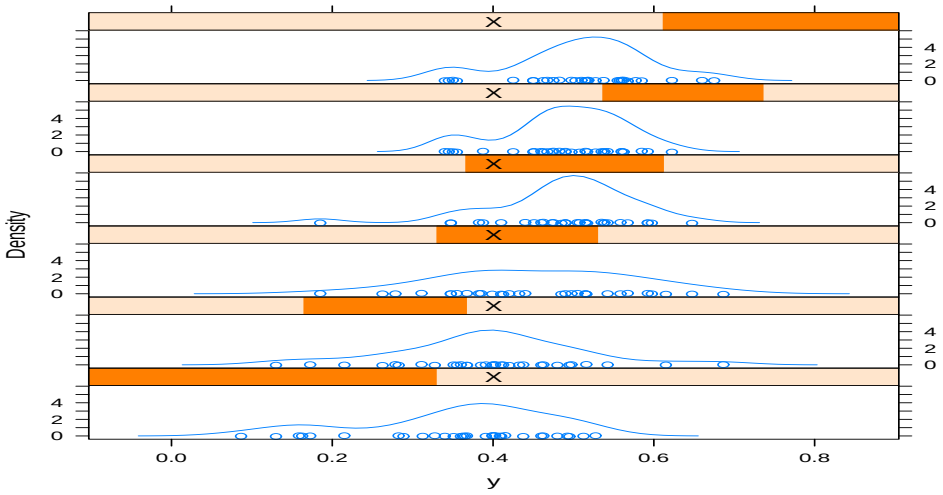
پارامترها	آماره آزمون		
	گلمن-روبین	جی‌وک	
	برآورد آماره آزمون	زنجیر۱	زنجیر۲
$\beta_۱$	۱۴۷/۲	۶/۲۰۳	−۳/۱۲۱
$\beta_۲$	۱/۰۰	۰/۱۴۰	−۱/۳۵۷
$\beta_۳$	۱/۰۰	۰/۴۱۸	۰/۲۰۳
$\beta_۴$	۱/۰۲	۲/۲۰۹	−۰/۳۹۹
$\beta_۵$	۱/۰۰	−۰/۲۳۷	−۱/۸۳۸
ϕ	۱/۰۰	۲/۱۵۳	−۰/۵۴۸
$\Sigma_{b_{۱۱}}$	۱/۰۱	۲/۰۲۴	۰/۲۲۱
$\Sigma_{b_{۱۲}}$	۱/۰۱	−۱/۰۸۸	−۲/۴۴
$\Sigma_{b_{۱۳}}$	۱/۰۲	۰/۳۳۷	−۲/۳۱۵
$\Sigma_{b_{۱۴}}$	۱/۰۰	۲/۱۱۴	۰/۸۷۴
$\Sigma_{b_{۱۵}}$	۱/۰۰	−۰/۲۱۰	۰/۹۰۴
$\Sigma_{b_{۲۲}}$	۱/۰۰	۰/۶۶۳	۱/۲۹۱
$\Sigma_{b_{۲۳}}$	۱/۰۲	۲/۴۱۲	۰/۷۷۸
$\Sigma_{b_{۲۴}}$	۱/۰۱	−۰/۰۵۵	−۱/۸۲۶
$\Sigma_{b_{۲۵}}$	۱/۰۰	−۰/۲۰۲	−۰/۶۷۳
$\Sigma_{b_{۳۳}}$	۱/۰۱	−۲/۱۱۳	−۰/۲۱۷
$\Sigma_{b_{۳۴}}$	۱/۰۱	۰/۳۷۶	−۰/۳۸۷
$\Sigma_{b_{۳۵}}$	۱/۰۰	۰/۱۹۷	−۲/۶۲۹
$\Sigma_{b_{۴۴}}$	۱/۰۰	۲/۸۰۱	۰/۲۱۲
$\Sigma_{b_{۴۵}}$	۱/۰۲	۱/۲۶۵	۱/۵۴۵
$\Sigma_{b_{۵۵}}$	۱/۰۰	−۱/۴۳۷	۱/۴۱۸



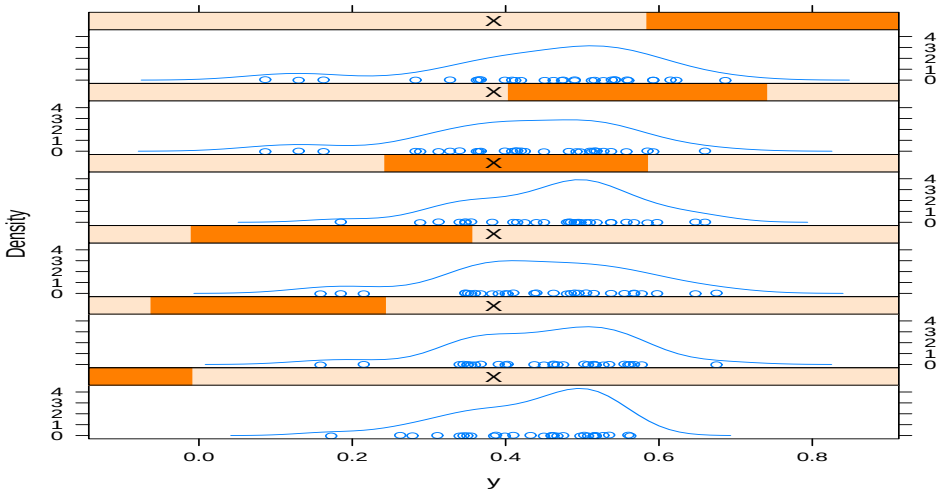
شکل ۱۱.۴: نمودار جعبه‌ای و چگالی متغیر پاسخ در زمان تکرار اول در دسته‌های مختلف متغیر ماسه



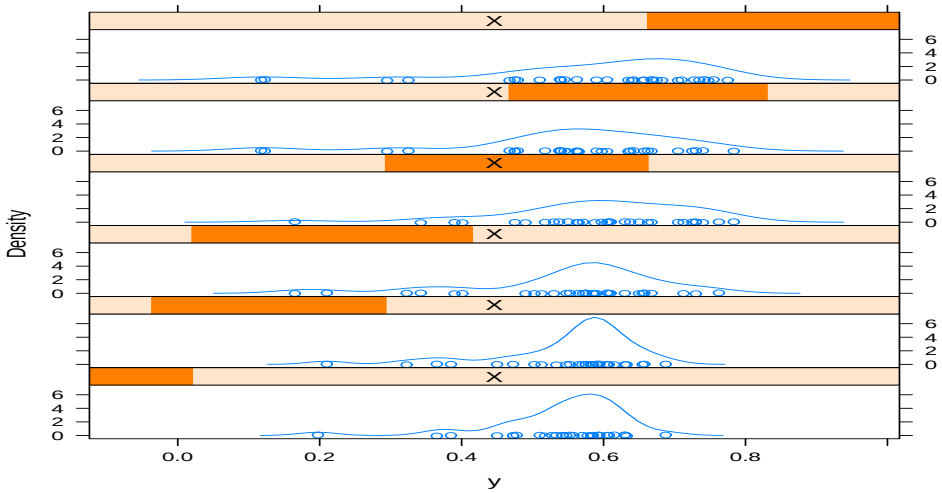
شکل ۱۲.۴: نمودار جعبه‌ای و چگالی متغیر پاسخ در زمان تکرار اول در دسته‌های مختلف متغیر آب



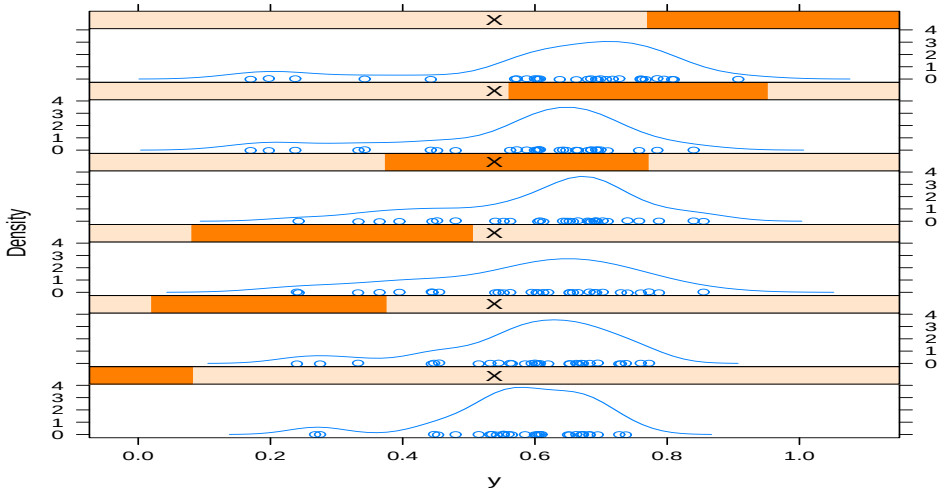
شکل ۱۳.۴: نمودار جعبه‌ای و چگالی متغیر پاسخ در زمان تکرار اول در دسته‌های مختلف متغیر سیمان



شکل ۱۴.۴: نمودار جعبه‌ای و چگالی متغیر پاسخ در زمان تکرار اول در دسته‌های مختلف متغیر میکروسیلیس



شکل ۱۵.۴: نمودار جعبه‌ای و چگالی متغیر پاسخ در زمان تکرار دوم در دسته‌های مختلف متغیر میکروسیلیس



شکل ۱۶.۴: نمودار جعبه‌ای و چگالی متغیر پاسخ در زمان تکرار سوم در دسته‌های مختلف متغیر میکروسیلیس

جدول ۱۲.۴: مدل‌های مختلف ϕ و مقادیر DIC برای مدل دوم در داده‌های مقاومت بتن

مدل	$\ln(\phi_{ij})$	DIC
$e\cdot ۱$	$\delta_۱ + \delta_۲ S_{ij} + \delta_۳ W_{ij} + \delta_۵ C_{ij}$	$-۱۷۳۳/۳$
$e\cdot ۲$	$(\delta_۱ + d_{i۱}) + \delta_۲ S_{ij} + \delta_۳ W_{ij} + \delta_۵ C_{ij}$	$-۱۷۳۳/۵$
$e\cdot ۳$	$\delta_۱ + (\delta_۲ + d_{i۲}) S_{ij} + (\delta_۳ + d_{i۳}) W_{ij} + (\delta_۵ + d_{i۵}) C_{ij}$	$-۱۷۲۵/۳$
$e\cdot ۴$	$(\delta_۱ + d_{i۱}) + (\delta_۲ + d_{i۲}) S_{ij} + (\delta_۳ + d_{i۳}) W_{ij} + (\delta_۵ + d_{i۵}) C_{ij}$	$-۱۷۲۰/۱$

تولیدشده برای هر پارامتر، جز $\beta_۱$، همگرا است.

با توجه به نتایج حاصل از هر دو مدل، می‌توان نتیجه گرفت که افزایش میزان ماسه و میکروسیلیس، به‌طور متوسط، بر افزایش میزان مقاومت بتن تاثیر مستقیم دارند در حالی‌که آب و سیمان، به‌طور متوسط، بر افزایش میزان مقاومت بتن تاثیر معکوس دارند. همچنین با مشاهده نتایج مدل $e\cdot ۲$ نیز می‌توان گفت میزان دقت مدل با افزایش میزان ماسه، آب و سیمان، با توجه به علامت مثبت این ضرایب، افزایش می‌یابد. البته باید توجه داشت که برآوردهای پارامترهای مدل رگرسیونی دقت مدل بر استنباط ساختار رگرسیونی میانگین مدل نیز تاثیرگذار است و اگر به‌درستی انتخاب شود، می‌تواند موجب افزایش کارایی استنباط ساختار میانگین شود.

۳.۴ نتیجه‌گیری

در این پایان‌نامه یک مدل رگرسیون آمیخته بتا برای پاسخ‌های وابسته و محدود به فاصله $(۰,۱)$ ارائه شد. از آن‌جا که این مدل در رده $GLMMs$ قرار دارد، استنباط آن از دیدگاه بیزی و استفاده از روش‌های نمونه‌گیری $MCMC$ صورت گرفت. عملکرد این مدل با یک مطالعه شبیه‌سازی مورد ارزیابی قرار گرفت و کاربرد مدل در دو مثال واقعی نشان داده شد. با توجه به نتایج به‌دست آمده از شبیه‌سازی و پیاده‌سازی مدل روی مثال‌های واقعی، به‌طور خلاصه، می‌توان نتایج زیر را عنوان کرد:

۱. استنباط بیزی مدل رگرسیون آمیخته بتا نسبت به انتخاب توزیع‌های پیشین استوار است.

۲. دقت استنباط‌های بیزی برای پارامترهای رگرسیونی بالا و قابل اعتماد است.

۳. اریبی برآوردهای پارامترهای وابستگی اثرات تصادفی می‌تواند قابل توجه باشد، ولی از آن‌جا که

این پارامترها معمولا مورد علاقه محققین نیستند، این اریبی می‌تواند چندان مهم نباشد. البته باید توجه داشت که در مواردی از جمله تحلیل داده‌های فضایی، برآورد کارای این پارامترها با اهمیت خواهد بود. زیرا این برآوردها هم بر کارایی برآوردگرهای پارامترهای رگرسیونی تاثیر مستقیم دارند و هم بر کارایی پیشگویی‌های فضایی.

۴. در موقعیت‌های کاربردی که تکیه‌گاه متغیر پاسخ محدود به یک فاصله است و ماهیت آن‌ها وابسته هستند، مدل پیشنهادشده در این پایان‌نامه و رهیافت استنباط مطرح‌شده می‌تواند به عنوان یک انتخاب محقق مد نظر قرار گیرد.

۵. با توجه به نتایج شبیه‌سازی‌ها و مثال‌های کاربردی، می‌توان همگرایی زنجیرهای پارامترهای مدل را قابل حصول دانست و در نتیجه به استنباط‌های به‌دست آمده اطمینان داشت.

۴.۴ پیشنهادات برای آینده تحقیق

با توجه به مباحث نظری و مطالعه شبیه‌سازی در این پایان‌نامه، می‌توان به بندهای زیر به عنوان آینده تحقیق اشاره کرد:

۱. در این پایان‌نامه، تمرکز بر تحلیل داده‌های طولی است. مدل و روش استنباط پیشنهادی را می‌توان برای سایر ساختار داده‌های وابسته مانند داده‌های فضایی و فضایی-زمانی مورد مطالعه قرار داد.

۲. استفاده از الگوریتم‌های $MCMC$ برای تحلیل بیزی $GLMMs$ می‌تواند بسیار هزینه‌بر باشد. در این موارد بهره‌گیری از روش‌های بیزی تقریبی مانند تقریب لاپلاس آشیانه‌ای جمع‌بسته[۷] ($INLA$) می‌تواند موجب کاهش این هزینه‌ها شود. مطالعه امکان به‌کارگیری این نوع روش‌ها از جمله زمینه‌های تحقیقات آینده می‌باشد.

[۷] Integrated Nested Laplace Appproximation

جدول ۱۳.۴: نتایج برازش مدل ۲ $\cdot e$ برای داده‌های مقاومت بتن

پارامترها	استنباط پسین		
	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
$\beta_۱$	−۱۰٫۶۷۶۷	۰٫۹۹۰۸	(−۱۲٫۵۳۵۶, −۸٫۸۸۸)
$\beta_۲$	۸٫۲۳۱۰۶	۰٫۴۵۷	(۷٫۳۰۳۱, ۹٫۰۶۹۸)
$\beta_۳$	−۱٫۱۹۸۶	۰٫۶۹۹۷	(−۲٫۵۵۱۷, ۰٫۱۷۶۱)
$\beta_۴$	۱۵٫۵۲۷۱	۱٫۰۳۷۱	(۱۳٫۵۲۲۲, ۱۷٫۵۷۳۳)
$\beta_۵$	−۱۱٫۲۸۶۹	۰٫۵۰۴۷	(−۱۳٫۲۴۳۴, −۱۰٫۲۳۶۹)
$\delta_۱$	۱٫۱۳۸	۱٫۸۸۳۹	(−۲٫۰۰۸۱, ۵٫۳۷۲۴)
$\delta_۲$	۱٫۵۶۰۱	۱٫۲۹۳۲	(−۰٫۸۲۴۶, ۴٫۰۹۵۷)
$\delta_۳$	۲٫۱۱۶۱	۲٫۰۸۸۸	(−۱٫۷۹۵۹, ۶٫۴۲۸۲)
$\delta_۵$	۳٫۵۶۲۹	۱٫۵۷۶۴	(۰٫۴۸۸۲, ۶٫۸۰۷۳)
$\Sigma_{b_{۱۱}}$	۴۰٫۵۲۶۱	۲٫۸۹۸۹	(۲۳٫۶۹۹۷, ۶۹٫۵۶۰۳)
$\Sigma_{b_{۱۲}}$	−۰٫۰۴۲۱	۰٫۵۳۱۲	(−۱۹٫۲۰۲۴, ۱۹٫۴۳۷۹)
$\Sigma_{b_{۱۳}}$	۰٫۰۸۵۱	۰٫۴۱۱۲	(−۱۹٫۱۷۲, ۱۹٫۰۴۰۲)
$\Sigma_{b_{۱۴}}$	۰٫۲۲۸۷	۰٫۶۶۵۲	(−۱۹٫۳۹۶۴, ۲۰٫۱۲۷۳)
$\Sigma_{b_{۱۵}}$	−۰٫۰۹۲۳	۰٫۳۷۲	(−۱۸٫۵۷۴۹, ۱۸٫۸۳۵۹)
$\Sigma_{b_{۲۲}}$	۵۲٫۰۸۲۱	۲٫۲۸۴۴	(۲۷٫۹۵۰۲, ۹۷٫۳۳۳۵)
$\Sigma_{b_{۲۳}}$	−۰٫۰۷۷۶	۰٫۰۸۴۸	(−۲۴٫۷۴۶۵, ۲۵٫۰۴۴)
$\Sigma_{b_{۲۴}}$	۰٫۸۲۹۷	۰٫۶۳۸۹	(−۲۳٫۵۵۹۶, ۲۷٫۶۷۷۶)
$\Sigma_{b_{۲۵}}$	−۰٫۴۲۱۸	۰٫۵۸۸۵	(−۲۵٫۲۴۳۸, ۲۴٫۹۴۳۴)
$\Sigma_{b_{۳۳}}$	۵۲٫۱۴۵۴	۲٫۶۳۱۲	(۲۸٫۴۴۶۷, ۹۷٫۳۴۷۱)
$\Sigma_{b_{۳۴}}$	−۰٫۴۲۴	۰٫۲۰۱۲	(−۲۵٫۸۱۲۲, ۲۴٫۱۸۷۶)
$\Sigma_{b_{۳۵}}$	−۰٫۱۸۶	۰٫۳۹۷۴	(−۲۴٫۵۱۸۲, ۲۵٫۱۰۴۵)
$\Sigma_{b_{۴۴}}$	۵۲٫۴۴۹۷	۲٫۳۸۲	(۲۸٫۳۹۸۲, ۹۷٫۷۶۷۲)
$\Sigma_{b_{۴۵}}$	−۰٫۳۶۳	۰٫۳۴۷	(−۲۶٫۸۷۳۵, ۲۳٫۷۰۸۹)
$\Sigma_{b_{۵۵}}$	۵۲٫۰۳۴۱	۲٫۰۸۶۵	(۲۸٫۲۵۴۴, ۹۴٫۲۳۶۸)

جدول ۱۴.۴: نتایج آزمون‌های همگرایی برای مدل ۲ $\cdot e$ برای داده‌های مقاومت بتن

پارامترها	آماره آزمون		
	گلمن-روبین	جی‌وک	
	برآورد آماره آزمون	زنجیر۱	زنجیر۲
β_1	۱۲۳٫۲	−۰٫۸۸۹	−۵٫۳۰۹
β_2	۱٫۰۰	۱٫۴۲۸	−۰٫۱۱۱
β_3	۱٫۰۰	−۱٫۰۲۱	−۰٫۴۱۸
β_4	۱٫۰۰	۱٫۶۴۳	−۱٫۷۵۶
β_5	۱٫۰۰	۱٫۱۸۴	۱٫۵۱۱
δ_1	۱٫۰۳	۱٫۷۱۱	−۳٫۱۸۴
δ_2	۱٫۰۲	−۰٫۰۶۲	−۲٫۵۸۸
δ_3	۱٫۰۰	۱٫۶۵۴	−۰٫۴۳۸
δ_5	۱٫۰۰	−۰٫۹۹۸	۰٫۶۴۴
$\Sigma_{b_{11}}$	۱٫۰۰	−۰٫۹۵۶	−۰٫۳۸۸
$\Sigma_{b_{12}}$	۱٫۰۰	۰٫۰۰۴	−۱٫۰۵۲
$\Sigma_{b_{13}}$	۱٫۰۰	۱٫۲۸۶	۱٫۶۲۷
$\Sigma_{b_{14}}$	۱٫۰۰	−۱٫۱۶۲	−۰٫۶۹۸
$\Sigma_{b_{15}}$	۱٫۰۲	۳٫۴۴۲	−۱٫۷۵۷
$\Sigma_{b_{22}}$	۱٫۰۰	۲٫۱۲۸	−۰٫۹۲۹
$\Sigma_{b_{23}}$	۱٫۰۰	۰٫۹۲۹	−۰٫۷۸۵
$\Sigma_{b_{24}}$	۱٫۰۰	۰٫۱۲۲	−۱٫۹۷۴
$\Sigma_{b_{25}}$	۱٫۰۰	−۰٫۷۸۳	۰٫۹۵۷
$\Sigma_{b_{33}}$	۱٫۰۰	۱٫۷۸۱	−۱٫۹۸۸
$\Sigma_{b_{34}}$	۱٫۰۳	−۱٫۰۰۸	−۰٫۷۷۶
$\Sigma_{b_{35}}$	۱٫۰۰	۱٫۷۵۹	۰٫۵۵۸
$\Sigma_{b_{44}}$	۱٫۰۲	۰٫۸۷۵	−۳٫۰۰۳
$\Sigma_{b_{45}}$	۱٫۰۱	−۱٫۸۰۳	۰٫۱۶۴
$\Sigma_{b_{55}}$	۱٫۰۱	۰٫۸۴۱	−۲٫۲۸۵

پیوست آ

نرم‌افزار $JAGS$

یکی از مهم‌ترین مسایلی که در کاربردهای گسترده مدل‌های بیزی مطرح می‌شود، یافتن توزیع‌های شرطی کامل است که اغلب نیازمند انتگرال‌گیری از توابعی با ابعاد بالاست و مستلزم محاسبات پیچیده‌ای نیز می‌باشد. در مجموعه روش‌های نمونه‌گیری $MCMC$، نمونه‌گیر گیبز برای تولید نمونه مونت کارلو از توزیع پسین توام نیازمند محاسبه توزیع‌های شرطی کامل می‌باشد. این توزیع‌ها به آسانی و به‌طور خودکار در نرم‌افزار $JAGS$ محاسبه می‌شوند. پس از محاسبه توزیع‌های شرطی کامل، تولید نمونه از آن‌ها انجام می‌شود. اگر این توزیع‌ها، از جمله توزیع‌های شناخته‌شده باشند، تولید نمونه به راحتی و به‌طور مستقیم صورت می‌پذیرد. در غیر این‌صورت تولید نمونه با روش‌های $MCMC$ مانند متروپولیس-هستینگس انجام می‌گیرد.

$JAGS$، به نقل از نویسنده برنامه، مارتین پلامر[1]، برنامه‌ای برای تحلیل مدل‌های بیزی با استفاده از نمونه‌گیری گیبز است. این برنامه دارای دستور زبانی مشابه، اما کمی متفاوت، با زبان‌های $WinBUGS$ و $OpenBUGS$ (نسخه‌های برنامه $BUGS$[2]) است. $JAGS$ یک جایگزین برای $WinBUGS$ است که می‌تواند بعد از نصب، از طریق R و با استفاده از بسته‌های $rjags$ و $r2jags$ مورد استفاده قرار گیرد.

اجرای یک برنامه در $JAGS$ شامل سه مرحله است: واردکردن داده‌ها، فرمول‌بندی مدل و نمونه‌گیری از زنجیرها. بعد از مشخص کردن داده‌ها و تعیین مدل، برای شروع شبیه‌سازی $MCMC$، ابتدا باید

---

[1] JAGS home page: http:// mcmc-jags.sourceforge.net

[2] Bayesian inference Using Gibbs Sampling

چندین مقدار آغازی را به مدل ارائه دهیم. سپس نوبت به تدوین مدل و راه‌اندازی زنجیرها برای تعدادی از تکرارها می‌رسد. یک خروجی $MCMC$ شامل دو بخش است. بخش اول شامل دوره داغیدن است و باقیمانده اجرا، خروجی است که برای همگرا شدن به توزیع مانا (توزیع پسین توام) مشاهده می‌شود. به منظور کنترل همگرایی و برای تولید نمونه‌هایی با وابستگی کمتر، به‌طور معمول به چندین زنجیر نیاز داریم. هر یک از زنجیرها شامل دنباله‌ای از نمونه‌های تولیدشده از توزیع پسین هستند. در نهایت، بعد از تولید نمونه، بررسی همگرایی مدل ضروری است.

با توجه به مراحل ذکر شده، دستورالعمل کلی اجرای یک برنامه در $JAGS$ به صورت زیر خواهد بود:

```
data <- list( y=y, N=length(y), ...)
inits <- list( list( var1, var2 ),
               list( var1, var2 ),
               list( var1, var2 ) )
cat( " model {
##Likelihood
...
##Priors
...
}", file="model.txt " )
```

برای تشریح این دستورالعمل، یک مدل رگرسیون خطی ساده را با $JAGS$ برازش می‌دهیم.

مثال آ.۱.۰. مدل رگرسیون خطی ساده $y_i = \alpha + x_i\beta + e_i, i = ۱, \ldots, ۵۰$ را با شرایط

$$\alpha, \beta \sim N(۰, ۰/۰۰۱), \quad x_i \sim N(۰, ۳۶), \quad e_i \sim N(۰, \sigma_e), \quad \sigma_e \sim U(۰, ۱۰۰)$$

در نظر بگیرید. شبیه‌سازی این مدل در $JAGS$ و بسته $rjags$ به صورت زیر قابل انجام است:

۱. فراخوانی بسته‌های مورد نیاز از جمله $rjags$:

```
library(rjags)
library(coda)
library(lattice)
```

۲. شبیه‌سازی مجموعه داده به حجم ۵۰ از مدل:

```
## simulate covariate data
n <- 50
sdx <- 6
obsx <- rnorm(n, 0, sdx)
## simulate response data
alpha <- 0
beta <- 10
sdy <- 20
error <- rnorm(n, 0, sdy)
obsy <- alpha + beta*truex + error
```

۳. معرفی داده‌ها:

```
jags_d <- list(x = obsx, y = obsy, n = length(obsx))
```

۴. معرفی مقادیر اولیه برای دو زنجیر:

```
jags_ini <- list(list(alpha=-2,beta=5,sdy=15),
                 list(alpha=1,beta=10,sdy=25) )
```

۵. شناساندن مدل بیزی رگرسیون خطی ساده به $JAGS$:

```
cat("
model{
## Likelihood
for (i in 1:n) {
mu[i] <- alpha + beta * x[i]
y[i] ~ dnorm(mu[i], tauy)  }
```

```
## Priors
alpha ~ dnorm(0, .001)
beta ~ dnorm(0, .001)
sdy ~ dunif(0, 100)
tauy <- 1 / (sdy * sdy)
}" ,  file="yerror.txt")
mod <- jags.model("yerror.txt", data=jags_d,
                  inits=jags_ini,
                  n.chains=2,
                  n.adapt=1000)
```

۶. تولید نمونه $MCMC$ به حجم ۱۰۰۰ از توزیع پسین توام پارامترها:

```
out <- coda.samples(mod,
                    var=c("alpha", "beta", "sdy"),
                    n.iter=10000,
                    thin=10)
```

۷. محاسبه کمیت‌های لازم از توزیع پسین توام:

```
summary(out)
```

جدول آ.۱، خلاصه‌ای از کمیت‌های پسین پارامترهای مدل را نمایش می‌دهد.

جدول آ.۱: نتایج شبیه‌سازی مثال (آ.۱.۰)

پارامترها	استنباط پسین			
	مقدار واقعی	میانگین	انحراف معیار	فاصله اعتبار ۹۵٪
α	۰	$-۵/۱۱۲$	۰/۱۰۸۰۸	$(-۱۶/۸۱۴, ۶/۷۸۶)$
β	۱۰	۶/۵۱۹	۰/۰۱۴۵۱	$(۵/۰۳۳, ۸/۰۲۹)$
σ_e	۲۰	۴۲/۱۰۹	۰/۰۷۹۵۳	$(۳۴/۴۰۶, ۵۱/۴۴۵)$

تمامی مدل‌های موجود در فصل‌های سوم و چهارم با استفاده از بسته $rjags$ در نرم افزار R برازش شده‌اند و دستورات مربوط به آن‌ها در پیوست ب گزارش شده‌اند.

پیوست ب

دستورهای نرم‌افزار $JAGS$

در این پیوست دستورات مربوط به برازش مدل‌های موجود در فصل چهارم ارایه شده است.
ابتدا بسته‌های زیر فراخوانی می‌شوند.

```
library(rjags)
library(coda)
library(lattice)
```

دستورات مربوط به داده‌های بنزین به صورت زیر می‌باشند:

```
data("GasolineYield",package="betareg")
```

- برای مدل اول با توزیع پیشین $b \cdot 1$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi,(1-mu[i])*phi)
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
```

```
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
phiinv ~ dgamma(0.01,0.01)
phi <- 1/phiinv
}
```

- برای مدل اول با توزیع پیشین ۲ $\cdot b$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi,(1-mu[i])*phi)
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
phiuni ~ dunif(0,50)
phi <- phiuni*phiuni
}
```

- برای مدل اول با توزیع پیشین ۳ $\cdot b$:

```
model {
```

```
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi,(1-mu[i])*phi)
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
phib ~ dbeta(1.1,1.1)
phi <- (50*phib)^2
}
```

- برای مدل اول با توزیع پیشین $b \cdot ۴$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi,(1-mu[i])*phi)
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
phib ~ dbeta(1.5,1.5)
```

```
phi <- (50*phib)^2
}
```

- برای مدل ۱ $\cdot c$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi[i],(1-mu[i])*phi[i])
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
log(phi[i]) <- delta[1]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
d[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
meanphi <- mean(phi)
}
```

- برای مدل ۲ $\cdot c$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi[i],(1-mu[i])*phi[i])
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
```

```
log(phi[i]) <- delta[1]+d[j,1,1]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
d[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
meanphi <- mean(phi)
}
```

- برای مدل ۳ ـ c:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi[i],(1-mu[i])*phi[i])
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
log(phi[i]) <- delta[2]*x[i]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
d[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
```

```
meanphi <- mean(phi)
}
```

- برای مدل ۴ $\cdot c$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi[i],(1-mu[i])*phi[i])
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
log(phi[i]) <- delta[1]+(delta[2]*x[i])
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
d[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
meanphi <- mean(phi)
}
```

- برای مدل ۵ $\cdot c$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi[i],(1-mu[i])*phi[i])
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
```

```
log(phi[i]) <- d[j,1,1]+delta[2]*x[i]
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
d[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
meanphi <- mean(phi)
}
```

- برای مدل $c\cdot۶$:

```
model {
for(j in 1:m) {
for(i in (N[j]+1) : N[j+1] ) {
y[i] ~ dbeta(mu[i]*phi[i],(1-mu[i])*phi[i])
logit(mu[i]) <- beta[1]+b[j,1,1] + (beta[2]+b[j,1,2])*x[i]
log(phi[i]) <- (delta[1]+d[j,1,1])+(delta[2]*x[i])
}
b[j,1,1:2] ~ dmnorm(mb[],sigb[,])
d[j,1,1:2] ~ dmnorm(mb[],sigb[,])
}
beta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:2] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:2,1:2] <- inverse(sigbeta[,])
sigb[1:2,1:2] ~ dwish(R0[,],4)
sigma.b[1:2,1:2] <- inverse(sigb[,])
```

```
meanphi <- mean(phi)
}
```

دستورات مربوط به داده‌های مقاومت بتن به صورت زیر می‌باشند:

- برای مدل اول با توزیع پیشین ۱ $\cdot d$:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
          (beta[3]+b[i,1,3])*W[i]+
          (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
phiinv ~ dgamma(0.01,0.01)
phi <- 1/(phiinv)
}
```

- برای مدل اول با توزیع پیشین ۲ $\cdot d$:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
```

```
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
          (beta[3]+b[i,1,3])*W[i]+
              (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
phiuni ~ dunif(0,50)
phi <- phiuni*phiuni
}
```

- برای مدل اول با توزیع پیشین $d \cdot ۳$:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]
                    +(beta[3]+b[i,1,3])*W[i]+
             (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
```

```
phib ~ dbeta(1.1,1.1)
phi <- (50*phib)^2
}
```

- برای مدل اول با توزیع پیشین ۴ ـ d:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
           (beta[3]+b[i,1,3])*W[i]+
              (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
phib ~ dbeta(1.5,1.5)
phi <- (50*phib)^2
}
```

- برای مدل ۱ ـ e:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
```

```
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
        (beta[3]+b[i,1,3])*W[i]+
             (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
log(phi[i,j]) <- delta[1]+(delta[2]*S[i])+(delta[3]*W[i])+
  (delta[5]*C[i])
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
d[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
meanphi <- mean(phi[,])
}
```

- برای مدل ۲ $\cdot e$:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
           (beta[3]+b[i,1,3])*W[i]+
              (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
log(phi[i,j])<-delta[1]+d[i,1,1]+(delta[2]*S[i])+
(delta[3]*W[i])+(delta[5]*C[i])
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
```

```
d[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
meanphi <- mean(phi[,])
}
```

- برای مدل ۳ ـ e:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
               (beta[3]+b[i,1,3])*W[i]+
                (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
log(phi[i,j])<- delta[1]+(delta[2]+d[i,1,2])*S[i]+
 (delta[3]+d[i,1,3])*W[i]+(delta[5]+d[i,1,5])*C[i]
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
d[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
```

```
meanphi <- mean(phi[,])
}
```

• برای مدل ۴ ۰ e:

```
model {
for(i in 1:m) {
for(j in 1:n) {
y[i,j] ~ dbeta(mu[i,j] * phi,(1-mu[i,j])* phi)
logit(mu[i,j])<-beta[1]+b[i,1,1]+(beta[2]+b[i,1,2])*S[i]+
                    (beta[3]+b[i,1,3])*W[i]+
                (beta[4]+b[i,1,4])*SF[i]+(beta[5]+b[i,1,5])*C[i]
log(phi[i,j])<-(delta[1]+d[i,1,1])+(delta[2]+d[i,1,2])*S[i]+
(delta[3]+d[i,1,3])
              *W[i]+(delta[5]+d[i,1,5])*C[i]
}
b[i,1,1:5] ~ dmnorm(mb[],sigb[,])
d[i,1,1:5] ~ dmnorm(mb[],sigb[,])
}
beta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
delta[1:5] ~ dmnorm(mbeta[],sigbeta1[,])
sigbeta1[1:5,1:5] <- inverse(sigbeta[,])
sigb[1:5,1:5] ~ dwish(R0[,],25)
sigma.b[1:5,1:5] <- inverse(sigb[,])
meanphi <- mean(phi[,])
}
```

مراجع

[۱] رضایی م و همکاران، (۱۳۹۱)، پایان‌نامه کارشناسی ارشد، تخمین خواص بنیادین بتن‌های hsc با استفاده از روش‌های آماری و شبکه‌های عصبی مصنوعی، دانشگاه صنعتی سهند.

[2] Birkhoff, G. D. (1942), What is the ergodic theorem?, *American Mathematical Monthly*, 49, 222-226.

[3] Branscum, A. J., Johnson, O. J., Thurmond, C. M. (2007), Bayesian beta regression: applications to household expenditure data and genetic distance between foot-and-mouth disease viruses, *Australian and New Zealand Journal of Statistics* , 49, 287–301.

[4] Breslow, N. E., Clayton, D. G. (1993), Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association*, 88, 9-25.

[5] Breslow, N. E., Lin, X. (1995), Bias correction in generalized linear mixed models with a single component of dispersion, *Biometrika*, 8, 81-91.

[6] Brooks, S. P., Gelman, A. (1998), General methods for monitoring convergence of iterative simulations, *Journal of Computational and Graphical Statistics*, 7, 434-455.

[7] Bury, K. (1999), *Statistical Distributions in Engineering*, Cambridge University Press, New York.

[8] Dey, D. K., Ghosh, S. K., Mallick, B. K. (2000), *Generalized Linear Models: A Bayesian Perspective, 1st Edition*, Marcel Dekker, New York.

[9] Ferrari, S. L. P., Cribari-Neto, F. (2004), Beta regression for modelling rates and proportions, *Journal of Applied Statistics*, 31, 799-815.

[10] Figueroa, J. I., Arellano-Valle, R. B., Ferrari, S. L. P. (2013), Mixed beta regression: A Bayesian perspective, *Computational Statistics and Data Analysis*, 61, 137-147.

[11] Fong, Y., Rue, H., Wakefield, J. (2010), Bayesian inference for generalized linear mixed models, *Biostatistics*, 11, 397-412.

[12] Gelfand, A., Smith, A. (1990), Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association*, 85, 398-409.

[13] Gelman, A. (2006), Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis*, 1, 515–533.

[14] Gelman, A., Kass, R. E., Carlin, B. P., Neal, R. (1998), Markov chain Monte Carlo in practice: a roundtable discussion, *The American Statistician*, 52, 93-100

[15] Gelman, A., Rubin, D. B. (1992), Inference from iterative simulation using multiple sequences, *Statistical Science*, 7, 457-511.

[16] Geman, S., Geman, D. (1984), Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6, 721-741.

[17] Geweke, J. (1992), Evaluating the accuracy of sampling-based approaches to calculating posterior moments, *In Bayesian Statistics*, (eds JM Bernado, JO Berger, AP Dawid and AFM Smith), 4, 169–193.

[18] Hastings, W. (1970), Monte Carlo sampling methods using Markov chains and their application, *Biometrika*, 57, 97-109.

[19] Heidelberger, P., Welch, P. D., (1983), Simulation run length control in the presence of an initial transient, *Operations Research*, 31, 1109-1144.

[20] James, A. T. (1964), Distributions of matrix variates and latent roots derived from normal samples, *Annals of Mathematical Statistics*, 35, 475-501.

[21] Johnson, N. L., Kotz, S., Balakrishnan, N. (1995), *Continuous Univariate Distributions, 2nd Edition*, Wiley, New York.

[22] Kotz, S., Nadarojah, S. (2004), *Multivariate t Distributions and Their Applications*, Cambridge University Press, New York.

[23] McCulloch, C. E., Searle, S. R. (2001), *Generalized, Linear, and Mixed Models*, John Wiley & Sons, New York.

[24] McCullach, P., Nelder, J. A. (1989), *Generalized Linear Models, 2nd Edition*, Chapman and Hall, London.

[25] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E. (1953), Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, 21, 1087-1092.

[26] Nelder, J. A., Wedderburn, R. W. M. (1972), Generalized linear models, *Journal of the Royal Statistical Society, Series A*, 135, 3, 370-384.

[27] Pinheiro, J. C., Bates, D. M. (1995), Approximations to the log-likelihood function in the nonlinear mixed-effects models, *Jornal of Computational and Graphical Statistics*, 4, 12-35.

[28] Plummer, M. (2003), JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Vienna, Austria.

[29] Plummer, M., Best, N., Cowles, K., Vines, K. (2006), The coda package, *R Project*, http://cran.r-project.org/doc/packages/coda.pdf.

[30] Prater, N. H. (1956), Estimate gasoline yields from crudes, *Petroleum Refiner*, 35, 236-238.

[31] Robert, C. P., Casella, G. (2004), *Monte Carlo Statistical Methods, 2nd Edition*, Springer, New York.

[32] Sarkar, D. (2008). The lattice package, *R Project*, http://r-forge.r-project.org/projects/lattice/.

[33] Smithson, M., Verkuilen, J. (2006), A better lemon squeezer? maximum likelihood regression with beta-distributed dependent variables, *Psychological Methods*, 11, 54-71.

[34] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., Van Der Lind, A. (2002), Bayesian measures of model complexity and fit, *Jornal of the Royal Statistical Society, Series B*, 64, 583-639.

[35] Venables, W. N. (2000), Exegeses on linear models, available at http://www.stats.ox.ac.uk/pub/MASS3/Exegeses.pdf.

نمایه

نمودار
  اثر، ۳۵
  جعبه‌ای، ۶۸
  جی‌وک، ۳۴
  خودهمبستگی، ۳۶
  میانگین تجمعی، ۳۵
  پراکنش، ۵۳
  چگالی، ۵۷
هسته مارکوف، ۸

Surname: Ajam Name: Mahnaz

Title: Bayesian Analysis of Beta Regression Models with Dependent Responses

Supervisor: Dr. Hossein Baghishani

Degree: Master of Science Subject: Statistics
Field: Bayesian Statistics

Shahrood University Faculty of Mathematical Sciences
Date: September 2014 Number of pages: 105

Keywords: Bayesian analysis; Beta regression; Mixed models

Abstract

In most applications including analysis of rate responses and proportions like inflation rate, growth rate, punctual customeres ratio in a financial business and a disease percent in a special region, we are interested in detecting effective elements on response variable. This aim is usually obtainable by regression models. The scale of such responses is continuous and restricted to the interval (0,1). Logistic and probit regression models are the popular models to analyze this type of responses. However, the real distribution of such responses are usually greatly skewed and then the aformentioned models are not appropriate for analyzing them. A new proposed regression model that has a logical and appropriate adaptation with the essence of rate responses is beta regression model which has recently become popular due to high flexibility of beta distribution in modeling various kinds of skewnesses. The beta regression model is an element of generalized linear models. An essential assumption in generalized linear models is independency of response variables. However, in various situations, e.g. in analyzing longitudinal data, spatial data and time series, the response variables are dependent. In these cases, generalized linear mixed models have been used in which the dependency structure of responses is considered by using latent variables. Fitting these models is easier by the Bayesian approach. This simplicity is due to the revolution of MCMC sampling methods. Therefore, our main aim is Bayesian analysis of dependent responses with continuous scale and restricted support.

Shahrood University
Faculty of Mathematical Sciences

Dissertation Submitted in Partial Fulfillment of The Requirements For The Degree of Master of Science in Statistics

Bayesian Analysis of Beta Regression Models with Dependent Responses

Supervisor

Dr. Hossein Baghishani

by

Mahnaz Ajam

September 2014