

پایش نیمرخ بیزی در حضور متغیر پنهان: مسئله پرتاب نیزک

علی‌رضا ظاهریون^۱، غزل آزادی^۲

تاریخ دریافت: ۱۴۰۰/۰۱/۱۷

تاریخ پذیرش: ۱۴۰۰/۰۸/۲۲

چکیده:

به‌طور معمول، پایش نیمرخ از طریق نمودارهای کنترل صورت می‌گیرد و در اغلب آن‌ها، متغیر پاسخ، قابل مشاهده است. در این مقاله، با مسئله مشابهی مواجهیم که در آن به‌جای مشاهده بردار پاسخ، مقادیر تابع پاداش را مشاهده می‌کنیم که برای تقریب به ذهن از مدل پرتاب نیزک استفاده کرده‌ایم. با فرض وجود حداکثر یک نقطه تغییر، دنباله‌ای مستقل از امتیازهای حاصل از پرتاب، مشاهده می‌شود و برآورد پارامتر دقت پرتاب‌ها و نقطه تغییر (در صورت وجود)، با دو رویکرد فراوانی‌گرا و بیزی ارائه می‌شوند. در هر دو رویکرد، دو حالت ممکن پارامتر اسکالر دقت و ماتریس دقت، به تفکیک بررسی شده‌اند. نتایج ارائه‌شده از طریق یک مطالعه عددی بررسی شده‌اند و این روش‌ها روی داده‌های واقعی حاصل از پرتاب، پیاده شده‌اند.

واژه‌های کلیدی: الگوریتم EM، الگوریتم MCMC، توزیع آمیخته، متغیر پنهان، نقطه تغییر، نمونه‌گیر گیبز.

۱ مقدمه

ذرات معلق با قطر کمتر از ۲٫۵ میکرومتر دقت کنید. عدد این شاخص، محصول غلظت چنین ذراتی، صرف‌نظر از نوع و جرم ذره است حال آنکه این غلظت توسط طیف گسترده‌ای از عناصر تولید می‌شود و ما هیچ اطلاعی از میزان مشارکت هریک از این عناصر در هوای امروز نداریم. آنچه به‌عنوان مشاهده در اختیار ما قرار می‌گیرد میانگین وزنی اجرام اندازه‌گیری شده توسط دستگاه طیف‌سنج است که اطلاعات مربوط به بسیاری دیگر از عناصر به‌طور موزج تشخیص داده‌شده توسط این دستگاه، در این شاخص گزارش نمی‌شود. فروگاهی اطلاعات چندمتغیره طیف‌سنج به مشاهدات یک‌متغیره، هرچند خلاصه‌سازی مفیدی در گزارش شاخص‌های کیفی هوا ایجاد می‌کند، اما اطلاعات مفیدی را از دست می‌دهد که این اطلاعات می‌توانند در پایش نیمرخ و در برآورد نقطه تغییر بسیار مفید واقع شوند. مثال پرتاب نیزک^۳، بهترین مثال برای بررسی این موضوع است. اصولاً در داده‌های امتیازهای پرتاب نیزک به سمت تخته آنچه ثبت می‌شود امتیاز حاصل از پرتاب است نه محل اصابت نیزک؛ حال آنکه عملکرد یک پرتاب‌کننده حرفه‌ای از محل فرود آمدن پرتاب او روی تخته بررسی می‌شود. از آنجایی‌که قرار است اصل موضوع روی همین مثال توضیح داده شود، شکل امتیازدهی

پایش نیمرخ^۴ یک فرایند یک متغیره در فاز II که یک مرحله آینده‌نگر است، به‌طور گسترده‌ای مورد بررسی قرار گرفته است [۲]. نقش روش‌های مبتنی بر تابع درست‌مایی در بسیاری از نمودارهای کنترل بسیار برجسته است تا آنجا که نمودارهای T^2 -هتلینگ، LRT و EWMA به‌طور مستقیم بر اساس تابع درست‌مایی حاصل می‌شوند. در نمودارهای کنترل مقدماتی، تغییرات نقطه‌به‌نقطه میانگین موضعی و واریانس موضعی در پنجره‌ای به طول ثابت و معمولاً متناسب با طول دنباله مشاهدات، امکان وجود نقطه تغییر در فرایند، بررسی و حدود آن تخمین زده می‌شود اما حتی در مشاهداتی از توزیع‌های شناخته‌شده نیز بسیار بعید است که روش‌های تشخیص نقطه تغییر نسبت به یک مشاهده حساس باشند و بتوانند به‌جای تشخیص نقطه تغییر در طول یک پنجره، نقطه دقیق را به‌عنوان پارامتر نقطه تغییر برآورد و ارائه کنند. مسئله زمانی بغرنج‌تر می‌شود که مشاهدات در دست، محصول متغیرهایی باشند که در عمل مشاهده نمی‌شوند و ما تنها برآیندی از اثرهای متقابل آن‌ها را مشاهده می‌کنیم. برای تقریب به ذهن، به مثال شاخص آلودگی

^۱دانشگاه شهید بهشتی، دانشکده علوم ریاضی، گروه آمار (نویسنده مسئول: a_taheriyoun@sbu.ac.ir)

^۲دانشگاه نروژی علوم طبیعی، پردیس اوس

^۳profile monitoring

^۴dart

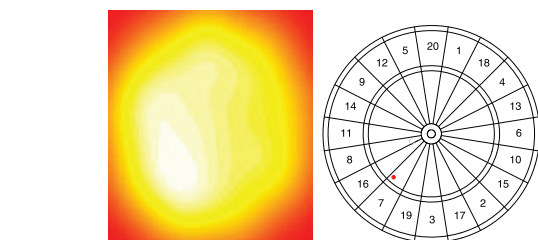
به پرتاب‌ها را کمی بیشتر مورد کنکاش قرار می‌دهیم. بازی نیزک،

$$P(\mathbf{Z} \in S) = \frac{\text{مساحت}(S \cap B)}{\text{مساحت}(B)}. \quad (1)$$

از طریق پرتاب نیزک به صفحه‌ای دایره‌ای شکل (تخته نیزک) انجام می‌شود. برخورد نیزک به قسمت‌های مختلف تخته نیزک امتیازهای متفاوتی را برای بازیکن به همراه دارد که البته رابطه یک‌به‌یکی بین محل اصابت نیزک و امتیاز کسب‌شده وجود ندارد یعنی ممکن است نشستن نیزک در دو نقطه مجزای تخته، منجر به امتیازهای یکسانی شود. تخصیص امتیاز و قوانین مربوط به این بازی در هر کشوری می‌تواند متفاوت باشد اما بر اساس قوانین موجود در سازمان نیزک بریتانیا، برخورد نیزک به مرکز تخته، DB، ۵۰ امتیاز، و حلقه مجاور، SB، ۲۵ امتیاز را به همراه دارد. سایر قسمت‌های تخته نیزک به ۲۰ قسمت تقسیم شده است که هر یک حاوی امتیازی از ۱ تا ۲۰ است. دو حلقه دو برابر و سه برابر نیز در این قسمت‌ها وجود دارد که در صورت برخورد نیزک به هر یک از این حلقه‌ها، امتیاز بازیکن به ترتیب دو و سه برابر خواهد شد.

$$\mathbf{Z} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}),$$

اگر $s(\mathbf{Z})$ امتیاز حاصل از نقطه محل برخورد باشد آنگاه تحت این مدل یکنواخت $E[s(\mathbf{Z})] \approx 12.82$. البته عاقلانه نیست که پرتاب‌کننده به تصادف نیزک را به سمت تخته رها کند و امید ریاضی امتیازها نیز مؤید این موضوع است. از این رو در [۳]، فرض را بر آن گذاشتند که پرتاب‌کننده مرکز تخته را نشانه می‌گیرد و مدل کروی زیر را در نظر گرفتند



شکل ۲: نمودار حرارت با فرض $\sigma = 42.09$. نقطه قرمز، نشانگر $\boldsymbol{\mu}$ است.

که $\mathbf{I}_{2 \times 2}$ ماتریس همانی است. نقطه $\boldsymbol{\mu} = (\mu_x, \mu_y)'$ نشان‌دهنده نقطه‌ای است که بازیکن به‌عنوان هدف خود انتخاب می‌کند. یک نمودار حرارت از چنین مدلی در شکل ۲ ارائه شده است.

به علاوه اندازه مقدار خطا توسط σ کنترل می‌شود. در واقع می‌توان گفت هر چه مقدار σ کوچک‌تر باشد، بازیکن پرتاب‌های خود را دقیق‌تر انجام داده است. لذا امید ریاضی امتیاز پرتاب‌ها تحت این مدل عبارت است از

$$\mathbb{E}_{\boldsymbol{\mu}, \sigma^2}[s(\mathbf{Z})] = \int_{\mathbb{R}^2} \frac{1}{2\pi\sigma^2} \exp\left[-\frac{\|(x, y) - \boldsymbol{\mu}\|^2}{2\sigma^2}\right] s(x, y) dx dy. \quad (2)$$

نگرش [۳] بر آن بود که اگر $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ داده‌های مشاهده نشده ما باشند که در آن $\mathbf{Z}_i = (Z_{i,x}, Z_{i,y})'$ مختصات محل اصابت نیزک است. در این صورت لگاریتم درستنمایی این داده‌های مشاهده نشده عبارت است از

$$\begin{aligned} \ell(\sigma^2; \mathbf{X}, \mathbf{Z}) &= \log L(\sigma^2; \mathbf{X}, \mathbf{Z}) \\ &= -n \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n [Z_{i,x}^2 + \mu_x^2 - 2Z_{i,x}\mu_x \\ &\quad + Z_{i,y}^2 + \mu_y^2 - 2Z_{i,y}\mu_y], \end{aligned} \quad (3)$$

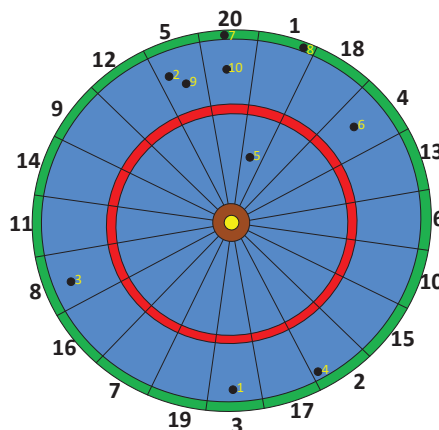
شکل ۱: تخته نیزک: ■، منطقه دو برابر؛ فاصله مرکز تخته تا نوار داخلی این منطقه ۱۶۲mm است. ■، منطقه سه برابر؛ فاصله مرکز تخته تا نوار داخلی این منطقه ۹۹mm است. پهنای هر دو منطقه دو برابر و سه برابر ۲۰mm است. ■، منطقه یک برابر؛ منطقه SB، با شعاع ۱۵۸mm و ■، منطقه DB که دایره‌ای به شعاع ۶۳۵mm است.

می‌توان در شکل ۱ وضعیت تخصیص امتیازها و فاصله مرکز تا هر ناحیه را بر اساس استاندارد سازمان نیزک بریتانیا دید. مدل‌بندی داده‌هایی با این نوع مشاهدات، پیش‌از این در [۳] ارائه شده است. اگر نیزک‌ها به تصادف به سمت تخته پرتاب شوند با فرض آنکه در تخته فرود آید، یک مدل احتمالی از محل برخورد نیزک روی تخته می‌تواند برآمده از یک فضای احتمال یکنواخت باشد و در این صورت اگر B نشان‌دهنده تخته نیزک و جفت مرتب $\mathbf{Z} = (Z_x, Z_y)'$ مختصات محل اصابت نیزک بر اساس دستگاه دکارتی دو بُعدی‌ای باشد که مرکز آن روی مرکز تخته منطبق است، آنگاه به ازای هر

به دست آوردن امتیاز موردنظر باید دقت پرتاب‌ها را به هر شکل ممکن بالا برد و به بیانی دیگر واریانس پرتاب را کاهش داد. پس، به‌وضوح می‌توان دید تأثیر ناهمگونی واریانس در یک بازی که توسط فرد خاصی انجام می‌شود تا چه حد بالاست. چون با تشخیص شماره پرتابی که از آن به بعد پرتاب‌ها با دقت متفاوتی انجام شده‌اند، می‌توان تا حد زیادی به بازیکن کمک کرد تا نخست متوجه شود که از آن لحظه بعد توانسته است دقیق‌تر پرتاب کند یا خیر و دوم اینکه در کدامین پرتاب شرایط روند بازی او تغییر چشمگیری داشته است. اگر متوجه شود از پرتاب r ام به بعد توانسته است بهتر بازی کند یا در اصطلاح بازی را یاد بگیرد، در دفعات بعدی سعی خواهد کرد زودتر از r امین پرتاب، دقت پرتاب‌های خود را بالا ببرد.

۱۰۲ تحلیل نشانه‌گیری بدون در نظر گرفتن دقت بازیکن

هنگام پرتاب نیزک دو مسئله مهم پیش می‌آید: یکی نشانه‌گیری و دیگری دقت. در این بخش ما تمام توجه خود را به مسئله نشانه‌گیری معطوف می‌کنیم و سپس در قسمت بعد لزوم در نظر گرفتن دقت را شرح می‌دهیم. سؤال اساسی آن است که اگر پرتاب‌کننده مجبور به انتخاب باشد، کدامیک را برمی‌گزیند: پرتاب با دقت کم ولی نشانه‌گیری مناسب یا پرتاب با دقت بالا ولی نشانه‌گیری‌های ضعیف؟ برای پاسخ به این سؤال در ابتدا تنها بحث نشانه‌گیری را مورد مذاقه قرار می‌دهیم و پس از آن خواهیم دید که در این خصوص، انتخاب دیگری وجود ندارد.



که در آن Z ماتریس حاصل از زیر هم قرار دادن مقادیر Z_1', \dots, Z_n' است. آنچه از نتیجه پرتاب نیزک ثبت می‌شود، امتیاز پرتاب است و نه مختصات مکان اصابت. این مسئله را نمی‌توان و نباید تنها در چارچوب مسئله پرتاب نیزک دید و یک مثال مشهود، محصول جایگزینی تخته نیزک با محل اصابت تسلیحات و تبدیل تابع امتیاز به تابع زیان است و این بازی از هر دو طرف (پرتاب‌کننده و دریافت‌کننده) اهمیت می‌یابد. موضوعی که در این مقاله به آن می‌پردازیم، بررسی پایش نیمرخ دقت اصابت‌ها در طول زمان و به‌ویژه بررسی تغییر در دقت پرتاب‌ها از طریق پایش پارامتر σ^2 است. در خصوص مسئله پرتاب نیزک، تغییر در دقت می‌تواند در اثر خستگی پرتاب‌کننده در دوره طولانی پرتاب و در مسئله شلیک، اثرات دمایی و یا استهلاک باشد. به‌طور معمول و معقول نقطه هدف یا μ در طول زمان ثابت در نظر گرفته می‌شود. از این‌رو، در بخش ۲ مدلی با واریانس متغیر معرفی می‌کنیم و الگوریتم EM را برای مسئله برآورد پارامتر واریانس و نقطه تغییر مدل، توسعه می‌دهیم. در همین بخش، موضوع برآورد را به ماتریس دقت یا کوواریانس تعمیم می‌دهیم. بخش ۳ برآورد پارامترهای مدل معرفی شده را به روش بیزی بررسی می‌کند. در این بخش، برآورد پارامترها تحت دو حالت بدون نقطه تغییر و حالت وجود یک نقطه تغییر به روش بیزی ارائه شده است و الگوریتم‌های شبیه‌سازی از توزیع‌های پسین به‌طور موجز بیان شده است. در بخش ۴ عملکرد برآوردگرهای مطرح شده در یک مطالعه شبیه‌سازی شده بررسی شده است. همچنین از یک مجموعه داده‌های واقعی امتیاز پرتاب نیزک برای به‌کارگیری روش، استفاده شده است.

۲ پرتاب با واریانس متغیر

با بررسی پرتاب‌های هر بازیکن در بازی نیزک، می‌توان به‌آسانی مشاهده کرد که فرد، پس از مدتی، بازی را به‌طور متفاوتی ادامه می‌دهد. متفاوت از این منظر که می‌تواند یا بهتر بازی کند و یا حتی بدتر. بهتر بازی کردن در رقابت نیزک بدین معناست که توانسته است امتیاز موردنظر را کسب نماید. امتیاز موردنظر همواره امتیاز بالا نیست. در بسیاری از بازی‌های نیزک باید پس از تعداد مشخصی پرتاب، امتیاز خاصی را به دست آورد؛ بنابراین چنانچه امتیاز شما بالاتر و یا پایین‌تر از حد مشخص شده باشد، عملاً بازی را باخته‌اید. حال اگر موارد ذکر شده را بخوانیم به زبان آماری بیان کنیم، ناگزیر می‌بایست از واریانس استفاده کنیم؛ چراکه برای

آنگاه می‌توان ادعا کرد که پرتاب‌کننده در جهت یادگیری بازی عمل کرده است و اگر واریانس افزایش یافته باشد، مسلماً دقت بازیکن رو به کاهش بوده است.

۲.۲ برآورد دقت با فرض مختصات مستقل و ناهمگونی واریانس

ابتدا حالتی را تصور کنید که مختصات افقی و عمودی محل اصابت مستقل از هم باشند. با توجه به توجیه استفاده از توزیع گاوسی، مدل زیر را می‌نویسیم

$$\mathbf{Z}_i = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, n, \quad (4)$$

که در آن

$$\boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_\tau \stackrel{iid}{\sim} N_{\mathbb{R}^2}(\mathbf{0}, \sigma_{\tau}^2 \mathbf{I}_{2 \times 2}) \quad \text{و} \quad \boldsymbol{\varepsilon}_{\tau+1}, \dots, \boldsymbol{\varepsilon}_n \stackrel{iid}{\sim} N_{\mathbb{R}^2}(\mathbf{0}, \sigma_{\tau+1}^2 \mathbf{I}_{2 \times 2}),$$

و برای برآوردپذیری پارامترها ضروری است که $\tau \in \{3, \dots, n-2\}$. اگر نقطه هدف‌گیری یعنی $\boldsymbol{\mu}$ را معلوم فرض کنیم، که البته منطقی هم همین است، لگاریتم تابع درست‌نمایی داده‌های کامل (مشاهده‌شده و مشاهده نشده) پارامترها تحت چنین مدلی به صورت زیر خواهد بود

$$\begin{aligned} \ell(\boldsymbol{\mu}, \sigma_{\tau}^2, \sigma_{\tau+1}^2, \tau | \mathbf{X}, Z) &= \log L(\boldsymbol{\mu}, \sigma_{\tau}^2, \sigma_{\tau+1}^2, \tau | \mathbf{X}, Z) \\ &= -\tau \log(\pi \sigma_{\tau}^2) - \frac{1}{2\sigma_{\tau}^2} \sum_{i=1}^{\tau} (Z_{i,x}^2 + Z_{i,y}^2) \\ &\quad - (n - \tau) \log(\pi \sigma_{\tau+1}^2) - \frac{1}{2\sigma_{\tau+1}^2} \sum_{i=\tau+1}^n (Z_{i,x}^2 + Z_{i,y}^2). \end{aligned} \quad (5)$$

به پیروی از رهیافت [۳]، برای تلفیقی از داده‌های مشاهده‌شده و مشاهده نشده، از الگوریتم EM برای ماکسیم کردن تابع لگاریتم درست‌نمایی استفاده می‌کنیم. به همین منظور، در تکرار t ام الگوریتم و در گام گرفتن امید ریاضی، با داده‌های مشاهده نشده مانند پارامتر برخورد می‌کنیم و امید ریاضی لگاریتم درست‌نمایی را به شرط داده‌های مشاهده‌شده امتیازهای $\mathbf{X} = (X_1, \dots, X_n)'$ محاسبه می‌کنیم. اگر عملگر \mathbb{E}_t را عملگر امید ریاضی نسبت به توزیعی با مقادیر پارامتری $\sigma_{\tau}^2 = \widehat{\sigma_{\tau}^2}$ ، $\sigma_{\tau+1}^2 = \widehat{\sigma_{\tau+1}^2}$ یعنی آخرین به‌روزرسانی‌های برآوردها در الگوریتم در نظر بگیریم، در این صورت

$$\begin{aligned} &\mathbb{E}_{t-1}[\ell(\sigma_{\tau}^2, \sigma_{\tau+1}^2, \tau; \mathbf{X}, Z | \mathbf{X})] \\ &= -\tau \log(\pi \sigma_{\tau}^2) - \frac{1}{2\sigma_{\tau}^2} \mathbb{E}_{t-1} \left[\sum_{i=1}^{\tau} (Z_{i,x}^2 + Z_{i,y}^2) | \mathbf{X} \right] \\ &\quad - (n - \tau) \log(\pi \sigma_{\tau+1}^2) - \frac{1}{2\sigma_{\tau+1}^2} \mathbb{E}_{t-1} \left[\sum_{i=\tau+1}^n (Z_{i,x}^2 + Z_{i,y}^2) | \mathbf{X} \right]. \end{aligned}$$

شکل ۳: مثال پرتاب نیزک؛ محل نشستن نیزک با دایره‌های مشکی توپُر و ترتیب هر پرتاب با اعداد زرد رنگ در کنار محل اصابت مشخص شده است.

اگر توزیع محل فرود نیزکها در صفحه، یک توزیع یکنواخت روی دایره‌ای به مرکز $\boldsymbol{\mu}$ باشد، می‌توان وضعیت نشانه‌گیری بازیکن را با اندازه‌گیری مقدار شعاع ساخته‌شده پرتاب‌ها سنجید. به‌عنوان مثال فرض کنید ۱۰ نیزک به‌صورتی که در شکل ۳ نشان داده‌شده است، پرتاب‌شده‌اند. به‌طور واضح بردار امتیازهای کسب‌شده در این پرتاب‌ها به‌صورت $\mathbf{X} = (3, 5, 8, 2, 1, 4, 4, 2, 5, 2)$ است. همان‌گونه که مشاهده می‌کنید بدون نیاز به محاسباتی بیشتر در صورتی که هدف، تنها کمتر کردن شعاع پرتاب‌ها باشد، از پرتاب پنجم به بعد، بازیکن به نقطه‌ای از یادگیری بازی رسیده است که می‌تواند به‌صورتی که در شکل نشان داده‌شده است، شعاع توزیع یکنواخت پرتاب‌ها را کاهش دهد و با نشانه‌گیری قسمت فوقانی تخته نیزک، پرتاب‌های قابل‌قبولی را به نمایش بگذارد. اگر τ را شماره پرتابی در نظر بگیریم که از آن به بعد نشانه‌گیری بهبود پیدا کرده است، آنگاه در این مثال $\tau = 5$ و بنابراین چنانچه بازیکن از این موضوع مطلع شود می‌تواند در مراحل بعدی بازی این عدد را کاهش دهد و عادت نشانه‌گیری درست را در پرتاب‌های اولیه به کار گیرد. مسئله‌ای که در این قسمت پیش می‌آید این است که در عمل، افزایش دقت در بازی نیزک، با کم کردن شعاع پرتاب یک توزیع یکنواخت مسجل نمی‌شود؛ چراکه دقت در بازی، تابعی از یک فاصله (شعاع در توزیع یکنواخت) نیست و شعاع این توزیع تنها توانایی نشان دادن وضعیت نشانه‌گیری بازیکن را به شکلی که در قسمت قبل توضیح دادیم، دارد. هرچند که همه مدل‌ها غلط هستند و برخی مفید، ما در این بین به دنبال یافتن مدل مفیدتریم. از این‌روست که به نظر می‌رسد، جایگزینی مدل مبتنی بر توزیع یکنواخت با مدل مبتنی بر توزیع گاوسی، انطباق بیشتری با ماهیت نشانه‌گیری یک نقطه داشته باشد. چراکه تلاش پرتاب‌کننده، هدف قرار دادن نقطه نشانه‌گیری است و نواحی نزدیک به نقطه هدف، علی‌القاعده می‌بایست شانس بیشتری برای مورد اصابت واقع شدن داشته باشند تا نواحی دورتر. در این بخش فرض می‌کنیم بازیکن تنها یک‌بار تغییر دقت (واریانس) را در پرتاب‌های خود تجربه می‌کند؛ به بیانی دیگر اگر τ را شماره پرتاب تغییر واریانس از σ_{τ}^2 در نظر بگیریم، از پرتاب $\tau + 1$ به بعد پرتاب‌ها با دقت و در نتیجه واریانس متفاوت $\sigma_{\tau}^2 \neq \sigma_{\tau+1}^2$ انجام خواهند شد. حال اگر $\sigma_{\tau} < \sigma_{\tau+1}$

درست‌نمایی به صورت

$$\begin{aligned} \ell(\mu, \sigma_1^2, \sigma_2^2, \tau | \mathbf{X}, Z) &= \log L(\mu, \sigma_1^2, \sigma_2^2, \tau | \mathbf{X}, Z) \\ &= -n \log(\sqrt{2\pi}) - \frac{\tau}{\sqrt{2\pi}} \log |\Sigma_1| - \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{\tau} \mathbf{Z}'_i \Sigma_1^{-1} \mathbf{Z}_i \quad (7) \\ &= \frac{-n - \tau}{\sqrt{2\pi}} \log |\Sigma_2| - \frac{1}{\sqrt{2\pi}} \sum_{i=\tau+1}^n \mathbf{Z}'_i \Sigma_2^{-1} \mathbf{Z}_i, \end{aligned}$$

است. به طریق مشابه، در مرحله t ام الگوریتم، برای پیاده‌سازی الگوریتم EM و در گام امیدگیری، امید ریاضی لگاریتم درست‌نمایی داده‌های کامل را به شرط داده‌های مشاهده‌شده بر پایه برآوردهای مرحله قبل محاسبه می‌کنیم. بنابراین، اگر در اینجا \mathbb{E}_t را امید ریاضی تحت فرض $\Sigma_2 = \widehat{\Sigma}_{2,t}$ و $\Sigma_1 = \widehat{\Sigma}_{1,t}$ تعریف کنیم، در این صورت

$$\begin{aligned} &\mathbb{E}_{t-1}[\ell(\Sigma_1, \Sigma_2, \tau; \mathbf{X}, Z | \mathbf{X})] \\ &= \frac{-\tau}{\sqrt{2\pi}} \log |\Sigma_1| - \frac{1}{\sqrt{2\pi}} \text{tr} \left(\Sigma_1^{-1} \sum_{i=1}^{\tau} \mathbb{E}_{t-1}(\mathbf{Z}_i \mathbf{Z}'_i) \right) \\ &= \frac{-n - \tau}{\sqrt{2\pi}} \log |\Sigma_2| - \frac{1}{\sqrt{2\pi}} \text{tr} \left(\Sigma_2^{-1} \sum_{i=\tau+1}^n \mathbb{E}_{t-1}(\mathbf{Z}_i \mathbf{Z}'_i) \right). \end{aligned}$$

در گام ماکسیم‌سازی، با ماکسیم کردن امید لگاریتم درست‌نمایی داده‌های کامل، و با توجه به استقلال پرتاب‌ها، برآوردهای

$$\begin{aligned} \widehat{\Sigma}_{1,t} &= \frac{1}{\tau} \sum_{i=1}^{\tau} \mathbb{E}_{t-1}(\mathbf{Z}_i \mathbf{Z}'_i | X_i) \\ \widehat{\Sigma}_{2,t} &= \frac{1}{n - \tau} \sum_{i=\tau+1}^n \mathbb{E}_{t-1}(\mathbf{Z}_i \mathbf{Z}'_i | X_i), \quad (8) \end{aligned}$$

حاصل می‌شود. دوباره برآورد τ یک مسئله $O(n)$ مضاعف است که به روش بخش قبل انجام می‌گیرد.

۳ رهیافت بیزی

در این بخش رهیافت بیزی برآورد پارامترها را تحت وجود یک نقطه تغییر موردبررسی قرار می‌دهیم. برای این منظور دقت داریم که تابع درست‌نمایی داده‌های کامل شامل بخشی مشاهده نشده است. لذا آنچه در ظاهر امر به نظر می‌رسد آن است که می‌توان مختصات‌های محل اصابت را به‌عنوان پارامتر در نظر گرفت و با وضع یک توزیع پیشین مناسب روی آن‌ها می‌توان پارامترهای موردعلاقه در مدل فراوانی‌گرایی (۴) را برآورد کرد.

۱.۳ برآورد بیزی مدل بدون نقطه تغییر

ابتدا برای سادگی و دریافت درک درستی از فضای مشاهدات، فرض کنید $\sigma_1 = \sigma_2$ یا $\tau = 1$. در این حالت تغییری در دقت پرتاب‌ها

در گام ماکسیم‌سازی، مشتق جزئی امید ریاضی را نسبت به پارامترهای موردعلاقه محاسبه و معادلات برآورد را به دست می‌آوریم:

$$\begin{aligned} &\frac{\partial \mathbb{E}_{t-1}[\ell(\sigma_1^2, \sigma_2^2, \tau; \mathbf{X}, Z | \mathbf{X})]}{\partial \sigma_1^2} \\ &= \frac{-\tau \sqrt{2\pi}}{2\pi \sigma_1^2} + \frac{2 \mathbb{E}_{t-1}[\sum_{i=1}^{\tau} (Z_{i,x}^2 + Z_{i,y}^2)]}{4\sigma_1^2} = 0, \\ &\frac{\partial \mathbb{E}_{t-1}[\ell(\sigma_1^2, \sigma_2^2, \tau; \mathbf{X}, Z | \mathbf{X})]}{\partial \sigma_2^2} \\ &= \frac{-(n - \tau) \sqrt{2\pi}}{2\pi \sigma_2^2} + \frac{2 \mathbb{E}_{t-1}[\sum_{i=\tau+1}^n (Z_{i,x}^2 + Z_{i,y}^2)]}{4\sigma_2^2} = 0, \end{aligned}$$

که برآوردهای به‌روزرسانی شده مرحله t ام $\widehat{\sigma}_{1,t}^2$ و $\widehat{\sigma}_{2,t}^2$ را به صورت

$$\begin{aligned} \widehat{\sigma}_{1,t}^2 &= \frac{\mathbb{E}_{t-1}[\sum_{i=1}^{\tau} (Z_{i,x}^2 + Z_{i,y}^2) | \mathbf{X}]}{2\tau}, \\ \widehat{\sigma}_{2,t}^2 &= \frac{\mathbb{E}_{t-1}[\sum_{i=\tau+1}^n (Z_{i,x}^2 + Z_{i,y}^2) | \mathbf{X}]}{2(n - \tau)}, \end{aligned}$$

تولید می‌کند. حال در پیاده‌سازی الگوریتم، با دو مقدار آغازین $\sigma_{1,0}$ و $\sigma_{2,0}$ الگوریتم را آغاز و تا همگرا شدن مقادیر برآورد، آن را ادامه می‌دهیم. با توجه به متناهی و شمارا بودن فضای پارامتر τ ، می‌توان این برآوردها را به ازای جمیع مقادیر τ محاسبه کرد. پس با جایگذاری τ از مجموعه $\{3, \dots, n-2\}$ می‌توان در هر مرحله برآورد منحصر به فرد وابسته به τ از مقادیر σ_1 و σ_2 به دست آورد. با قرار دادن مقدار برآوردهای به‌دست‌آمده از هر یک از این‌ها در $\ell(\mu, \sigma_1^2, \sigma_2^2, \tau | \mathbf{X}, Z)$ مقدار τ ماکسیم کننده انتخاب می‌شود. البته این کار منجر به اضافه شدن یک عبارت $O(n)$ در هزینه محاسباتی خواهد شد.

۳.۲ برآورد دقت با فرض مختصات وابسته و

ناهمگونی واریانس

اگر از فرض استقلال مؤلفه‌های محل اصابت صرف نظر کنیم، توزیع خطا در مدل (۴) به احتساب وجود یک نقطه تغییر به صورت

$$\varepsilon_1, \dots, \varepsilon_{\tau} \stackrel{iid}{\sim} N_2(\mathbf{0}, \Sigma_1) \quad \text{و} \quad \varepsilon_{\tau+1}, \dots, \varepsilon_n \stackrel{iid}{\sim} N_2(\mathbf{0}, \Sigma_2), \quad (6)$$

است که در آن Σ_1 و Σ_2 ماتریس‌های کوواریانس نامعلوم و N_2 نماد توزیع نرمال دو متغیره است. لذا به‌مانند گذشته، لگاریتم تابع

در می‌آید که تداعی‌کننده پیشین مزدوج وارون ویشارت است. به همین منظور برای تکمیل مدل بیزی، پیشین

$$\Sigma \sim IW(\Psi, m),$$

پیشنهاد می‌شود که توزیع وارون ویشارت با ماتریس مقیاس Ψ و درجه آزادی m است و پسین به شرط داده‌های کامل

$$\Sigma | \mathbf{X}, Z \sim IW\left(\Psi + \sum_{i=1}^n \mathbf{Z}_i \mathbf{Z}'_i, n + m\right), \quad (13)$$

را به دست می‌دهد. این بار نیز برای راه‌یابی به یک الگوریتم نمونه‌گیر گیبز نیاز به توزیع‌های تمام‌شرطی از جمله توزیع شرطی داده‌های مشاهده نشده به شرط داده‌های مشاهده شده و پارامترهای نامعلوم یعنی $Z | \mathbf{X}, \Sigma$ داریم که مانند (۱۱) یک توزیع نرمال دومتغیره مرکزی تحدیدشده روی $R_{\mathbf{x}}$ با ماتریس مقیاس Σ است. لذا، توزیع‌های تمام‌شرطی برای بهره‌مندی از یک الگوریتم نمونه‌گیر گیبز به صورت

$$\begin{aligned} \mathbf{Z}_{i,k} &\sim \pi_{\mathbf{Z}_i | X_i, \Sigma}(\cdot | x_i, \Sigma_{k-1}), \quad i = 1, \dots, n, \\ \Sigma_k &\sim IW\left(\Psi + \sum_{i=1}^n \mathbf{Z}_{i,k} \mathbf{Z}'_{i,k}, n + m\right), \end{aligned}$$

حاصل می‌شود که نمونه‌های تولیدشده بعد از یک اجرای به اندازه کافی طولانی، به‌طور مجانبی نمونه‌هایی تصادفی از توزیع‌های پسین $\pi_{\Sigma | \mathbf{X}}$ و $\pi_{\mathbf{Z}_i | X_i}$ خواهد بود و می‌توان ویژگی‌های گشتاوری این توزیع‌های پسین و به‌ویژه پسین دوم را با استفاده از گشتارهای نمونه‌ای حاصل از نمونه‌های شبیه‌سازی شده و قانون قوی اعداد بزرگ به دست آورد.

۲.۳ برآورد بیزی یک نقطه تغییر و

پارامترهای دقت

وجود نقطه تغییر، بُعد فضای پارامتر را تغییر می‌دهد. برای سادگی تعداد نقطه تغییرها برابر با یک در نظر گرفته می‌شود. باحالت استقلال مؤلفه‌های محل اصابت شروع می‌کنیم و به همین منظور، مدل (۴) را در نظر بگیرید که به دنبال آن تابع درستمایی پارامترها برحسب مشاهدات کامل، به صورت

$$\begin{aligned} f_{\mathbf{X}, Z | \sigma_1^2, \sigma_2^2, \tau}(\mathbf{x}, z) &= (2\pi)^{-n} (\sigma_1^2)^{-\tau} (\sigma_2^2)^{\tau-n} \quad (14) \\ &\times \exp\left(-\sum_{i=1}^{\tau} \mathbf{z}'_i \mathbf{z}_i / (2\sigma_1^2) - \sum_{i=\tau+1}^n \mathbf{z}'_i \mathbf{z}_i / (2\sigma_2^2)\right) \\ &\times \prod_{i=1}^n I\{s(\mathbf{z}_i) = x_i\}, \end{aligned}$$

رخ نداده است و اگر $\mathbf{X} = (X_1, \dots, X_n)$ بردار امتیازهای مستقل مشاهده شده و $Z = (Z_1 | \dots | Z_n)$ مختصات‌های مشاهده نشده یا ثبت نشده محل اصابت باشند که $\mathbf{Z}'_i = (Z_{i,x}, Z_{i,y})$ در این صورت بدون کاستن از کلیت با فرض $\mu = (0, 0)'$ در ادبیات آمار بیزی داریم

$$\begin{aligned} f_{\mathbf{X}, Z | \sigma^2}(\mathbf{x}, z) &= (2\pi)^{-n} (\sigma^2)^{-n} \quad (9) \\ &\times \exp\left(-\sum_{i=1}^n \mathbf{z}'_i \mathbf{z}_i / (2\sigma^2)\right) \prod_{i=1}^n I\{s(\mathbf{z}_i) = x_i\}. \end{aligned}$$

تابع درستمایی حاصل برای پارامتر مورده‌علاقه σ^2 ، پیشین مزدوج گاما وارون را پیشنهاد می‌دهد. لذا با پیشین $\sigma^2 \sim IG(a, b)$ برای $a, b > 0$ مفروض خواهیم داشت

$$\sigma^2 | \mathbf{X}, Z \sim IG\left(n + a, b + \sum_{i=1}^n \mathbf{Z}'_i \mathbf{Z}_i / 2\right). \quad (10)$$

پسین حاصل، به شرط مشاهدات کامل است که در عمل قابل به‌کارگیری نیست. این مسئله برای بهره‌گیری از یک نمونه‌گیر گیبز، نیاز به یک توزیع شرطی $Z | \mathbf{X}, \sigma^2$ دارد که به این منظور داریم

$$\pi_{\mathbf{Z}_i | X_i, \sigma^2}(\mathbf{z}_i | x_i, \sigma^2) \propto \exp\left(-\frac{\mathbf{z}'_i \mathbf{z}_i}{2\sigma^2}\right) I\{s(\mathbf{z}_i) = x_i\}, \quad (11)$$

که تابع چگالی نرمال دومتغیره مرکزی تحدیدشده با ماتریس کوواریانس $\sigma^2 I_{2 \times 2}$ روی ناحیه

$$R_{\mathbf{x}} = \{(x, y) \in \mathbb{R}^2 | s(x, y) = x_i, i = 1, \dots, n\},$$

است. شبیه‌سازی از این توزیع کار سختی نیست و می‌توان با استفاده از الگوریتم پذیرش و رد [۱، ص. ۲۹۲]، نمونه‌های تصادفی از آن شبیه‌سازی کرد. از این رو می‌توان تنظیمات زیر را برای پیاده‌سازی یک نمونه‌گیر گیبز فراهم کرد:

$$\begin{aligned} \mathbf{Z}_{i,k} &\sim \pi_{\mathbf{Z}_i | X_i, \sigma^2}(\cdot | x_i, \sigma_{k-1}^2), \quad i = 1, \dots, n, \\ \sigma_k^2 &\sim IG\left(n + a, b + \sum_{i=1}^n \mathbf{Z}'_{i,k} \mathbf{Z}_{i,k} / 2\right). \end{aligned}$$

در این صورت، به ازای $k^* \in \mathbb{N}$ به اندازه کافی بزرگ و $k > k^*$ می‌توان چنین در نظر گرفت که $\mathbf{Z}_{i,k}$ ها و σ_k^2 از توزیع مانای فرایند مارکوف فوق یعنی به ترتیب از $\pi_{\mathbf{Z}_i | \mathbf{X}}$ و $\sigma^2 | \mathbf{X}$ تولیدشده‌اند. تعمیم این موضوع به حالتی که مؤلفه‌های مختصات محل اصابت وابسته‌اند، با الهام از مدل بیزی گفته شده، کار چندان دشواری نیست. توجه داریم که در این حالت تابع درستمایی به صورت

$$\begin{aligned} f_{\mathbf{X}, Z | \Sigma}(\mathbf{x}, z) &= (2\pi)^{-n} |\Sigma|^{-n/2} \quad (12) \\ &\times \exp\left(-\text{tr}\left(\Sigma^{-1} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}'_i\right) / 2\right) \prod_{i=1}^n I\{s(\mathbf{z}_i) = x_i\}, \end{aligned}$$

در می‌آید که شبیه‌سازی نخست به ازای $i = 1, \dots, n$ انجام می‌شود. دوباره به ازای یک k^* به اندازه کافی بزرگ، نمونه‌های تولیدشده بعد از تکرار k^* ام را می‌توان به مثابه نمونه‌هایی تصادفی از توزیع‌های پسین σ_1^2, σ_2^2 و τ در نظر گرفت. پس از تعیین مقدار τ ، تنها نمونه‌هایی که به شرط مقدار τ تعیین شده تولید شده‌اند، مورد استفاده قرار می‌گیرند. نکته این تنظیمات، بخش آخر آن است که توزیع تمام شرطی τ ، توزیع شناخته‌شده‌ای نیست و می‌بایست برای تولید از این توزیع یک سازوکار دیگر معرفی کرد. به همین منظور از الگوریتم MCMC استفاده می‌کنیم و توزیع پیشنهادی الگوریتم را به گونه‌ای در نظر می‌گیریم که احتمال‌های پذیرش الگوریتم، قابل ملاحظه باشد. به منظور ساختن یک الگوریتم MCMC برای تولید نمونه تصادفی از توزیع (۱۸) نیاز به یک توزیع پیشنهادی داریم که بتواند مقادیر فعلی تولیدشده را، که با τ^c نمایش می‌دهیم، به مقادیر پیشنهادی τ^p ، نسبت دهد که این نسبت از توزیع شرطی پیشنهادی متناظر با تابع چگالی $q_{\tau^p|\tau^c}(\cdot|\cdot)$ حاصل می‌شود. لذا در مرحله زام، نمونه زام یعنی $\tau^{(j)}$ ، به جای تولید از توزیع (۱۸)، از توزیع معلوم $q_{\tau^p|\tau^c}(\cdot|\tau^{(j-1)})$ تولید می‌شود و نمونه تولیدشده با احتمال $\min\{1, A\}$ پذیرفته می‌شود که در آن

$$A = \frac{\pi_{\tau|\mathbf{X}, \sigma_1^2, \sigma_2^2, Z}(\tau^{(j)}|z, \mathbf{x}, \sigma_1^2, \sigma_2^2) q_{\tau^p|\tau^c}(\tau^{(j-1)}|\tau^{(j)})}{\pi_{\tau|\mathbf{X}, \sigma_1^2, \sigma_2^2, Z}(\tau^{(j-1)}|z, \mathbf{x}, \sigma_1^2, \sigma_2^2) q_{\tau^p|\tau^c}(\tau^{(j)}|\tau^{(j-1)})}$$

دقت داریم که τ نقطه تغییر فرایند محل اصابت پرتاب‌ها است و نباید دقت برآورد را فدای سرعت محاسبات کرد. اهمیت این عبارت را به زودی درمی‌یابیم. در این مجال ما از یکی از ساده‌ترین توزیع‌های پیشنهادی بهره می‌بریم که اتفاقاً ریشه در تعریف ما از پارامتر نقطه تغییر دارد. تصور کنید در مرحله ۱ - زام الگوریتم MCMC برای تولید نمونه از توزیع مطرح شده در (۱۸) هستیم. زمانی که به دنبال تغییر مکان نقطه تغییر شبیه‌سازی شده و در واقع به دنبال نقل مکان نقطه تغییر باشیم، این کار را با احتیاط و گام‌های کوچک انجام می‌دهیم. از آنجایی که کاربرد نسبت به محل دقیق نقطه تغییر آگاه نیست، در مورد پیشنهاد مقصد تغییر به صورت ناآگاهانه عمل می‌کند و تحت یک فضای هم‌شانس، می‌تواند یک قدم به چپ، یعنی $\tau^{(j-1)-1}$ ، یا به راست، یعنی $\tau^{(j-1)+1}$ برود. در نتیجه،

به دست می‌آید. دور از انتظار نیست که مقدار تولیدشده برای دقت (وارون واریانس) مستقل از زمان وقوع تغییر در مقادیر دقت، یعنی τ باشد. بنابراین، برای حصول یک مدل بیزی تحت فرض استقلال مؤلفه‌های مختصات محل اصابت، ساختار پیشینی

$$\sigma_j^2|\tau \sim IG(a_j, b_j), \quad j = 1, 2 \quad (15)$$

$$\tau \sim U(\{3, \dots, n-2\}), \quad (16)$$

پیشنهاد می‌شود که در آن $U(A)$ توزیع یکنواخت گسسته روی مجموعه شمارای A است. همچنین با فرض استقلال پیشین‌های σ_1 و σ_2 و با توجه به درستی (۱۴)، داریم

$$\sigma_1^2|\mathbf{X}, Z, \tau \sim IG\left(\tau + a_1, b_1 + \sum_{i=1}^{\tau} \mathbf{z}'_i \mathbf{z}_i / 2\right)$$

$$\sigma_2^2|\mathbf{X}, Z, \tau \sim IG\left(n - \tau + a_2, b_2 + \sum_{i=\tau+1}^n \mathbf{z}'_i \mathbf{z}_i / 2\right).$$

همچنین، توزیع تمام شرطی مختصات‌های محل اصابت در این حالت به صورت

$$\pi_{\mathbf{z}_i|\mathbf{X}_i, \sigma_1^2, \sigma_2^2, \tau}(\mathbf{z}_i|x_i, \sigma_1^2, \sigma_2^2, \tau) \quad (17)$$

$$\propto \exp\left(-\frac{\mathbf{z}'_i \mathbf{z}_i}{2\sigma_1^2}\right) I\{s(\mathbf{z}_i) = x_i\} I\{\tau \leq i\}$$

$$+ \exp\left(-\frac{\mathbf{z}'_i \mathbf{z}_i}{2\sigma_2^2}\right) I\{s(\mathbf{z}_i) = x_i\} I\{\tau > i\},$$

است که توزیع آمیخته دو بردار گاوسی دومتغیره مرکزی به ترتیب با ماتریس‌های کوواریانس $\sigma_1^2 I_{\tau \times \tau}$ و $\sigma_2^2 I_{\tau \times \tau}$ و به ترتیب با احتمال آمیختگی $\alpha = P(\tau \leq i)$ و $1 - \alpha$ است که روی ناحیه R_X تحدید شده است. برای ساختن یک نمونه‌گیر گیز مشابه قبل، نیاز به شبیه‌سازی از توزیع تمام شرطی $\tau|Z, \mathbf{X}, \sigma_1^2, \sigma_2^2$ داریم. دقت داریم که با ضرب تابع درستی‌مایی در چگالی‌های پیشین، این توزیع پسین به صورت

$$\pi_{\tau|\mathbf{X}, \sigma_1^2, \sigma_2^2, Z}(t|z, \mathbf{x}, \sigma_1^2, \sigma_2^2) \propto \left(\frac{\sigma_1^2}{\sigma_2^2}\right)^{-t} \quad (18)$$

$$\times \exp\left(-\sum_{i=1}^t \mathbf{z}'_i \mathbf{z}_i / (2\sigma_1^2) - \sum_{i=t+1}^n \mathbf{z}'_i \mathbf{z}_i / (2\sigma_2^2)\right),$$

حاصل می‌شود. لذا تنظیمات نمونه‌گیر گیز در این حالت به شکل

$$\mathbf{Z}_{i,k} \sim \pi_{\mathbf{z}_i|\mathbf{X}_i, \sigma_{1,k-1}, \sigma_{2,k-1}, \tau_{k-1}}(\cdot|x_i, \sigma_{1,k-1}, \sigma_{2,k-1}, \tau_{k-1});$$

$$\sigma_{1,k} \sim IG\left(\tau_{k-1} + a_1, b_1 + \sum_{i=1}^{\tau_{k-1}} \mathbf{z}'_{i,k} \mathbf{z}_{i,k} / 2\right);$$

$$\sigma_{2,k} \sim IG\left(n - \tau_{k-1} + a_2, b_2 + \sum_{i=\tau_{k-1}+1}^n \mathbf{z}'_{i,k} \mathbf{z}_{i,k} / 2\right);$$

$$\tau_k \sim \pi_{\tau|\mathbf{X}, \sigma_{1,k}, \sigma_{2,k}, Z}(\cdot|z_k, \mathbf{x}, \sigma_{1,k}, \sigma_{2,k});$$

توزیع پیشنهادی را به صورت

$$q_{\tau^c | \tau^c}(t | \tau^{(j-1)}) = \begin{cases} 1/2, & t = \tau^{(j-1)} - 1, \tau^{(j-1)} + 1 \\ & \text{و } \tau^{(j-1)} \neq 3, n-2, \\ 1, & t = \tau^{(j-1)} + 1 \text{ و } \tau^{(j-1)} = 3, \\ 1, & t = \tau^{(j-1)} - 1 \text{ و } \tau^{(j-1)} = n-2, \end{cases} \quad (19)$$

پیشنهاد می‌دهیم. همان‌طور که دیده می‌شود، در هر مرحله، مقدار پیشنهادی نقطه تغییر تنها یک گام جابه‌جا می‌شود که می‌توان این موضوع را به گام‌های بیشتر نیز تعمیم داد که البته در این صورت، سرعت همگرایی آماره‌های نمونه‌ای پسینی نیز کاهش می‌یابد. نتایج شبیه‌سازی در بخش بعد، مؤید کارایی این الگوریتم است. به پیروی از این روش، می‌توان ایده برآورد بیزی را برای حالتی که مختصات‌های محل اصابت وابسته باشند نیز تعمیم داد. برای این منظور، در مدل (۴)، توزیع خطا را به صورت (۶) در نظر گیرید. لذا تابع درست‌نمایی‌ای به صورت زیر داریم:

$$f_{\mathbf{X}, Z | \Sigma}(\mathbf{x}, z) = (\tau\pi)^{-n} |\Sigma_1|^{-\tau/\tau} |\Sigma_2|^{-(n-\tau)/\tau} \times \exp\left(-\text{tr}\left(\Sigma_1^{-1} \sum_{i=1}^{\tau} \mathbf{z}_i \mathbf{z}_i + \Sigma_2^{-1} \sum_{i=\tau+1}^n \mathbf{z}_i \mathbf{z}_i\right) / \tau\right) \times \prod_{i=1}^n I\{s(\mathbf{z}_i) = x_i\}. \quad (20)$$

با همان توجه قبل، پیشین‌های Σ_1, Σ_2 را مستقل از هم و به صورت

$$\Sigma_j \sim \mathcal{IW}(\Psi_j, m_j), \quad j = 1, 2,$$

و پیشین τ را نیز همچنان یکنواخت گسسته روی $\{3, \dots, n-2\}$ در نظر می‌گیریم. در این حالت، توزیع پسین Σ_1 و Σ_2 عبارت خواهد بود از

$$\Sigma_1 | \mathbf{X}, Z, \tau \sim \mathcal{IW}\left(\Psi_1 + \sum_{i=1}^{\tau} \mathbf{Z}_i \mathbf{Z}_i', \tau + m_1\right),$$

$$\Sigma_2 | \mathbf{X}, Z, \tau \sim \mathcal{IW}\left(\Psi_2 + \sum_{i=\tau+1}^n \mathbf{Z}_i \mathbf{Z}_i', n - \tau + m_2\right).$$

چگالی مختصات‌های مشاهده نشده $Z = (Z_1, \dots, Z_n)$ به شرط سایر پارامترها \mathbf{X} نیز چگالی ماتریسی با ستون‌های دو بُعدی مستقل از هم است که چگالی ستون i ام، $i = 1, \dots, n$ ، به صورت

$$\pi_{\mathbf{z}_i | X_i, \Sigma_1^i, \Sigma_2^i, \tau}(\mathbf{z}_i | x_i, \Sigma_1^i, \Sigma_2^i, \tau) \propto \exp(\mathbf{z}_i' \Sigma_1^{-1} \mathbf{z}_i / 2) I\{s(\mathbf{z}_i) = x_i\} I\{\tau \leq i\} + \exp(-\mathbf{z}_i' \Sigma_2^{-1} \mathbf{z}_i / 2) I\{s(\mathbf{z}_i) = x_i\} I\{\tau > i\},$$

است که توزیع آمیخته دو بردار گاوسی دومتغیره مرکزی به ترتیب با ماتریس‌های کوواریانس Σ_1 و Σ_2 و به ترتیب با احتمال آمیختگی α و $1 - \alpha$ است که روی ناحیه R_X تحدید شده است. به روشی مشابه توزیع تمام شرطی τ از تناسب زیر حاصل می‌شود:

$$\pi_{\tau | \mathbf{X}, \Sigma_1, \Sigma_2, Z}(t | z, \mathbf{X}, \Sigma_1, \Sigma_2) \propto \left(\frac{|\Sigma_1|}{|\Sigma_2|}\right)^{-t/\tau} \times \exp\left(-\sum_{i=1}^t \mathbf{z}_i' \Sigma_1^{-1} \mathbf{z}_i / 2 - \sum_{i=t+1}^n \mathbf{z}_i' \Sigma_2^{-1} \mathbf{z}_i / 2\right). \quad (21)$$

تکرارهای الگوریتم نمونه‌گیر گیبز در این حالت، نمونه‌هایی به شکل

$$\mathbf{Z}_{i,k} \sim \pi_{\mathbf{Z}_i | \mathbf{X}, \Sigma_1, \Sigma_2, Z, \tau_{k-1}}(\cdot | x_i, \Sigma_1, \Sigma_2, \tau_{k-1});$$

$$\Sigma_{1,k} \sim \mathcal{IW}\left(\Psi_1 + \sum_{i=1}^{\tau} \mathbf{Z}_{i,k} \mathbf{Z}_{i,k}', \tau_{k-1} + m_1\right);$$

$$\Sigma_{2,k} \sim \mathcal{IW}\left(\Psi_2 + \sum_{i=\tau+1}^n \mathbf{Z}_{i,k} \mathbf{Z}_{i,k}', n - \tau_{k-1} + m_2\right);$$

$$\tau_k \sim \pi_{\tau | Z, \mathbf{X}, \Sigma_1, \Sigma_2, \tau_{k-1}}(\cdot | z, \mathbf{X}, \Sigma_1, \Sigma_2, \tau_{k-1});$$

است که برای تولید نمونه از چگالی (۲۱) از الگوریتم MCMC با همان چگالی پیشنهادی (۱۹) استفاده می‌کنیم.

۴ مطالعه عددی

۱.۴ مطالعه شبیه‌سازی

مطالعه عددی مباحث گفته‌شده به تفکیک دو حالت مختصات‌های محل برخورد مستقل و وابسته ارائه می‌شود. در هر دو حالت، برآورد پارامترها تحت دو رهیافت فراوانی‌گرا و بیزی برای نمونه‌های شبیه‌سازی‌شده و برای یک دنباله از امتیازهای واقعی حاصل از پرتاب‌های یک پرتاب‌کننده بررسی می‌شود. در مطالعات شبیه‌سازی، ابتدا با استفاده از مدل (۴) و تحت دو ساختار ماتریس کوواریانس قطری و غیر قطری داده‌های محل اصابت، Z ، تولید می‌شود و سپس با استفاده از ساختار امتیازدهی بریتانیا، به هر مختصات، امتیازی تعلق می‌گیرد. بعد از این مرحله تمام فرایند برآورد مبتنی بر امتیازهای مشاهده‌شده \mathbf{X} است و مراجعه‌ای به مقادیر Z صورت نمی‌گیرد. در فرایند این مطالعه عددی، نکته مهم زمانی است که ماتریس کوواریانس مؤلفه‌های محل برخورد قطری نبوده و دو مؤلفه از طریق یک مقدار کوواریانس به هم وابسته‌اند. در این حالت و به‌ویژه زمانی که قدرمطلق مقدار ضریب همبستگی خطی دو مؤلفه مختصات نسبتاً بزرگ است، از آنجایی که بیضی‌های کانتر

و سپس مختصات Z_q^* را از توزیع نرمال دو متغیره $N_2((0,0)', \Sigma_{t-1})$ محدود شده روی ناحیه انتخاب شده C_l تولید کنید. این کار را به ازای $q = 1, \dots, 10000$ تکرار می‌کنیم و با استفاده از قانون قوی اعداد بزرگ، امید ریاضی شرطی را به صورت زیر تقریب می‌زنیم:

$$\mathbb{E}_{t-1}[Z_{i,x}^* + Z_{i,y}^* | X_i = 12] \approx \sum_{q=1}^{10000} \|Z_q^*\|^2 / 10000.$$

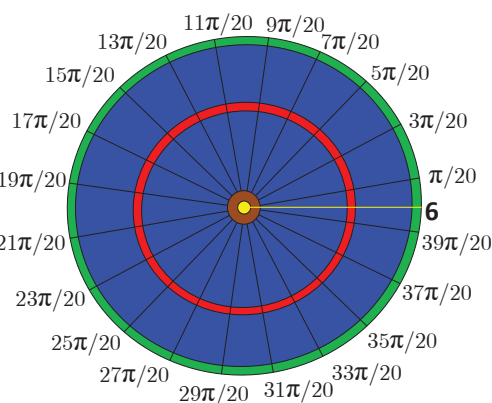
یکی از مشکلات این نوع محاسبه زمانی رخ می‌دهد که مقدار ضریب همبستگی دو مؤلفه خیلی شدید باشد. چراکه در این حالت کانتورهای توزیع نرمال دو متغیره Z_i ها بیضی‌های بسیار کشیده‌ای خواهند بود و احتمال وقوع برخی از نواحی نزدیک به صفر است. یکی از مقادیر p_l بسیار کوچک است ولی اگر در انتخاب ناحیه برای شبیه‌سازی از Z^* ، یک C_l با احتمال کوچک انتخاب شود، فرایند تولید نمونه 10000 تایی از این ناحیه بسیار زمان‌بر خواهد شد. هرچند این اتفاق با احتمال کوچکی رخ می‌دهد ولی در شبیه‌سازی‌های طولانی، ممکن است بروز کند. نکته دیگری که می‌بایست به آن بپردازیم، محاسبه احتمال‌های p_l گفته شده در بالا است. به همین منظور، ابتدا ناحیه C_l را به مستطیل‌هایی تقسیم‌بندی می‌کنیم. به بیانی دیگر، ما به دنبال پوشش دادن C_l با استفاده از یک توری منظمیم؛ به این صورت که خط‌هایی موازی (با شیب یکسان) روی قاچ متناظر با C_l ترسیم می‌کنیم و معادله این خطوط را می‌نویسیم. عملاً جزء تعیین‌کننده در این خط‌ها، اندازه عرض از مبدأ آن‌ها خواهد بود؛ چراکه شیب برای همه آن‌ها یکسان است. به عبارت دیگر معادله خط $y = cx + a_\delta$ را در نظر بگیرید که در آن $\delta = 1, \dots, D$ و c طوری تعیین می‌شود که شیب خط‌ها برابر با شیب یکی از یال‌های قاچ موردنظر باشد. مقدار D که برابر با تعداد خطوط است، با توجه به میزان دقت موردنظر ما از محاسبه p_l تغییر می‌کند؛ D بزرگ‌تر، توری دقیق‌تری را تولید می‌کند که دقت تخمین p_l را افزایش می‌دهد. برای تکمیل توری، کافی است در گام بعدی خط‌هایی با شیب $-c$ با معادلات $y = -cx + b_{\delta'}$ ترسیم کنیم که $\delta' = 1, \dots, D'$ (شکل ۵ را ببینید). مختصات گره‌های توری، طول و عرضی به صورت

$$x = \frac{b_{\delta'} - a_\delta}{2c},$$

$$y = \frac{b_{\delta'} + a_\delta}{2c},$$

دارند. در مرحله آخر باید احتمال این را که نقطه اصابت نیزک درون این مستطیل باشد، حساب و به ازای مستطیل‌های درون ناحیه موردبررسی باهم جمع کنیم. اگر A یکی از این مستطیل‌ها

توزیع نرمال دو متغیره Z_i ها بسیار کشیده است، برخی نقاط روی تخته، احتمال اصابت تقریباً صفر خواهند داشت. این موضوع اتفاقاً در بسته darts در R که توسط نویسندگان [۳] نوشته شده است، بدون آنکه تمهیدی در این خصوص اندیشیده شده باشد، رخ می‌دهد. این موضوع محاسبه $L(\Sigma_1, \Sigma_2, \tau | \mathbf{X}, Z)$ را دچار مشکل می‌کند. برای محاسبه امتیاز، کافی است که مختصات محل برخورد را در دستگاه قطبی محاسبه کنیم و بر اساس شعاع و زاویه مختصات قطبی می‌توان امتیاز را به سادگی بر اساس شکل ۴ محاسبه کرد. مثلاً امتیاز ۱۲ را در نظر بگیرید. این امتیاز یا از طریق برخورد به ناحیه تک امتیازی ۱۲ کسب می‌شود؛ یا برخورد به منطقه دو برابر ۶ امتیازی و یا اصابت به ناحیه سه برابر ۴ امتیازی. همان‌طور که در شکل می‌توان دید در منطقه تک امتیازی ۱۲، شعاع در محدوده $[159, 99]$ یا $[107, 162]$ قرار دارد. به علاوه زاویه تشکیل دهنده این قسمت در محدوده $[13\pi/20]$ در حالت دوم که منطقه دو برابر ۶ امتیازی را هدف محاسبه احتمال قرار دادیم، شعاع باید در محدوده $[162, 170]$ و زاویه در محدوده $[39\pi/20, 40\pi/20]$ یا $[0, \pi/20]$ باشد. حالت سوم نیز شعاع در محدوده $[99, 107]$ و زاویه در بازه $[3\pi/20, 5\pi/20]$ است.



شکل ۴: زاویه قطعه‌های تخته نیزک در مختصات قطبی

برای محاسبه کمیت $\mathbb{E}_{t-1}[Z_{i,x}^* + Z_{i,y}^* | X_i]$ از روش مونت‌کارلو استفاده می‌کنیم. این روش را به اختصار برای حالت کلی $Z_i \sim N_2((0,0)', \Sigma)$ وقتی که $X_i = 12$ توضیح می‌دهیم. تحت این شرط، نیزک ممکن است در هر یک از سه ناحیه گفته شده فرود آمده باشد. این سه ناحیه را با C_1 ، C_2 و C_3 نمایش دهید. تحت فرض $\Sigma = \Sigma_{t-1}$ که برآورد ماتریس کوواریانس متناظر در اجرای مرحله قبل است، احتمال هر یک از این سه ناحیه با اندکی ملاحظه قابل محاسبه است و این سه احتمال را به ترتیب با p_1 ، p_2 و p_3 نشان می‌دهیم. ناحیه C_l را با احتمال p_l به ازای $l = 1, 2, 3$ انتخاب کنید

شکل ۵: تقسیم‌بندی قاچی از صفحه نیزک به مستطیل‌های تشکیل‌دهنده

برای مطالعه عددی ابتدا $n = 50, 100, 200$ امتیاز از پرتاب‌هایی تحت مدل (۴) شبیه‌سازی می‌کنیم که در آن $\tau = 30, \sigma_1^2 = 100mm^2$ و $\sigma_2^2 = 90mm^2$. برای هر نمونه امتیازی تولیدشده، مقادیر پارامتری برآورد می‌شود و این کار را برای هر یک از این تنظیمات شبیه‌سازی، ۱۰۰ بار تکرار می‌کنیم. میانگین مقادیر برآوردشده و انحراف معیار تجربی در دو روش برآورد ماکسیمم درستنمایی با استفاده از الگوریتم EM و برآورد بیز در جدول ۱ ارائه شده‌اند. برای تعیین پارامترهای پیشین (۱۶) از نتایج روش ML استفاده می‌کنیم. در واقع پارامترهای a_i و b_i را طوری محاسبه می‌کنیم که امید ریاضی توزیع پیشین برابر با برآورد ماکسیمم درستنمایی و واریانس آن عدد قابل ملاحظه‌ای شود. مثلاً وقتی $n = 50$ مقدار برآورد ML پارامتر σ_2^2 برابر با 959210 است و بنابراین مقدار a_1 را برابر با ۳ و مقدار b_1 را برابر با $959210 \times (3-1)$ در نظر می‌گیریم. بدین ترتیب، واریانس پیشین نیز مقدار قابل ملاحظه‌ای خواهد داشت. همان‌طور که از این جدول پیداست، هر دو برآورد کم‌برازش دارند. هرچند که برآوردگر بیز وظیفه مینیمم کردن تابع مخاطره بیز را دارد، با این وجود از نگاه مخاطره تجربی نیز این برآوردگر عملکرد بهتری دارد. در تمامی برآوردگرهای بیز، ۱۰۰۰۰ نمونه تولیدشده نخست، به‌عنوان نمونه‌های پیش از داغیدن، کنار گذاشته شده است. همان‌طور که از این جدول پیداست، در هر دو برآوردگر، با افزایش اندازه نمونه، مقدار مخاطره تجربی از طریق کاهش اریبی و انحراف معیار برآوردگرها کاهش می‌یابد که به‌نوعی القاکننده سازگاری هر دو برآوردگر است.

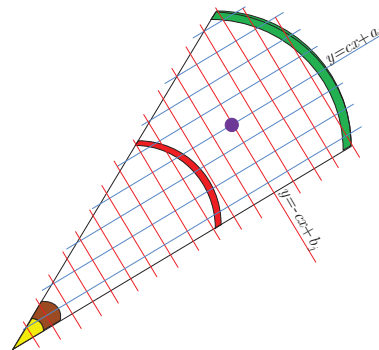
باشد، A را به اندازه $\theta = \arctan(c)$ دوران می‌دهیم تا اضلاع آن به‌صورت افقی و عمودی درآیند و مستطیل حاصل که آن را A^* می‌نامیم از طریق دو گره مقابل با مختصات‌های (x_1, y_1) و (x_2, y_2) قابل شناسایی باشد. ماتریس کوواریانس Σ_{t-1} را نیز به اندازه θ دوران می‌دهیم و تعریف کنید

$$\Sigma^* = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \widehat{\Sigma}_{t-1}$$

در این صورت اگر $Z^* \sim N_2((\cdot, \cdot)', \Sigma^*)$ آنگاه

$$\begin{aligned} P(Z_i \in A) &= P(Z^* \in A^*) \\ &= P(Z_{i,1}^* \leq x_2, Z_{i,2}^* \leq y_2) - P(Z_{i,1}^* \leq x_1, Z_{i,2}^* \leq y_2) \\ &\quad - P(Z_{i,1}^* \leq x_2, Z_{i,2}^* \leq y_1) + P(Z_{i,1}^* \leq x_1, Z_{i,2}^* \leq y_1). \end{aligned}$$

با جمع بستن مقادیر احتمال تمام مستطیل‌های محصور در ناحیه C_i مقدار p_i تقریب زده می‌شود. دقت دارید که تشخیص مستطیل‌های محصور در داخل یک ناحیه در مختصات قطبی به سبکی که برای محاسبه امتیازها گفتیم کار سختی نیست و بار محاسباتی ندارد.



جدول ۱. برآورد پارامترها با فرض استقلال مؤلفه‌های مختصات محل اصابت با دو روش ماکسیمم درستنمایی و بیزی. مقادیر داخل پراوتز، انحراف معیار تجربی حاصل از هر برآورد را نشان می‌دهد. در مورد برآورد τ مقدار نما در ۱۰۰ بار تکرار و در مورد دو پارامتر دیگر از میانگین نمونه‌ای استفاده شده است.

اندازه نمونه	برآورد ML			برآورد بیز		
	τ	σ_1^2	σ_2^2	τ	σ_1^2	σ_2^2
$n = 50$	۳۷(۳۷۹۸۸)	۸۶۴۲۶۲(۱۷۸۱۷۱)	۹۵۹۲۱۰(۲۱۱۲۰۷)	۲۶(۲۴۱۵۷)	۸۶۱۸۸۹(۱۵۶۲۵۷)	۹۵۸۰۹۲(۲۱۳۴۰۶۹)
$n = 100$	۳۶(۲۲۰۲۲)	۸۷۸۳۳۰(۱۰۷۲۳۳)	۹۷۶۳۴۸(۱۹۷۵۴۶)	۳۶(۱۴۶۸۴)	۸۸۳۸۶۱(۹۱۶۶۱)	۹۸۱۷۲۶(۱۹۳۲۹۰۳)
$n = 200$	۳۴(۱۰۰۲۱)	۸۸۹۳۸۹(۷۳۴۵۶)	۹۹۱۵۱۴(۱۲۳۴۵۸)	۳۴(۰۵۸۱۴)	۸۹۸۲۱۱(۶۲۰۸۲)	۹۹۷۰۲۱(۱۰۹۱۳۰)

مقدار کوواریانس دو مؤلفه را نمی‌توان عدد بزرگی قرار داد. البته این موضوع در عمل منعی ندارد ولی در مطالعه شبیه‌سازی،

شبیه‌سازی مشابهی را با همان تنظیم‌ها به‌علاوه افزودن یک پارامتر همبستگی در مختصات محل اصابت، تدارک دیده‌ایم.

مثلاً وقتی $n = 50$ برآورد ML ماتریس Σ_1 به صورت

$$\widehat{\Sigma}_1 = \begin{bmatrix} 955288 & 892224 \\ 892224 & 952198 \end{bmatrix},$$

است. میانگین یک ماتریس وارون و یشارت زمانی وجود دارد که درجه آزادی بزرگتر از مرتبه ماتریس به علاوه یک باشد. لذا با در نظر گرفتن $m_1 = 4$ ، پیشین را در این مطالعه برای $n = 50$ به صورت $\Sigma_1 \sim IW(\widehat{\Sigma}_1, 4)$ در نظر گرفتیم. البته استفاده از روش های بیز تجربی برای انتخاب ابرپارامترها بی تردید برآوردهای بهتری را بازمی گرداند لیکن اضافه کردن یک مسئله ماکسیم سازی برای محاسبه پارامترهای پیشین، هزینه محاسباتی را افزایش می دهد. با مقایسه جدول ۱ و جدول ۲، به طور کلی می توان گفت که برآورد مؤلفه های واریانس، در حالت مختصات های وابسته، از نظر اریبی و انحراف معیار تجربی بهتر از حالت مؤلفه های مستقل است. ولی دقت داریم که با در نظر گرفتن تابع زیان به صورت مجموع توان دو خطاهای همه پارامترها، مقدار مخاطره تجربی به دلیل بزرگ تر شدن بُعد فضای پارامتر، بزرگ تر می شود. به نظر می رسد که همین موضوع منجر به افزایش در مقدار اریبی برآورد نقطه تغییر در مقایسه با جدول ۱ شده است.

کوواریانس که قدر مطلق مقدار آن عدد قابل ملاحظه ای در مقایسه با مقادیر واریانس ها باشد، منجر به کاهش احتمال پذیرش نمونه گیری ها در تمام الگوریتم های پذیرش و رد و MCMC می شود. از این رو ماتریس کوواریانس مؤلفه های محل اصابت پیش و پس از نقطه تغییر را به صورت

$$\Sigma_1 = \begin{bmatrix} \sigma_{11,1} & \sigma_{12,1} \\ \sigma_{12,1} & \sigma_{22,1} \end{bmatrix} = \begin{bmatrix} 100 & 10 \\ 10 & 100 \end{bmatrix},$$

$$\Sigma_2 = \begin{bmatrix} \sigma_{11,2} & \sigma_{12,2} \\ \sigma_{12,1} & \sigma_{22,2} \end{bmatrix} = \begin{bmatrix} 90 & 5 \\ 5 & 90 \end{bmatrix},$$

در نظر گرفتیم. نقطه تغییر همچنان $\tau = 30$ در نظر گرفته شده است. با توجه به بزرگ تر شدن جدول، مقادیر برآورد و انحراف معیار تجربی پارامترهای $\sigma_{11,1}$ ، $\sigma_{22,1}$ و $\sigma_{12,1}$ در جدول ۲ گزارش شده اند، و مقادیر مربوط به پارامترها پس از نقطه تغییر، و همچنین برآورد نقطه تغییر در ادامه این جدول آورده شده است. نتایج مشابهی با حالت مؤلفه های وابسته مشاهده می شود و همچنان برآوردگر بیز عملکرد پذیرفتنی تری در مقایسه با برآوردگر ماکسیم درستیابی دارد. ایده مشابهی برای تعیین پیشین های ماتریس های Σ_1 و Σ_2 نظیر آنچه در مطالعه قبلی داشتیم استفاده شده است. به همین منظور

جدول ۲. میانگین برآوردها و انحراف معیار تجربی آن ها (اعداد داخل پرانتز) در ۱۰۰ بار تکرار آزمایش برای برآورد پارامترهای پیش از نقطه تغییر در مدل مؤلفه های مختصات وابسته محل اصابت با دو روش ماکسیم درستیابی و بیزی.

برآورد بیز			برآورد ML			اندازه نمونه
$\sigma_{12,1}$	$\sigma_{22,1}$	$\sigma_{11,1}$	$\sigma_{12,1}$	$\sigma_{22,1}$	$\sigma_{11,1}$	
۸۷۸۱۹(۳/۱۶۸۸)	۹۵۳۳۵۶(۲۳/۳۷۱)	۹۴۲۷۰۱(۲۷/۴۲۹۹)	۸۹۲۲۴(۳/۱۵۲۴)	۹۵۲۱۹۸(۲۷/۶۴۴۹)	۹۵۵۲۸۸(۲۹/۷۵۱۸)	$n = 50$
۹۳۹۴۷(۲/۴۶۱۲)	۹۶۳۰۵۸(۱۶/۹۸۰۰)	۹۶۰۸۴۷(۱۷/۹۴۷۵)	۹۳۵۵۲(۲/۹۷۰۴)	۹۶۹۸۱۴(۱۷/۳۱۳۹)	۹۶۱۸۶۱(۱۹/۸۸۱۶)	$n = 100$
۹۹۰۲۶(۱/۳۷۶۹)	۹۹۲۵۸۰(۱۰/۵۲۴۶)	۹۹۰۸۲۸(۱۰/۹۳۲۸)	۹۸۷۲۱(۱/۹۳۶۱)	۸۹۱۴۱۱(۱۱/۲۱۷۹)	۹۸۸۰۶۵(۱۲/۳۶۵۲)	$n = 200$

ادامه جدول ۲. میانگین برآوردها و انحراف معیار تجربی آن ها (اعداد داخل پرانتز) در ۱۰۰ بار تکرار آزمایش برای برآورد پارامترهای پس از نقطه تغییر در مدل مؤلفه های مختصات وابسته محل اصابت با دو روش ماکسیم درستیابی و بیزی.

برآورد بیز			برآورد ML			اندازه نمونه
$\sigma_{12,2}$	$\sigma_{22,2}$	$\sigma_{11,2}$	$\sigma_{12,2}$	$\sigma_{22,2}$	$\sigma_{11,2}$	
۴۰۶۰۱(۲/۳۹۶۳)	۸۵۹۳۶۲(۱۴/۵۴۸۲)	۸۵۸۶۲۰(۱۵/۰۹۹۴)	۴۰۷۲۸(۲/۴۹۳۸)	۸۶۲۳۸۶(۱۵/۵۲۳۷)	۸۵۹۲۱۷(۱۷/۳۶۲۴)	$n = 50$
۴۴۹۷۲(۰/۶۳۰۹)	۸۶۴۶۰۳(۱۱/۰۳۹۶)	۸۷۷۵۳۵(۱۱/۸۳۰۵)	۴۲۹۰۲(۰/۷۹۷۵)	۸۷۴۳۰۸(۱۱/۰۴۳۲)	۸۷۶۳۴۸(۱۲/۵۷۴۸)	$n = 100$
۴۹۲۸۲(۰/۵۲۲۳)	۸۹۳۴۱۵(۸/۷۴۳۸)	۸۹۱۸۵۹(۹/۲۲۰۳)	۴۹۳۲۷(۰/۶۱۷۷)	۸۹۳۹۴۹(۹/۴۱۸۶)	۸۹۱۹۳۰(۱۰/۰۸۰۴)	$n = 200$

شکل ۷: مقادیر ماکسیمم لگاریتم درستنمایی در مقابل مقادیر τ . مقدار ماکسیمم با نقطه قرمز رنگ در $\tau = ۶۵$ مشخص شده است.

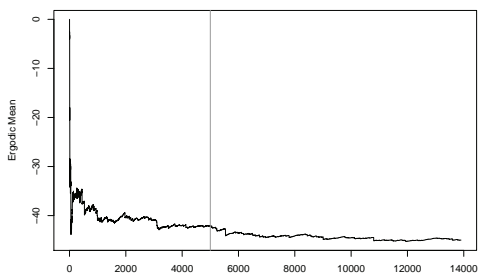
برای کم نکردن از کلیت مسئله، مدل (۴) را با توزیع خطاهای (۶) در نظر می‌گیریم. ابتدا برآوردهای ML حاصل از این مشاهدات را با استفاده از الگوریتم EM به ازای $\tau = ۳, \dots, ۹۸$ مختلف محاسبه کنیم و بررسی می‌کنیم که به ازای کدام مقدار τ مقدار تابع درستنمایی ماکسیمم می‌شود. نمودار لگاریتم درستنمایی به ازای مقادیر مختلف τ در شکل ۷ رسم شده است.

مقادیر حاصل، برای پارامترهای Σ_1 و Σ_2 به صورت

$$\hat{\Sigma}_1 = \begin{bmatrix} ۳۹۴۸۱۸۷ & -۱۱۶۶۲۵۲ \\ -۱۱۶۶۲۵۲ & ۲۰۱۵۷۳۱ \end{bmatrix},$$

$$\hat{\Sigma}_2 = \begin{bmatrix} ۳۶۵۲۶۸۹ & -۸۱۸۵۳۵ \\ -۸۱۸۵۳۵ & ۱۷۰۶۸۲۰ \end{bmatrix},$$

حاصل می‌شود. از همین مقادیر استفاده می‌کنیم و پیشین‌های Σ_1 و Σ_2 و $IW(\hat{\Sigma}_1, ۴)$ و $IW(\hat{\Sigma}_2, ۴)$ را به ترتیب برای ماتریس‌های Σ_1 و Σ_2 به کار می‌گیریم. همچنین نقطه آغازین الگوریتم برای مقدار τ را ۶۵ در نظر می‌گیریم. در هر مرحله از اجرای الگوریتم MCMC نمونه‌های مستقل از توزیع دومتغیره آمیخته نرمال تحدید شده را از $\pi_{z_i | X, \Sigma_1, k-1, \Sigma_2, k-1, \tau, k-1}(\cdot | x_i, \Sigma_1, k-1, \Sigma_2, k-1, \tau, k-1)$ تولید می‌کنیم و بر اساس آن نمونه‌های $\Sigma_{1,k}$ و $\Sigma_{2,k}$ تولید می‌شود.



شکل ۸: مقادیر شبیه‌سازی شده از مؤلفه (۱, ۲) ماتریس Σ_2 زمانی که $\tau = ۵۹$. مقدار داغیدن برابر با ۵۰۰۰ در نظر گرفته شده است.

با توجه به تعدد پارامترها، تنها نمودار میانگین تجمعی نمونه‌های تولید شده از مؤلفه $\sigma_{12,2}$ ماتریس Σ_2 را در شکل ۸ ترسیم کرده‌ایم. ۵۰۰۰ نمونه نخست، به عنوان نمونه‌های پیش از داغیدن کنار گذاشته شده‌اند. همچنین نمودار فروانی نسبی نمونه‌های تولید شده از پسین τ نیز در شکل ۹ آورده شده است. همان‌طور که از این شکل پیداست، مقدار τ برابر با ۵۹ برآورد شده است و بنابراین، ما

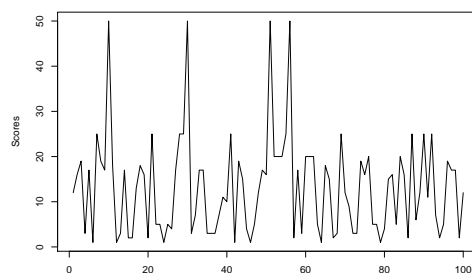
ادامه جدول ۲. برآورد نقطه تغییر با دو روش ماکسیمم درستنمایی و بیزی. مقادیر داخل پرانتز، مقدار انحراف معیار تجربی حاصل از هر برآورد را نشان می‌دهد.

اندازه نمونه	برآورد ML	برآورد بیزی
$n = ۵۰$	۴۲(۳۸۲۴۳)	۴۱(۲۸۶۲۴)
$n = ۱۰۰$	۴۰(۲۴۷۹۳)	۳۸(۱۷۵۹۷)
$n = ۲۰۰$	۳۸(۱۳۴۵۲)	۳۷(۰۸۷۷۴)

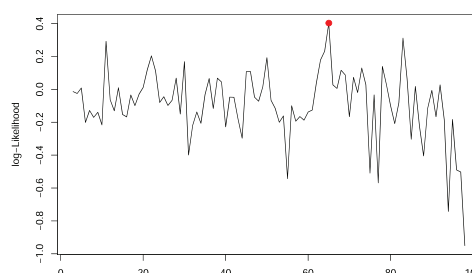
نکته قابل ملاحظه دیگری که در این مطالعه عددی می‌بایست مدنظر داشت آن است که هرچند در این مطالعه، برآوردگر بیز عملکرد بهتری از برآوردگر ML دارد، لیکن این عملکرد بهتر، یک هزینه محاسباتی نیز به واسطه استفاده از الگوریتم MCMC تحمیل می‌کند. هرچند که در عمل، این هزینه محاسباتی اضافی، چندان اهمیتی نمی‌یابد اما این موضوع می‌تواند به کارگیری آن را در فرایندهای پایش برخط، دشوارتر کند.

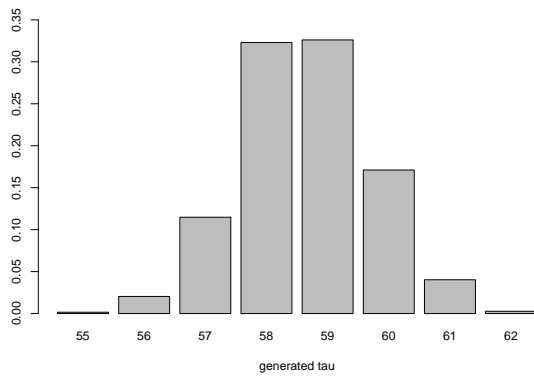
۲.۴ تحلیل داده‌های واقعی

برای جلوگیری از ارب شدن آزمایش، از داده‌های امتیازهای ثبت شده حاصل از ۱۰۰ پرتاب موجود در بسته darts از بسته نرم‌افزاری R استفاده می‌کنیم که در اثر نشانه‌گیری مرکز تخته نیزک ایجاد شده‌اند و امکان وجود نقطه تغییر در ۱۰۰ پرتاب پیاپی را بررسی می‌کنیم. این پرتاب‌ها در شکل ۶ نمایش داده شده‌اند.



شکل ۶: امتیازهای کسب شده در ۱۰۰ پرتاب پیاپی صورت گرفته در اثر نشانه گرفتن مرکز تخته نیزک.





تنها از نمونه‌های تولیدشده از توزیع‌های شرطی وقتی که $\tau = 59$ است استفاده می‌کنیم. تعداد این نمونه‌ها، ۳۲/۶۱٪ کل نمونه‌های تولیدشده را تشکیل می‌دهد. مقادیر برآورد بیزی پارامترها برای ماتریس‌های

$$\Sigma_1 \text{ و } \Sigma_2 \text{ به ترتیب عبارت‌اند از } \begin{bmatrix} 354,1453 & -80,0019 \\ -80,0019 & 188,6274 \end{bmatrix}$$

$$\cdot \begin{bmatrix} 371,0595 & -64,1798 \\ -64,1798 & 160,8655 \end{bmatrix}$$

شکل ۹: نمودار فراوانی نسبی مقادیر ۱۰۰۰۰ نمونه‌شده شبیه‌سازی شده از توزیع پسین τ .

References

- [1] Lehmann, E. L. and Casella, G., (1998), *Theory of Point Estimation*, 2nd ed. New York: Springer-Verlag.
- [2] Noorossana, R., Saghaei, A., and Amiri, A., (2011), *Statistical Analysis of Profile Ponitoring*, vol. 865, Wiley Series in Probability and Statistics, Wiley, New York.
- [3] Tibshirani, R. J., Price, A., and Taylor, J., (2011), A statistician plays darts, *J. Roy. Statist. Soc. Ser. A*, **174**, 213–226.

Bayesian profile monitoring in presence of latent variable: darts problem

Ali Reza Taheriyoun¹ and Ghazal Azadi²

Abstract:

Profile monitoring is usually faced by control charts and mostly the response variable is observable in those problems. We confront here with a similar problem where the values of the reward function are observed instead of the response variable vector and we use the dart model to make it easier to understand. Supposing there exists at most one change-point, a sequence of independent points resulted by darts throws is observed and the estimation of parameters and the change-point (if there exists any) are presented using the frequentist and Bayesian approaches. In both the approaches, two possible precision scalar and matrix are studied separately. The results are examined through a simulation study and the methods applied on a real data.

Keywords: change-point, EM algorithm, Gibbs sampler, latent variable, MCMC, mixture distribution.

¹Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

²Norwegian University of Life Sciences, Campus Ås