دانشگاه تربیت مدرس

دانشکده علوم ریاضی

پایان نامه دوره کارشناسی ارشد آمار

فرمول معکوس بیز برای تحلیل مدل‌های متغیر پنهان سلسله مراتبی

توسط

مسعود فریدی

استاد راهنما

دکتر مجید جعفری خالدی

بهمن ماه ۹۱

قدردانی

بدو گفت موبد که دانش به است که دانا به گیتی زهر کس مه است

سپاس بی‌کران یگانه هستی بخش را قدرت و فرصت درک دانش عطا فرمود تا بتوانم پس از طی مراحل مختلف کسبِ دانش و معرفت به تصنیف و نگارش این پایان‌نامه بپردازم.

نیست آب حیات جز دانش نیست باب نجات جز دانش

هر که این آب خورد باقی ماند چشم او در جمال ساقی ماند

در پی کشف این و آن رفتن جز به دانش کجا توان رفتن؟

استاد گرامی جناب آقای دکتر مجید جعفری خالدی بی تردید موفقیتم را مرهون زحمات بی شائبه شما می‌دانم که همچون چراغی دانش افروز در افق دیده‌گانم درخشیدید تا بتوانم راه را از بی‌راهه تمیز دهم.

مسعود فریدی

بهمن ۱۳۹۱

تقدیم به

آنان،

که

وجود،

کلامشان

خدمت ،

پدر و مادر.

و،

« »

است.

چکیده

در سال‌های اخیر استفاده از مدل‌های متغیر پنهان در زمینه‌های آماری مختلف مورد توجه قرار گرفته است. یک مدل متغیر پنهان دارای دو سطح است که در سطح اول آن توزیع مشاهدات به شرط متغیر پنهان و در سطح دوم توزیع متغیرهای پنهان مشخص می‌شود. یکی از رهیافت‌های تحلیل این گونه مدل‌ها، روش بیزی است که در آن با در نظر گرفتن توزیع پیشین برای پارامترهای مدل، یک سطح دیگر به دو سطح قبلی اضافه شده و بدین ترتیب یک مدل متغیر پنهان سلسله مراتبی شکل می‌گیرد. از آنجا که در این مدل‌ها، تابع درستنمایی به صورت انتگرال‌هایی پیچیده بر حسب متغیر پنهان است، تعیین تحلیلی توزیع پسین دشوار و بعضاً غیرممکن است. لذا اغلب سعی می‌شود با استفاده از روش‌های مونت کارلوی زنجیر مارکوفی، از توزیع پسین نمونه‌گیری شده و استنباط‌های بیزی مورد نظر انجام شوند. اما در این روش‌ها همگرایی زنجیر مساله‌ای اساسی به شمار می‌رود. با توجه با این مساله، در این پایان‌نامه روشی نامکرر بر اساس فرمول معکوس بیز معرفی می‌شود که با استفاده از آن می‌توان نمونه‌هایی مستقل از توزیع پسین پارامترهای یک مدل متغیر پنهان سلسله مراتبی شبیه‌سازی نمود. به عنوان یک فعالیت پژوهشی جدید، این روش برای استنباط بیزی یکی از مهمترین و پرکاربردترین مدل‌های متغیر پنهان، تحت عنوان مدل عاملی ارائه می‌شود. سپس عملکرد آن با استفاده از شبیه‌سازی مورد ارزیابی قرار گرفته و کاربست آن در تحلیل یک مجموعه داده واقعی نشان داده می‌شود.

واژه‌های کلیدی : مدل متغیر پنهان، بیز سلسله مراتبی، تحلیل بیزی عاملی، روش‌های مونت کارلوی زنجیر مارکوفی، فرمول معکوس بیز.

فهرست مندرجات

فصل ۱

مقدمات و مفاهیم مورد نیاز

۱.۱ مقدمه

مدل‌های متغیر پنهان[۱] مدل‌هایی هستند که دارای یک یا چند متغیر مشاهده نشده، گمشده یا پنهان‌اند. متغیر پنهان منبعی است که موجب تغییرات غیر قابل مشاهده بر متغیر پاسخ یا پارامترهای مدل می‌شود. در سال‌های اخیر استفاده از مدل‌های متغیر پنهان در زمینه‌های آماری مختلف مورد توجه قرار گرفته است (بارتلموف و همکاران، ۲۰۱۱ ؛ اسکراندل و رابه هسکه، ۲۰۱۲). یک مدل متغیر پنهان دارای دو سطح است که در سطح اول آن توزیع مشاهدات به شرط متغیر پنهان و در سطح دوم توزیع متغیرهای پنهان مشخص می‌شود. به دلیل وجود متغیر پنهان در این مدل‌ها، تابع درستنمایی به صورت انتگرال‌هایی بر حسب متغیر پنهان است. از این رو برآورد پارامترهای این گونه مدل‌ها مستلزم بکارگیری راه‌حل‌هایی است. یکی از رهیافت‌های مهم روش بیزی است که در آن با در نظر گرفتن توزیع پیشین برای پارامترهای مدل، یک سطح دیگر به دو سطح قبلی اضافه شده و بدین ترتیب یک

[۱] Latent Variable Models

مدل متغیر پنهان سلسله مراتبی[۲] شکل می‌گیرد. در تحلیل این حالت یکی از اساسی‌ترین موضوعات یافتن توزیع پسین بر اساس تابع درستنمایی و توزیع پیشین است. به دلیل انتگرال‌های پیچیده موجود در تابع درستنمایی، تعیین تحلیلی توزیع پسین دشوار و بعضاً غیرممکن است.

یکی از روش‌های معمول برای حل این مساله، شبیه‌سازی از توزیع پسین پارامترها با استفاده از روش‌های مونت کارلوی زنجیر مارکوفی[۳] (MCMC) است. بدین ترتیب برآورد بیزی پارامترهای مدل بر اساس نمونه‌های حاصل از توزیع پسین صورت می‌گیرد. از جمله مهمترین این روش‌ها می‌توان به الگوریتم‌های گیبس[۴] و متروپولیس – هستینگز[۵] اشاره نمود.

به طور کلی در روش‌های MCMC، بر اساس یک الگوریتم مکرر[۶]، زنجیری تقلیل ناپذیر[۷] و نادوره‌ای[۸] ساخته شده و دنباله‌ای از پارامترهامورد نظر تولید می‌شود به گونه‌ای که توزیع مانای آن همان توزیع پسین است. در واقع با تکرار متوالی مراحل الگوریتم یک نمونه وابسته مارکوفی از توزیع پسین تولید می‌شود. دنباله آغازین تکرارهای زنجیر مارکوف قبل از زمان همگرایی، دوره داغیدن نامیده می‌شود. مشاهدات بعد از این دوره را می‌توان به عنوان مشاهداتی از توزیع پسین دانست. در روش‌های $MCMC$ همگرایی زنجیر مساله‌ای اساسی به شمار می‌رود. در مدل‌های پیچیده از جمله مدل‌های متغیر پنهان سلسله مراتبی، معمولا دوره داغیدن برای همگرایی کلیه پارامترها طولانی و زمان همگرایی بسیار کند است. در نتیجه زمان انجام محاسبات به نحوی چشمگیر افزایش می‌یابد. ضمن آنکه تعیین زمان مشترکی که کلیه پارامترهای مدل در آن به همگرایی رسیده باشند، از دیگر مشکلات

---

[۲] Hierarchical Latent Variable Model

[۳] Markov Chain Monte Carlo Methods

[۴] Gibbs

[۵] Metropolis - Hastings

[۶] Iterative

[۷] Irreducibility

[۸] Aperiodic

این روش‌هاست. به علاوه، در الگوریتم متروپولیس – هستینگز لازم است یک تابع نامزد مناسب اختیار شود که خود مساله‌ای اساسی به شمار می‌رود.

با توجه به مسائل فوق، تان و همکاران (۲۰۰۳) روشی نامکرر[۹] بر اساس فرمول معکوس بیز[۱۰] (IBF) پیشنهاد دادند که با استفاده از آن می‌توان نمونه‌هایی مستقل از توزیع پسین پارامترهای یک مدل سلسله مراتبی شبیه‌سازی نمود. مبنای روش به این صورت است که در چارچوب فرمول معکوس بیز، توزیع پسین متغیرهای پنهان متناسب با نسبت توزیع شرطی کامل متغیرهای پنهان به توزیع شرطی کامل پارامترهای مدل نوشته می‌شود. سپس با استفاده از روش نامکرر نمونه‌گیری بازنمونه‌گیری مهم[۱۱] (SIR) از توزیع پسین متغیرهای پنهان، نمونه‌های مستقل شبیه‌سازی می‌شود. لازم به ذکر است که در این الگوریتم، توزیع شرطی کامل متغیرهای پنهان به شرط مد پسین پارامترها و مشاهدات، به عنوان توزیع مهم[۱۲] انتخاب می‌شود. برای محاسبه مد پسین پارامترها نیز الگوریتم امیدگیری – ماکسیمم‌سازی[۱۳] (EM) یا الگوریتم مونت کارلوی EM[۱۴] ($MCEM$) استفاده می‌شود. در انتها با جایگذاری نمونه‌های تولید شده در توزیع شرطی کامل پارامترها به شرط متغیرهای پنهان و مشاهدات، و نمونه‌گیری از آن، نمونه‌های مستقل از توزیع پسین حاصل می‌گردد.

در سال‌های اخیر روش IBF به دلیل ویژگی‌های مناسب آن بسیار مورد توجه قرار گرفته است. از جمله تان و همکاران (۲۰۰۶) این روش را برای مدل‌های سلسله مراتبی در داده‌های زوجی صفر و یک توسعه دادند. بعلاوه تان و همکاران (۲۰۰۷) و لاچاس و همکاران (۲۰۱۲) نیز با استفاده از

---

[۹] Non-Iterative

[۱۰] Inverse Bayes Formula

[۱۱] Sampling Importance Resampling

[۱۲] Importance Distribution

[۱۳] Expectation - Maximization

[۱۴] Monte Carlo Expectation Maximization

IBF به تحلیل بیزی مدل‌های آمیخته خطی پرداختند. تیان و همکاران (۲۰۰۷) نیز از روش نامکرر مبتنی بر IBF، برای استنباط بیزی مدل‌های متغیر پنهان هنگامی که داده گمشده وجود دارد استفاده نمودند. یانگ و همکاران (۲۰۰۸) دو الگوریتم غیر مکرر بر اساس IBF معرفی کردند. تیان و همکاران (۲۰۰۹) فرمول معکوس بیز را برای مدل‌های نقطه تغییر پواسن[۱۵] پیشنهاد دادند. برای مدل‌های آمیخته خطی چندمتغیره با خطاهای دارای توزیع تی استودنت، وانگ و فان (۲۰۱۲) از الگوریتم هیبرید[۱۶] IBF و گیبس استفاده نمودند. همچنین یو و تیان (۲۰۱۱) بر اساس روش IBF یک الگوریتم برای نمونه‌گیری از توزیع نرمال بریده شده معرفی کردند. آن و بنتلر (۲۰۱۲) بر مبنای مدل‌های متغیر پنهان، به تحلیل داده‌های صفر و یک بر اساس روش IBF پرداختند.

در این پایان‌نامه، فرمول معکوس بیز معرفی شده و نحوه استفاده از آن در تولید نمونه‌های مستقل از توزیع پسین پارامترهای یک مدل متغیر پنهان سلسله مراتبی بیان می‌شود. در ادامه به عنوان یک فعالیت پژوهشی جدید، برای یکی از مهمترین و پرکاربردترین مدل‌های متغیر پنهان تحت عنوان مدل عاملی، استنباط بیزی بر اساس فرمول معکوس بیز ارائه می‌شود. سپس با استفاده از شبیه‌سازی، عملکرد آن مورد ارزیابی قرار گرفته و با روش‌های $MCMC$ مقایسه می‌شود. در پایان کاربست روش بر مبنای تحلیل یک مجموعه داده واقعی نشان داده می‌شود. برای انجام این اهداف فصول پایان‌نامه به صورت زیر تدوین گردیده است. در ادامه این فصل برخی مفاهیم آمار بیزی و نحوه استفاده از روش‌های مکرر و نامکرر مونت کارلویی در شبیه‌سازی از توزیع پسین مدل‌های متغیر پنهان سلسله مراتبی بیان می‌شود. در فصل دوم فرمول معکوس بیز و نحوه استفاده از آن در شبیه‌سازی از توزیع پسین مدل‌های متغیر پنهان سلسله مراتبی بیان می‌شود. فصل سوم شامل تعمیم این روش برای مدل‌های عاملی است. در فصل چهارم با استفاده از مثال‌های شبیه‌سازی و کاربردی ویژگی‌های

[۱۵] Poisson Changepoints

[۱۶] Hybrid

نمونه‌گیری از پسین بر مبنای فرمول معکوس بیز را مورد بررسی و تحقیق قرار می‌دهیم.

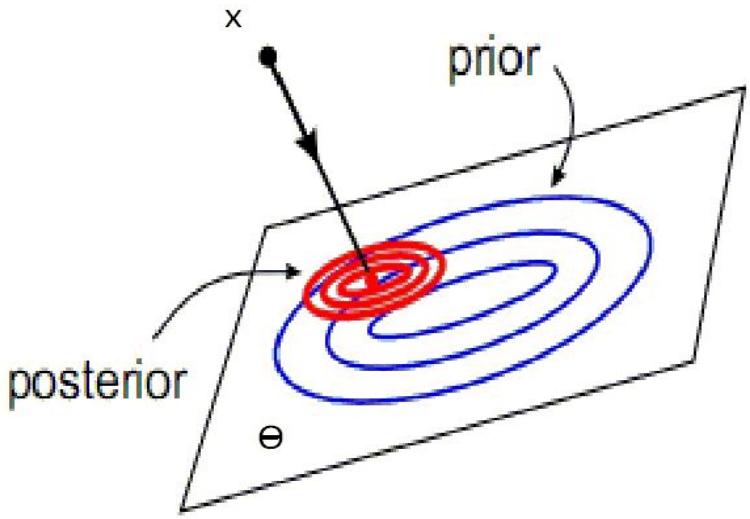
۲.۱ استنباط بیزی

فرض کنید $X_1, ..., X_n$ یک نمونه تصادفی از توزیع معلوم $p(x|\theta)$ باشد، که در آن $\theta \in \Theta$ پارامتری نامعلوم است. در استنباط بیزی پیش از استفاده از اطلاعات داده‌ها، بر اساس اطلاعات قبلی یا باور شخصی، محتمل‌ترین مقدار θ در Θ تعیین می‌شود. در این صورت θ به عنوان متغیری تصادفی با توزیع پیشین $\pi(\theta)$ در نظر گرفته می‌شود. اطلاعات پیشین با ترکیب اطلاعات داده‌ها در مورد θ به هنگام می‌شود. با ترکیب توزیع پیشین و تابع درستنمایی داده‌ها، توزیع پسین

$$\pi(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{\int p(x|\theta)\pi(\theta)d\theta} \qquad (۱.۲.۱)$$

حاصل می‌شود و بدین ترتیب می‌توان اطلاعات قبلی و اطلاعات نهفته در داده‌ها در مورد θ را با هم ترکیب کرد و به واقعیت نزدیک کرد. در شکل ۱.۲.۱، به طور شماتیک توزیع پیشین با اطلاعات نهفته در مشاهدات ترکیب شده و توزیع پسین حاصل به نمایش گذاشته شده است. همان گونه که از این شکل مشاهده می‌شود فضایی از پارامتر که محتمل‌ترین مقدار در آن اتفاق می‌افتد به نحوی چشمگیر کاهش یافته است. بدین ترتیب در صورت مناسب بودن توزیع پیشین اختیار شده، به شکل دقیق‌تری می‌توان پارامترهای مدل را برآورد کرد. هر اندازه اطلاعات قبلی مناسب‌تری در اختیار باشد، تأثیر توزیع پیشین در توزیع پسین افزایش می‌یابد یا به تعبیر دیگر میزان آگاهی بخشی[۱۷] آن افزایش می‌یابد. همچنان که در شکل ۱.۲.۱، خصوصیات متقارن بودن و زنگوله‌ای بودن شکل چگالی پیشین به چگالی پسین منتقل شده است که نشان می‌دهد تأثیر توزیع پیشین در پسین زیاد بوده است. با

[۱۷] Informative



شکل ۱.۲.۱: نمودار کانتور توزیع پسینی حاصل از اطلاعات پیشین و داده‌ها

این وجود، هرگاه اطلاعات قبلی در اختیار نبوده یا این اطلاعات مبهم باشند، در تحلیل‌های بیزی، توزیع‌های پیشین ناآگاهی بخش[۱۸] یا مبهم[۱۹] مورد استفاده قرار می‌گیرند. مثالی در این مورد حالتی است که یک مشاهده دارای توزیع برنولی با پارامتر θ باشد و توزیع پیشین θ توزیع یکنواخت بر بازه $(۰, ۱)$ اختیار شود که در این حالت توزیع پیشین تأثیری در توزیع پسینی پارامترها نخواهد داشت. توزیع‌های پیشین ناآگاهی بخش اغلب ناسره[۲۰] هستند، یعنی

$$\int \pi(\theta)d\theta = \infty.$$

از آنجا که پیشین ناسره توزیع‌های احتمال نیستند، لزوماً توزیع‌های پسین حاصل از آنها سره نخواهند بود. لذا هنگام استفاده از پیشین‌های ناسره باید توجه شود که تنها در صورت سره بودن توزیع پسین،

---

[۱۸] Non-Informative

[۱۹] Vague

[۲۰] Improper

استنباط بیزی میسر می‌شود. عموماً در مسائل بیزی محاسبه‌ی امید پسین تابعی از پارامتر، یعنی

$$\begin{aligned} E[g(\theta)|x] &= \int_{\Theta} g(\theta)\pi(\theta|x)d\theta \\ &= \frac{\int_{\Theta} g(\theta)p(x|\theta)\pi(\theta)d\theta}{\int_{\Theta} p(x|\theta)\pi(\theta)d\theta} \end{aligned} \qquad (۲.۲.۱)$$

مورد نظر است. به علت چند بعدی بودن پارامترها تعیین انتگرال‌ها در روابط (۱.۲.۱) و (۲.۲.۱) بسیار دشوار می‌باشد و نمی‌توان آنها را به صورت بسته محاسبه کرد. اگر خانواده‌ای از توزیع‌های پیشین به گونه‌ای باشد که توزیع پسین حاصل از آنها نیز متعلق به همان خانواده باشد، این خانواده از توزیع‌ها را خانواده توزیع‌های مزدوج برای تابع درستنمایی می‌نامند. انتخاب توزیع پیشین از خانواده توزیع‌های مزدوج، محاسبات مربوط به توزیع پسین را راحت‌تر می‌کند. ضمن آنکه اگر به مناسب بودن پیشین اختیار شده اعتقاد وجود داشته باشد توزیع پسین و پیشین می‌بایست از یک خانواده باشند. با این وجود در بعضی مسائل به علت چند بعدی بودن پارامتر θ کماکان محاسبه انتگرال‌هایی دشوار مورد نیاز است. در این موارد می‌توان روش‌های عددی مونت کارلو[۲۱] (MC) را به کار برد که بر مبنای تولید نمونه تصادفی برای حل انتگرال‌های پیچیده هستند.

۳.۱ روش‌های مونت کارلو

مونت کارلو نام محلی است که در آن بازی‌های شانسی انجام می‌شود و به همین دلیل بسیار شناخته شده است. شکل ۲.۳.۱ تصویری از این محل را نمایش می‌دهد. منطقه مونت کارلو ثروتمندترین منطقه در ناحیه خودمختار موناکو در کشور فرانسه است که تا پیش از سال ۱۸۸۶ دارای زبان رسمی ایتالیایی بود اما در یکصد سال گذشته به زبان فرانسوی تغییر یافت. ریشه نام مونت کارلو از زبان

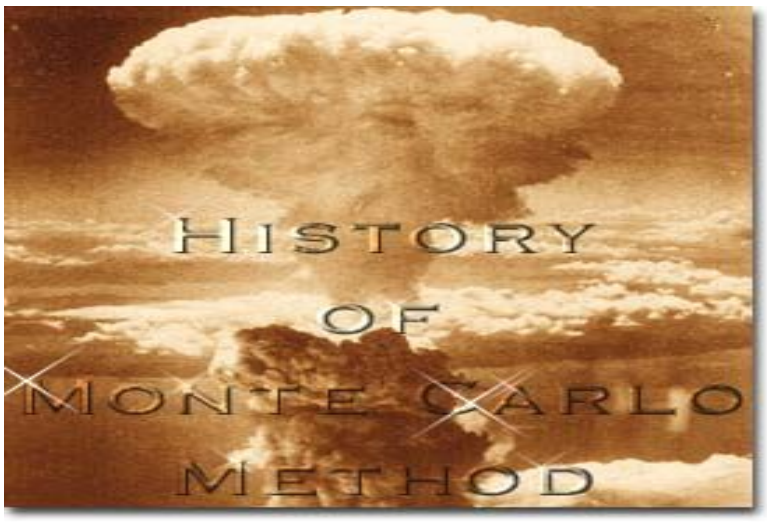
[۲۱] Monte Carlo

شکل ۲.۳.۱: تصویری از مونت کارلو

ایتالیایی است و به اصلیت شاهزاده کارلوی سوم از موناکو بر می‌گردد. روش‌های مونت کارلو یک کلاس از روش‌های عددی است که از اعداد تصادفی برای محاسبه نتایج استفاده می‌کند. روش‌های مونت کارلو معمولا برای شبیه سازی سیستم‌های فیزیکی، ریاضی، اقتصادی و ... است. در ریاضیات و آمار این روش کاربردهای بسیاری دارد از جمله مهمترین این کاربردها، حل انتگرال‌های معین، حل مسائل بهینه سازی، محاسبات ریاضیاتی (مانند تقریب عدد π)، حل دستگاهای معادلاتی پیچیده و ... است.

واژه مونت کارلو در دهه ۱۹۴۰ به وسیله فیزیکدانانی مانند جان وان نیومن، استانیسلاو اولام، نیکولاس متروپولیس که روی پروژه ساخت یک سلاح اتمی (پروژه منهتن) در آزمایشگاه ملی لوس آلاموس آمریکا کار می‌کردند رایج شد که شکل ۳.۳.۱ به این نکته اشاره دارد.

روش‌های مونت کارلو به صورت یک الگوریتم واحد نیست و شامل طیف وسیعی از روش‌ها است. اما



شکل ۳.۳.۱: تاریخچه روش‌های مونت کارلو

بیشتر این روش‌ها مراحل زیر را دارند:

۱) مجموعه‌ای از ورودی‌های ممکن تعریف می‌شود (تکیه‌گاه متغیر تصادفی تعیین می‌شود).

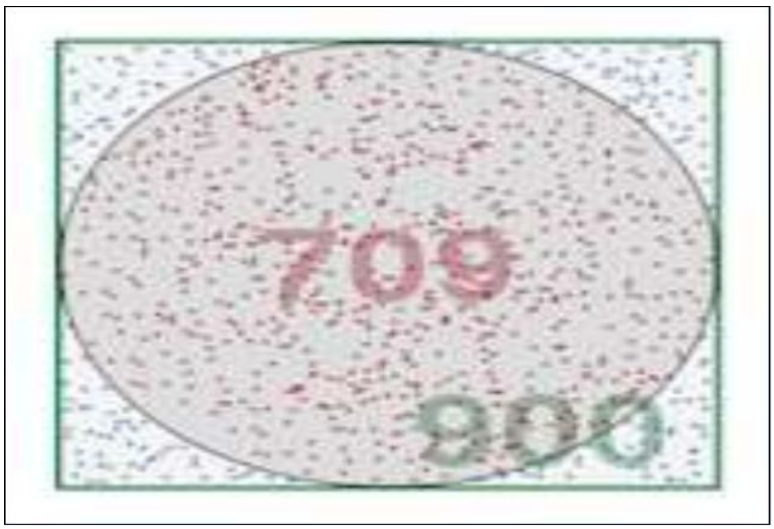
۲) از ورودی‌های ممکن، ورودی‌های تصادفی تولید می‌شود (با توجه به تکیه‌گاه متغیر تصادفی، نمونه تصادفی از متغیر تصادفی تولید می‌شود).

۳) با توجه به ورودی‌های تصادفی، محاسبات مورد نظر انجام می‌شود.

۴) نتایج هر یک از اجراهای محاسباتی ثبت شده، و بر این اساس پاسخ نهایی ارائه می‌گردد.

برای جزئیات بیشتر به ذکر مثالی می‌پردازیم که در آن از روش مونت کارلو برای تقریب عدد π استفاده می‌شود. دایره‌ای را متصور شوید که با یک مربع به اضلاع r، محاط شده باشد. می‌دانیم نسبت مساحت دایره به مربع

$$\frac{\pi(r/۲)^۲}{r^۲} = \frac{\pi}{۴}$$



شکل ۴.۳.۱: تقریب مونت کارلویی عدد π

است. حال برای تولید اعداد تصادفی از داخل مربع، از توزیع توام

$$p(x,y) = \frac{۱}{۴} I_{[-۱,۱]}(x) I_{[-۱,۱]}(y) \qquad (۳.۳.۱)$$

نمونه‌های $(x_۱,y_۱),(x_۲,y_۲)...,(x_n,y_n)$ را تولید می‌کنیم. می‌دانیم در صورت بزرگ بودن n نسبت نقاطی که داخل دایره واقع می‌شوند نزدیک $\frac{\pi}{۴}$ خواهد بود. اکنون یک تقریب از عدد π را می‌توان ۴ برابر نسبت اعداد داخل دایره به کل اعداد تولید شده دانست. شکل ۴.۳.۱ حاصل تولید ۹۰۰ نمونه تصادفی است که ۷۰۹ تا از نقاط نمونه‌ای در داخل دایره واقع شده‌اند. در این صورت با توجه به مطالب ذکر شده یک تقریب از π به صورت

$$۴ \times \frac{۷۰۹}{۹۰۰} = ۳/۵۱$$

حاصل می‌گردد. در ادامه بیشتر توجه بر روش‌های MC در تقریب انتگرال‌های پیچیده متمرکز است. فرض کنید محاسبه مقدار انتگرال $I = \int h(x)p(x)dx$ مورد نظر است که در آن $p(x)$ یک تابع چگالی است. اگر بتوان نمونه‌های $x_۱,x_۲,...,x_n$ را از توزیع $p(x)$ شبیه‌سازی کرد، تقریب مونت کارلوی

انتگرال I به صورت $\hat{I} = \frac{۱}{n}\sum_{i=۱}^{n} h(x_i)$ خواهد بود. در صورتی که این نمونه‌ها مستقل باشند بر اساس قانون قوی اعداد بزرگ وقتی $n \longrightarrow \infty$ ، با احتمال یک

$$\hat{I} = \frac{۱}{n}\sum_{i=۱}^{n} h(x_i) \longrightarrow I \qquad (۴.۳.۱)$$

از این رو با افزایش مقدار n، تقریب دقیق‌تر خواهد بود. بر اساس واریانس نمونه‌ای $h(x_i)$ ها و برای هر مقدار ثابت n برآورد دقت تقریب (۴.۳.۱) به صورت

$$\hat{Var}(\hat{I}) = \frac{۱}{n(n-۱)}\sum_{i=۱}^{n}(h(x_i) - \hat{I})^۲$$

خواهد شد. همچنین بر اساس قضیه حد مرکزی یک بازه اطمینان $۱۰۰(۱-\alpha)\%$ برای I نیز به صورت

$$\hat{I} \pm z_{\frac{\alpha}{۲}}\sqrt{\hat{Var}(\hat{I})}$$

خواهد بود. به عنوان مثال، برآورد تابع

$$p = P(a < x < b) = E[I_{(a,b)}(x)] = \int I_{(a,b)}(x)p(x)dx$$

به صورت

$$\hat{p} = \frac{\sum_{i=۱}^{n} I_{(a,b)}(x_i)}{n}$$

با واریانس $\hat{Var}(\hat{p}) \approx \frac{\hat{p}(۱-\hat{p})}{n}$ است.

روش‌های مونت کارلو به دو دسته مکرر و نامکرر تقسیم می‌شوند. از روش‌های مونت کارلوی نامکرر می‌توان نمونه‌گیری مهم[۲۲] (IS) و SIR را نام برد. مهم‌ترین روش‌های مکرر، روش‌های مونت کارلوی

[۲۲] Importance Sampling

زنجیر مارکوفی، مانند الگوریتم گیبس و نمونه‌گیری متروپولیس – هستینگز است. در ادامه به معرفی مختصر این روش‌ها می‌پردازیم.

۱.۳.۱ روش‌های نمونه‌گیری نامکرر

روش‌های IS و SIR از روش‌های نامکرر مونت کارلویی هستند که به صورت گسترده مورد استفاده قرار می‌گیرند. حتی در ساختار برخی از روش‌های مونت کارلوی مکرر نیز به کار می‌روند. فرض اساسی در این الگوریتم‌ها این است که نمونه‌گیری از توزیع هدف امکان پذیر نیست، اما می‌توان توزیعی نزدیک به آن یافت که نمونه‌گیری از آن ساده باشد.

روش IS

هنگامی که نمونه‌گیری مستقیم از توزیع $p(x)$ امکان پذیر نباشد، اما بتوان توزیعی مانند $g(x)$ (که به آن توزیع مهم گویند) را یافت به گونه‌ای که

۱) تکیه‌گاه $g(x)$ با $p(x)$ یکی باشد،

۲) نمونه‌گیری از آن امکان پذیر باشد،

۳) به توزیع $p(x)$ نزدیک باشد،

نمونه‌گیری IS مورد استفاده قرار می‌گیرد. چون

$$\begin{aligned} I &= E_p(h(X)) \\ &= \int h(x)p(x)dx \end{aligned}$$

$$
\begin{aligned}
&= \int h(x)p(x)\frac{g(x)}{g(x)}dx \\
&= \int \frac{h(x)p(x)}{g(x)}g(x)dx \\
&= E_g(\frac{h(X)p(X)}{g(X)}) \\
&= E_g(h(X)\omega(X)) \qquad (۵.۳.۱)
\end{aligned}
$$

که در آن $\omega(x) = \frac{p(x)}{g(x)}$، انتگرال I که امید تابعی از متغیر X با تابع چگالی $p(x)$ است به امید تابعی از متغیر X با تابع چگالی $g(x)$ به صورت (۵.۳.۱) تبدیل می‌شود. اکنون می‌توان نمونه‌های $x_۱, x_۲, ..., x_n$ را از توزیع $g(x)$ تولید نموده و بر اساس روش مونت کارلو به سادگی برآوردی برای I به صورت

$$\hat{I} = \frac{۱}{n}\sum_{i=۱}^{n} h(x_i)\omega(x_i)$$

معرفی کرد.

به عنوان کاربردی از IS فرض کنید که $x_۱, x_۲, ..., x_n$ مشاهداتی از توزیع $p(x|\theta)$ باشد که در آن $\theta \in \Theta$ دارای توزیع پیشین $\pi(\theta)$ است و بخواهیم $E[g(\theta)|x]$، امید تابعی از توزیع پسین پارامترها یعنی

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

را محاسبه کنیم و فرض کنیم که این انتگرال به روش‌های معمول قابل محاسبه نیست. در این صورت بر مبنای روش مونت کارلو می‌باید از توزیع پسین پارامترها $\pi(\theta|x)$ نمونه‌های $\theta_۱, \theta_۲, ..., \theta_m$ را تولید کرده و سپس $\frac{۱}{m}\sum_{i=۱}^{m} g(\theta_i)$ برآوردی از $E(g(\theta)|x)$ است. اما در بسیاری از مواقع، به دلیل چند بعدی بودن فضای پارامترها تعیین توزیع پسین مشکل است یا شکل استانداردی ندارد که بتوان به