دانشگاه پیام نور

گروه آمار

پایان‌نامه جهت اخذ درجه‌ی کارشناسی ارشد

رشته‌ی آمار

روش‌های محاسباتی برای انتخاب مدل بیزی

استاد راهنما:

دکتر علی شادرخ

استاد مشاور:

دکتر مسعود یارمحمدی

ساناز پناهی مرادکندی

خرداد ۱۳۹۱

بسم الله الرحمن الرحيم

شماره: ....
تاریخ: ....
پیوست: ....

دانشگاه پیام نور
دانشگاه پیام نور استان تهران
اللهم عجل لولیک الفرج و العافیة و النصر

جمهوری اسلامی ایران
وزارت علوم تحقیقات و فناوری
مرکز تهران شرق

صورتجلسه دفاع از پایان نامه کارشناسی ارشد

جلسه دفاع ازپایان نامه کارشناسی ارشد خانم ساناز پناهی مرادکندی
دانشجوی رشته آمار ریاضی به شماره دانشجویی: ۸۹۰۰۶۶۶۴۳
تحت عنوان:

روش های محاسباتی برای انتخاب مدل بیزی

جلسه دفاع با حضور داوران نامبرده ذیل در روز شنبه مورخ: ۹۱/۰۳/۲۷ ساعت : ۱۲-۱۱ درمحل تهران شرق برگزار شد. و پس از بررسی پایان‌نامه مذکور با نمره به عدد ۱۸/۷۵ به حروف هیجده و هفتاد و پنج صدم با درجه ارزشیابی عالی مورد قبول واقع شد ☑ نشد ☐

ردیف	نام و نام خانوادگی	هیات داوران	مرتبه دانشگاهی	دانشگاه/ موسسه	امضاء
۱	دکترعلي شادرخ	استاد راهنما	استادیار	پیام نور	
۲	دکترمسعود یارمحمدي	استاد مشاور	دانشیار	پیام نور	
۳	دکترصادق رضایي	استاد داور	دانشیار	دانشگاه صنعتي امیرکبیر	
۴	دکترپرویز نصیري	نماینده علمي گروه/ نماینده تحصیلات تکمیلی	دانشیار	پیام نور	

تهران ، خیابان کریمخان زند ،خیابان استاد نجات الهی ، خیابان شهید فلاح پور ، پلاک ۲۷ مرکز تهران شرق
تلفن : ۸۸۹۱۳۴۷۵
دورنگار: ۸۸۹۴۸۹۸۴
Tshargh.Tpnu.ac.ir
Tshargh@Tpnu.ac.ir

منّت خدای را عزّ و جل ...

درود و سپاس محضر استاد راهنمایم جناب آقای دکتر علی شادرخ که با یاری و همکاری ایشان این وجیزه‌ی ناچیز پای به این حال و مقام نمی‌نهاد و تقدیر و تشکری تام و تمام، خدمت استاد گرامی جناب آقای دکتر مسعود یارمحمدی که زحمت مشاورت این رساله را تقبّل فرمودند؛ و سپاسی صمیمانه محضر استادان ارجمند آقای دکتر صادق رضایی و جناب آقای دکتر پرویز نصیری که مسئولیت داوری این پایان نامه را پذیرا شدند.

تقدیم به پدرم کز او صبر و مادرم کز او مهر آموختم

و تقدیم به خواهران مهربانم

و همه ی یاران و دوستانی که در طی این طریق و حصول و وصول به این جایگاهی که در آنم، مرا هم یاری نمودند.

چکیده

فرآیند استنباط آماری از دو رویکرد کلاسیک و بیزی قابل بررسی است. صرف نظر از نقاط قوت و ضعف هر یک از این رویکردها در مسائل با ابعاد کوچک، رویکرد بیزی یک چارچوب بسیار قدرتمند برای تحلیل و تجزیه‌ی نتایج آزمایش‌های علمی با ابعاد بزرگ فراهم می‌کند. در این رویکرد بسیاری از کمیت‌های مورد علاقه به صورت انتگرال ظاهر می‌شوند که راه حل تحلیلی ندارند. یک رویکرد طبیعی برای حل این مشکلات محاسباتی، استفاده از تکنیک‌های مونت کارلو و روش‌های عددی است که در استنباط بیزی، روش‌های مونت کارلو بر روش‌های عددی ترجیح داده می‌شوند.

در میان این کمیت‌های مورد علاقه، عامل بیز به عنوان یک معیار، نقش کلیدی در آزمون فرض‌ها و انتخاب مدل ایفا می‌کند. در رابطه با آزمون فرض‌ها، عامل بیز می‌تواند به عنوان جانشین برای آزمون‌های معنی دار بیزی و کلاسیک تلقی گردد. هر چند عامل بیز برای مقایسه دو مدل است ولی در عمل می‌توان آن را برای مقایسه چند مدل رقیب نیز به کار برد.

در این پایان نامه ابتدا مروری بر رویکرد استنباط بیزی داشته و ضمن بیان مشکلات محاسباتی آن، به روش‌های تقریب مونت کارلو پرداخته و به خصوص عامل بیز را بررسی خواهیم نمود.

واژگان کلیدی

استنباط بیزی، توزیع پیشین، توزیع پسین، روش‌های مونت کارلو، نمونه‌گیری نقاط مهم، عامل بیز.

پیشگفتار

عامل بیز معیاری در استنباط آماری می‌باشد که برای آزمون فرض‌ها و انتخاب مدل مورد استفاده قرار می‌گیرد. این معیار اولین بار توسط جفریز در سال **۱۹۳۹** بیان شد و پس از آن مولفین متعددی آن را با پیشین‌های مختلف به کار بردند. هدف اصلی ما نیز، انتخاب مدل مناسب بیزی با استفاده از این معیار می‌باشد.

این پایان‌نامه از سه فصل تشکیل شده است.

فصل اول به کلیات و مفاهیم کلیدی که در فصل‌های آینده نیز مورد استفاده قرار می‌گیرند تعلق دارد. در این فصل به معرفی استنباط بیزی، توزیع‌های پیشین، انواع آنها و روش‌های انتخاب پیشین مناسب پرداخته و سپس کمیت‌های پسین را محاسبه می‌کنیم.

به دست آوردن مقدار کمیت‌های پسین (کمیت‌هایی که براساس توزیع پسین قابل محاسبه هستند) ممکن است بسیار سخت و یا حتی غیرقابل محاسبه باشند. یک رویکرد طبیعی برای حل این مشکل محاسباتی، استفاده از تکنیک‌های مونت کارلو و روش‌های عددی است. **فصل دوم** را به معرفی این روش‌ها اختصاص می‌دهیم.

در **فصل سوم** که فصل پایانی نیز می‌باشد، به بررسی آزمون فرض بیزی، عامل بیز و انتخاب مدل مناسب با استفاده از عامل بیز می‌پردازیم.

فهرست مطالب

لیست جداول

لیست تصاویر

فصل ۱

استنباط بیزی

۱.۱ مقدمه

فرض کنید یک مجموعه از داده‌ها در اختیار داریم که متأثر از یک سری تغییرات تصادفی می‌باشند؛ به مجموعه روش‌های آماری که برای به دست آوردن نتیجه گیری در خصوص یک ویژگی از این داده‌ها می‌توانند به کار روند استنباط آماری[۱] می‌گویند.
برای اینکه این نتیجه‌گیری‌ها معتبر بوده و روش‌های مورد استفاده بتوانند برای دسته‌ی وسیعی از داده‌ها به کار روند، لازم است ساختاری برای توصیف داده‌ها در نظر بگیریم.
فرض کنید $X_1, \ldots, X_n$ یک دسته داده‌ای باشند که می‌خواهیم در مورد آن‌ها تحلیلی انجام دهیم. ساختار مورد نظر، برای مثال، می‌تواند به این صورت باشد:

$$X_i = \mu + \varepsilon_i$$

که در آن μ جز ثابت داده‌ها و ϵ_iها جز تصادفی می‌باشند و از یک توزیع احتمال خاصی تبعیت می‌کند. به چنین ساختارهایی مدل[۲] می‌گوییم. در صورتی که فرض کنیم مدل، به صورت کامل معلوم باشد (پارامترهای مدل معلوم باشد) و مقدار مشاهدات نامعلوم، چنین ساختاری را مدل احتمالی[۳] می‌گوییم. در صورتی که داده‌ها معلوم باشند و توزیع داده‌ها نامعلوم، آن را مدل آماری[۴] می‌گویند. نامعلوم بودن توزیع داده‌ها می‌توانند به دو صورت در نظر گرفته شود:

۱. ساختار کلی توزیع داده‌ها معلوم نباشد

۲. پارامترهای توزیع داده‌ها معلوم نباشند.

روش‌های استنباط آماری می‌توانند با استفاده از این مدل، نتیجه‌ی معتبری در مورد داده‌ها (و همچنین جامعه‌ای که داده‌ها از آن آمده‌اند) به دست دهند. از آنجایی که طبیعت این نوع نتیجه‌گیری (یا استنباط) از جزء به کل رسیدن است، این نوع استنباط را استنباط استقرایی [۵] نیز می‌گویند.
دقت داشته باشید که وقتی داده‌ها از یک مدل احتمالی تبعیت می‌کنند، یعنی $x \sim f(x|\theta)$ در واقع فرض می‌کنیم، تمامی داده‌هایی که ساختار تصادفی مورد بررسی می‌تواند تولید کند، از توزیع احتمالی بالا قابل تولید است. در انجام استنباط آماری، عکس این موضوع را در نظر می‌گیریم، یعنی می‌گوییم اگر چنین مدلی را داشته باشیم، این مدل می‌تواند داده‌هایی را که در اختیار داریم (داده‌های مشاهده شده) تولید کند.
این ویژگی را در مقایسه با مدل احتمالی، ویژگی وارونگی[۶] استنباط آماری می‌گویند. این ویژگی در نظریه درستنمایی بطور کامل نمود پیدا می‌کند [۱۲]:

$$L(\theta|x) = f(X|\theta)$$

[۱] Statistical Inference

[۲] Model

[۳] Probabilistic Model

[۴] Statistical Model

[۵] Induction Inference

[۶] Invertion

سمت راست عبارت فوق در واقع مدل احتمالی است، و سمت چپ آن مدل آماری می‌باشد.

دو دیدگاه کلی برای انجام استنباط آماری در ادبیات آماری وجود دارد. استنباط کلاسیک یا فراوانی‌گرا[۷] و استنباط بیزی[۸]. دیدگاه فراوانی‌گرا بر این اساس است که احتمال، یک فراوانی حدی است، یعنی احتمال یک پیشامد برابر است با فراوانی نسبی وقوع آن پیشامد؛ این بدین معنی است که فراوانی‌گراها احتمال را برای پیشامدهای قابل تکرار در نظر می‌گیرند که در آن عدم حتمیت به خاطر تصادفی بودن پیشامدها می‌باشد. در صورتی که عدم حتمیت وقوع پیشامد، به خاطر عدم اطلاع در مورد آن پیشامد باشد، فراوانی‌گراها نباید به آن احتمالی نسبت دهند. برای مثال، در مورد این پیشامد که آیا در سیاره‌ی دیگری نیز حیات وجود دارد یا خیر، یک فراوانی‌گرا نمی‌تواند احتمالی را به آن نسبت دهد. از آن جایی که عدم حتمیت در مورد پارامتر به خاطر عدم شناخت است، استنباط فراوانی‌گرا نمی‌تواند به پارامتر مجهول در مدل، توزیع احتمالی نسبت دهد. برای مثال، از دیدگاه فراوانی‌گرا، اگر یک فاصله‌ی اطمینان ۹۵٪ برای میانگین مجهول جامعه، [۱ ,۱−] باشد، این بدین معنی نیست که احتمال اینکه میانگین در این فاصله باشد، برابر ۹۵٪ است؛ بلکه بدین معنی است که اگر به تعداد دفعات زیادی نمونه‌گیری کنیم و فاصله‌ی اطمینان ۹۵٪ را بسازیم، انتظار داریم در ۹۵٪ حالات، این فاصله شامل میانگین مجهول باشد.

از دیدگاه استنباط بیزی ، پارامترهای مجهول، متغیر تصادفی فرض می‌شوند. عدم حتمیت یا درجه‌ی اعتقاد نسبت به پارامترها، با استفاده از توزیع‌های احتمالی بیان می‌شود. یعنی اگر توزیع احتمالی پارامتر، به یک قسمتی از فضای پارامتر، وزن بیشتر می‌دهد، این بدان معنی است که از دیدگاه بیزی، اعتقاد بیشتری نسبت به تعلق پارامتر به آن محدوده از فضای پارامتری وجود دارد. این اعتقاد را اعتقاد پیشین[۹] و توزیع احتمال آن را توزیع پیشین[۱۰] می‌نامند. روش استنباط بیزی با استفاده از داده‌ها، توزیع پیشین را به‌روز[۱۱] کرده و توزیع احتمال جدید را توزیع پسین[۱۲] می‌نامیم.
صرف نظر از توانایی‌ها و ایرادهایی که بر هر کدام از این دو رویکرد وارد است، ما در این پایان‌نامه از رویکرد بیزی برای استنباط استفاده خواهیم نمود. برای کسب اطلاعات بیشتر در زمینه‌ی توانایی‌ها و ایرادهای این دو رویکرد، به مراجع [۲۱] و [۱۲] مراجعه کنید.
در این فصل ابتدا قضیه‌ی بیز را بیان می‌کنیم، سپس به معرفی انواع توزیع‌های پیشین می‌پردازیم و نهایتاً در قالب نظریه‌ی تصمیم بیزی، برآورد نقطه‌ای بیز را بررسی خواهیم کرد.

[۷] Frequentistic Inference

[۸] Bayesian Inference

[۹] Prior Belief

[۱۰] Prior Distribution

[۱۱] Update

[۱۲] Posterior Distribution

۲.۱ قضیه بیز

فرض کنید $A_1, ..., A_n$ یک افراز از فضای نمونه‌ای، و E یک پیشامد دلخواه باشد؛ قضیه بیز به صورت زیر بیان می‌شود:

$$P(A_i|E) = \frac{P(E|A_i)P(A_i)}{\Sigma_{i=1}^{n} P(E|A_i)P(A_i)} \tag{۱.۱}$$

در واقع این قضیه یک روش رسمی برای به روز رساندن شانس A_i از $P(A_i)$ به $P(A_i|E)$ ، وقتی E مشاهده شده است، می‌باشد. $P(A_i)$ را احتمال پیشین و $P(A_i|E)$ را احتمال پسین می‌نامیم. این قضیه برای تابع چگالی (مدل آماری) به صورت زیر تعریف می‌شود:

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{۲.۱}$$

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\sum f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{۳.۱}$$

که در آن $\pi(\boldsymbol{\theta})$، *توزیع پیشین* $\boldsymbol{\theta}$ ، بیان کننده‌ی اعتقادات و اطلاعات پیشین ما درباره‌ی مقدار نامعلوم $\boldsymbol{\theta}$ است؛ و $\pi(\boldsymbol{\theta}|\boldsymbol{x})$، *توزیع پسین* $\boldsymbol{\theta}$ ، بیان کننده‌ی اعتقادات و اطلاعات درباره‌ی $\boldsymbol{\theta}$ بعد از مشاهده‌ی نمونه‌ی $\boldsymbol{x}$ می‌باشد.

دقت کنید قضیه‌ی بیز بیانگر چگونگی تلفیق اطلاعات پیشین و اطلاعات حاصل از مشاهده‌ی جدید می‌باشد. در واقع اگر $\pi(\boldsymbol{\theta})$ و $f(\boldsymbol{x}|\boldsymbol{\theta})$ نشان دهنده اطلاعات شخصی مستدل و منطقی باشند، قضیه بیز یک روش بهینه برای به روز رسانی اطلاعات شخصی درباره‌ی $\boldsymbol{\theta}$، بعد از اطلاع جدید $\boldsymbol{x}$ می‌باشد [۲۱].

از آنجایی که مقدار انتگرال $\int f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ مستقل از $\boldsymbol{\theta}$ بوده و حاصل آن یک عدد ثابت می‌باشد، در محاسبه‌ی $\pi(\boldsymbol{\theta}|\boldsymbol{x})$ معمولاً از محاسبه‌ی انتگرال خودداری کرده و توزیع پسین را به صورت زیر بیان می‌کنند:

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) \propto f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{۴.۱}$$

نکته ۱.۲.۱ (ابرپارامتر) *اگر توزیع پیشین خود دارای پارامترهای نامعلومی باشد، به آنها ابر پارامتر*[۱۳] *گوییم.*

[۱۳] Hyper parameter

مثال ۲.۲.۱ فرض کنید $X \sim Bin(n,\theta)$، $\theta \in (0,1)$ و توزیع پیشین $Beta(\alpha,\beta)$ معلوم‌اند. توزیع پسین را بدست آورید.

می‌دانیم

$$f(x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x} \quad , \quad \pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

بنابراین

$$\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{\int_0^1 \pi(\theta)f(x|\theta)}$$

$$= \frac{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\binom{n}{x}\theta^x(1-\theta)^{n-x}}{\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\binom{n}{x}\theta^x(1-\theta)^{n-x}}$$

$$= \frac{\theta^{\alpha+x-1}(1-\theta)^{\beta-x+n-1}}{\int_0^1 \theta^{\alpha+x-1}(1-\theta)^{\beta-x+n-1}}$$

$$= \frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+x)\Gamma(\beta-x+n)}\theta^{\alpha+x-1}(1-\theta)^{\beta-x+n-1}$$

بنابراین توزیع پسین $Beta(\alpha+x, n+\beta-x)$ می‌باشد.

۱.۲.۱ توزیع‌های پیشین سره و ناسره

گوییم $\pi(\boldsymbol{\theta})$ ناسره[۱۴] (تعمیم یافته[۱۵]) است اگر:

$$\int_\Theta \pi(\boldsymbol{\theta})d\boldsymbol{\theta} = +\infty \qquad (5.1)$$

باشد. در صورتی که داشته باشیم

$$\int_\Theta \pi(\boldsymbol{\theta})d\boldsymbol{\theta} < \infty \qquad (6.1)$$

[۱۴] Improper

[۱۵] Generalized

آنگاه توزیع $\pi(\boldsymbol{\theta})$ را توزیع سره[۱۶] می‌گوییم. توجه داشته باشید که زمانی می‌توان از توزیع پیشین ناسره برای $\boldsymbol{\theta}$ استفاده کرد که توزیع پسین بدست آمده با استفاده از قضیه بیز، توزیع سره باشد [۱۲].

۳.۱ انتخاب توزیع پیشین

یک گام اساسی در رویکرد بیزی، انتخاب توزیع پیشین برای پارامتر $\boldsymbol{\theta}$ می‌باشد. برای تعیین توزیع پیشین، مکاتب فکری مختلفی وجود دارد که این مکاتب را در دو دسته طبقه بندی می‌کنیم. مکتب اول، رویکرد *عینی*[۱۷] *و ذهنی*[۱۸] و مکتب دوم، رویکرد *تجربی*[۱۹] *و سلسله مراتبی*[۲۰] می‌باشد [۱۷].

۱.۳.۱ توزیع‌های پیشین ذهنی و عینی

الف) **توزیع‌های پیشین ذهنی**

این رویکرد تلاش می‌کند تا تمام اطلاعات قبلی ممکن را در انجام استنباط در نظر بگیرد. این اطلاعات می‌تواند به صورت داده‌های آزمایشی قبلی، یا تجربه‌ای از تخصص‌های فردی باشند. جمع آوری این گونه اطلاعات به شکل یک تابع چگالی در اغلب حالات امکان پذیر نیست؛ اما در برخی از خانواده‌های توزیع‌ها مانند خانواده‌ی توزیع نمایی می‌توان توزیع‌های پیشین ذهنی را تعیین کرد. این توزیع‌های پیشین ذهنی می‌توانند *توزیع‌های پیشین مزدوج*[۲۱] باشند.

تعریف ۱.۳.۱ (توزیع‌های پیشین مزدوج) *$\pi(\theta)$ را برای $f(x|\theta)$ مزدوج گوییم هرگاه توزیع پسین آن متعلق به خانواده‌ی توزیع پیشین $\pi(\theta)$ باشد.*

در جداول (۱.۱) و (۲.۱) برخی توزیع‌های مزدوج در حالت‌های گسسته و پیوسته آمده است. [۱۳]

مزایای استفاده از توزیع‌های پیشین مزدوج توزیع‌های پیشین مزدوج حداقل دو مزیت زیر را دارند: [۱۷]

۱. استفاده از آنها در استنباط بیزی منجر به انتگرال‌هایی می‌شود که جواب تحلیلی دارند.

---

[۱۶] Proper

[۱۷] Objective

[۱۸] Subjective

[۱۹] Emprical

[۲۰] Hierarchical

[۲۱] Conjugate priors

جدول ۱.۱: توزیع‌های درستنمایی گسسته

توزیع	پارامترهای مدل	توزیع پیشین مزدوج	ابرپارامترهای پیشین	ابرپارامترهای پسین
برنولی	p(احتمال)	بتا	α, β	$\alpha + \Sigma_{i=1}^{n} x_i, \beta + n - \Sigma_{i=1}^{n} x_i$
دوجمله‌ای	p(احتمال)	بتا	α, β	$\alpha + \Sigma_{i=1}^{n} x_i, \beta + \Sigma_{i=1}^{n} N_i - \Sigma_{i=1}^{n} x_i$
دوجمله‌ای منفی	p(احتمال)	بتا	α, β	$\alpha + rn, \beta + \Sigma_{i=1}^{n} x_i$
پواسون	λ(نرخ)	گاما	θ, k	$k + \Sigma_{i=1}^{n} x_i, \frac{\theta}{n\theta + 1}$
چند جمله‌ای	$\mathbf{p}$(بردار احتمال)	دریکله	$\overrightarrow{\alpha}$	$\overrightarrow{\alpha} + \Sigma_{i=1}^{n} \overrightarrow{x}^{(i)}$
هندسی	p(احتمال)	بتا	α, β	$\alpha + n, \beta + \Sigma_{i=1}^{n} x_i$

جدول ۲.۱: توزیع‌های درستنمایی پیوسته

توزیع	پارامترهای مدل	توزیع پیشین مزدوج	ابرپارامترهای پیشین	ابرپارامترهای پسین
یکنواخت	θ	پارتو	x_m, k	$max\{x_{(n)}, x_m\}, k + n$
نمایی	λ(نرخ)	گاما	α, β^*	$\alpha + n, \beta + \Sigma_{i=1}^{n} x_i$
نرمال (با واریانس معلوم σ^2)	μ(میانگین)	نرمال	μ_0, σ_0^2	$\left(\frac{\mu_0}{\sigma_0^2} + \frac{\Sigma_{i=1}^{n} x_i}{\sigma^2}\right) \Big/ \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right), \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)^{-1}$

۲. اگر محاسبه‌ی توزیع پسین امکان پذیر باشد، آنگاه می‌توان اطمینان داشت که استفاده از این توزیع به عنوان توزیع پیشین برای استنباط‌های بعدی منجر به توزیع پیشین سره‌ای خواهد شد.

در حالت کلی برای تعیین توزیع‌های پیشین ذهنی چهار روش مختلف وجود دارد. [۱۴]

۱. رویکرد درستنمایی نسبی

۲. رویکرد هیستوگرام

۳. انطباق با صورت تابعی داده شده

۴. تعیین تابع توزیع تجمعی

در ادامه، با یک مثال رویکرد درستنمایی نسبی را توضیح می‌دهیم. برای اطلاعات بیشتر به [۱۴] مراجعه کنید.

این رویکرد اغلب زمانی که فضای پارامتری، زیر مجموعه‌ای از خط حقیقی است، (یعنی: $\Theta \subseteq \mathbb{R}$) مورد استفاده قرار می‌گیرد. در این رویکرد به مقایسه‌ی درستنمایی‌های تک تک نقاط مختلف فضای پارامتری پرداخته و مستقیماً از این توابع برای تعیین توزیع پیشین ذهنی استفاده می‌کنند. [۱۴]

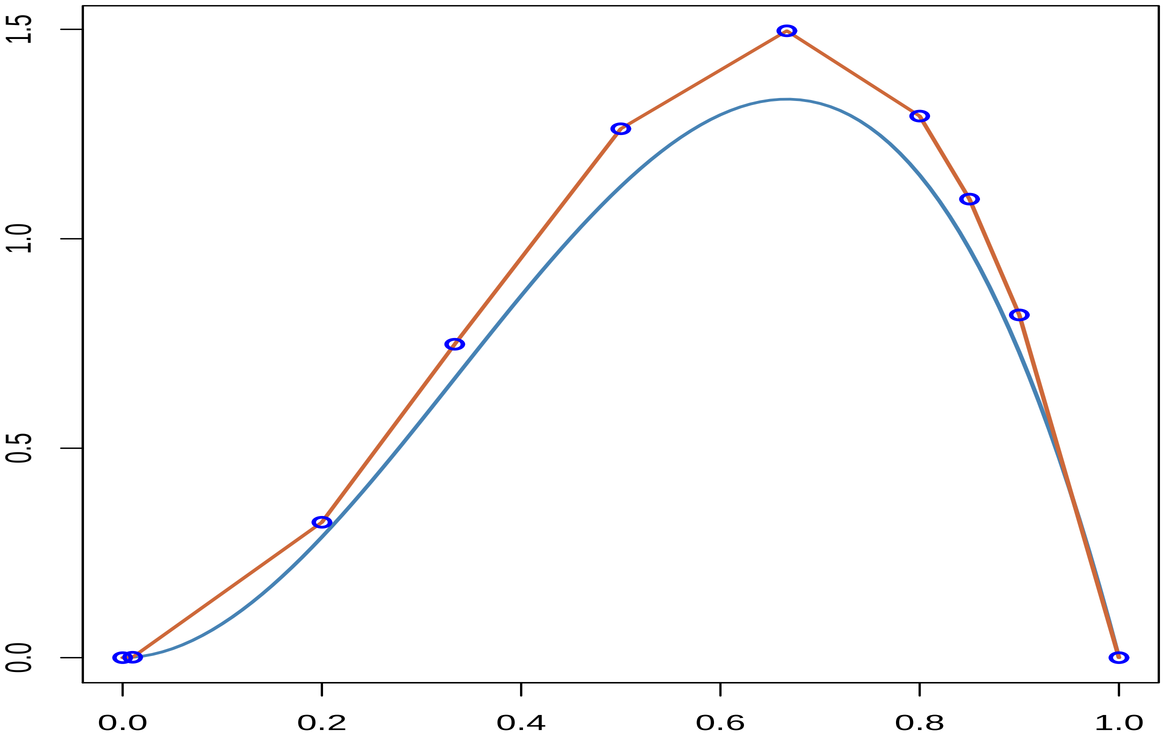
مثال ۲.۳.۱ *فرض کنید $\Theta = [۰, ۱]$ باشد. همچنین فرض کنید که مقادیر درستنمایی برای نقاط:*

$$\theta = ۰, ۰/۲, ۰/۳۳, ۰/۵, ۰/۶۶, ۰/۸, ۰/۸۵, ۰/۹, ۱$$

به صورت زیر باهم در ارتباط باشند: (به صورت ذهنی فرض می‌کنیم.)

$$L(۰|X) = L(۱|X) = ۰$$
$$L(۰/۲|X) = ۰/۳۲L(۰/۰۱|X)$$
$$L(۰/۳۳|X) = ۰/۷۵L(۰/۰۱|X)$$
$$L(۰/۵|X) = ۱/۲۶L(۰/۰۱|X)$$
$$L(۰/۶۶|X) = ۱/۵L(۰/۰۱|X)$$
$$L(۰/۸|X) = ۱/۲۹L(۰/۰۱|X)$$
$$L(۰/۸۵|X) = ۱/۰۹L(۰/۰۱|X)$$
$$L(۰/۹|X) = ۰/۸۲L(۰/۰۱|X)$$

(این اطلاعات به صورت ذهنی بوده و تماماً بر پایه‌ی حدس و گمان خودمان می‌شد، که این یکی از ایرادهای مدل ذهنی است) با در نظر گرفتن یک مقدار ثابت، مثلاً ۱ ، برای $L(۰/۰۱|X) = ۱$ نمودار درستنمایی به ازای نقاط متناظر از فضای پارامتر می‌تواند به صورت شکل (۱.۱) باشد. (منحنی با خطوط شکسته)



شکل ۱.۱: نمودار درستنمایی (خطوط شکسته) و منحنی توزیع پیشین برازش داده شده

هر چقدر نقاط بیشتری از فضای پارامتری را در نظر گرفته و مقدار متناظر تابع درستنمایی آنها را بدست آوریم، نمودار درستنمایی به منحنی نزدیک و نزدیکتر خواهد شد. حال اگر بتوانیم شکل تابعی نمودار را حدس زده و یا آن را تقریب بزنیم، از این شکل تابعی می‌توان به عنوان توزیع پیشین استفاده نمود. در این مثال، این شکل تابعی به صورت مضربی از $\theta^{۲}(۱-\theta)$ *تقریب زده شده است.*

$$\pi(\theta) \propto \theta^{۲}(۱-\theta)$$

شکل (۱.۱)، این فرم تابعی و نقاط متناظر تابع درستنمایی را نشان می‌دهد.

ب) توزیع‌های پیشین عینی

هدف رویکرد عینی در مقابل رویکرد ذهنی قرار دارد. به جای تلاش برای فرمول‌بندی حجم وسیعی از اطلاعات به صورت یک توزیع پیشین، رویکرد عینی سعی در استفاده از حداقل اطلاعات در بدست آوردن توزیع پیشین دارد. این رویکرد به داده‌ها اجازه می‌دهد تا حد امکان وزن زیادی در توزیع پسین داشته باشند. اصطلاحاً در این رویکرد به داده‌ها اجازه داده می‌شود، خود در مورد پارامتر صحبت کنند [۲۲]. [۱۷] توزیع‌های پیشین عینی معمولاً توزیع‌های ناسره‌ای هستند که منجر به توزیع‌های پسین سره می‌شوند. در

[۲۲]letting the data speak for itself