دانشگاه پیام نور

دانشکده علوم

گروه آمار

عنوان پایان نامه

برآورد بیزی پارامترهای توزیع نمایی تعمیم‌یافته

استاد راهنما

دکتر پرویز نصیری

استاد مشاور

دکتر مسعود یارمحمدی

نگارش

سیده محبوبه حسینی

پایان نامه

برای دریافت درجه کارشناسی ارشد

مرداد ماه ۱۳۸۹

چکیده

در سالهای اخیر یک توزیع جدید با نام توزیع نمایی تعمیم یافته معرفی شده و به طور وسیع مورد مطالعه قرار گرفته است . توزیع نمایی تعمیم یافته مانند توزیع گاما و وایبل دارای دو پارامتر(شکل و مقیاس) است . بنابراین توزیع نمایی تعمیم یافته را می‌توان با عنوان جایگزینی برای توزیع‌های گاما و وایبل به کار برد . پیدایش این مدل ، ویژگی‌های مختلف ، روش‌های مختلف برآورد و ویژگی‌های آنها در این پایان نامه بحث خواهد شد. در این پایان نامه به بررسی برآورد بیزی پارامترهای نامعلوم توزیع نمایی تعمیم یافته تحت انتخاب توزیع پیشین گاما روی پارامترهای شکل و مقیاس می پردازیم . برآوردگرهای بیز توزیع نمایی تعمیم یافته بصورت فرم صریحی بدست نمی‌آیند . برآوردگرهای بیز تقریبی بوسیله ایده لیندلی و روش MCMC محاسبه می‌شوند در پایان نیز شبیه سازی انجام شده است .

کلید واژه ها : توزیع نمایی تعمیم یافته، برآوردگرهای ماکزیمم درستنمایی ، برآوردگرهای بیز، توزیع گاما، لگاریتم تابع چگالی مقعر، تابع زیان مربع خطا، تابع چگالی پسین

فهرست مطالب

عنوان .... صفحه

فهرست نمودارها و جداول

عنوان .... صفحه

فصل اول

مفاهیم اساسی و تعاریف پایه ای

۱- ۱ مقدمه و تاریخچه

توزیع‌های گاما و وایبل سه پارامتری، رایج‌ترین توزیع‌ها برای تجزیه و تحلیل داده‌های طول عمر می‌باشند که در مقایسه با دیگر توزیع‌های طول عمر کاربرد بیشتری دارند. پارامترهای این توزیع‌ها عبارتند از پارامتر شکل، مقیاس و مکان. که به علت وجود این پارامترها انعطاف‌پذیری بالاتری برای تجزیه و تحلیل انواع داده‌های طول عمر دارند. این توزیع‌ها وابسته به شکلشان دارای نرخ خطر صعودی و یا نزولی‌اند.

توزیع‌های گاما و وایبل سه پارامتری دارای نقاط قوت و ضعیف زیر می‌باشند. ایراد اصلی توزیع گاما این است که نمی‌توان تابع توزیع و تابع بقاء آن را هنگامی که پارامتر شکل صحیح نباشند، به سادگی محاسبه کرد. برای محاسبه‌ی تابع توزیع و تابع بقاء و تابع نرخ خطر توزیع گاما باید از جداول ریاضی یا نرم افزارهای کامپیوتری استفاده کرد. این در حالی است که تابع توزیع، تابع بقاء و تابع نرخ خطر وایبل به سادگی محاسبه می‌شود. افزون براین در عمل نیز برای بسیاری از داده‌های واقعی توزیع وایبل در مقایسه با توزیع گاما بهتر برازش داده می‌شود. علی‌رغم مزیت‌هایی که برای توزیع وایبل در فوق برشمردیم ایرادهایی نیز بر این توزیع وارد است، که به این صورت است که بین[۱] در سال ۱۹۷۸ نشان داد که برآوردگر**های** درستنمایی ماکزیمم این توزیع ممکن است برای همه مقادیر پارامترها رفتار خوبی داشته باشند، حتی هنگاهی که پارامتر مکان صفر باشد.

اگر چه تابع توزیع وایبل دارای شکل ساده‌ای می‌باشد، با این وجود براساس یک نمونه‌ی تصادفی nتایی از این توزیع، توزیع مجموع و همچنین توزیع میانگین نمونه را نمی‌توان به راحتی محاسبه کرد. در حالی که برای

---

[۱] Bain

توزیع گاما محاسبه‌ی توزیع مجموع و توزیع میانگین نمونه به راحتی امکان‌پذیر است. توزیع وایبل سه پارامتری با یک پارامتر مقیاس و دو پارامتر شکل توسط مود هلکار[۲] و سریواستاوا در سال ۱۹۹۵ معرفی شد. آنها مجموعه داده‌های معینی را تجزیه و تحلیل کردند و نشان دادند که توزیع وایبل نمایی سه پارامتری در مقایسه با توزیع وایبل دو پارامتری(با پارامتر مکان صفر) و توزیع نمایی که هر دوی این‌ها حالت خاص توزیع وایبل نمایی‌اند، بهتر به این داده‌ها برازش می‌شود. حالت خاص از مدل وایبل نمایی، مدل نمایی تعمیم یافته می‌باشد که توسط گوپتا و کوندا[۳] در سال ۱۹۹۷ معرفی شد. این مدل ویژگی‌های مطلوب مدل‌های گاما و وایبل را دارا است. از این رو از آن می‌توان برای تجزیه و تحلیل داده‌های طول عمر استفاده کرد. اغلب توزیع نمایی به عنوان توزیع مدت زمانی که یک اتفاق خاصی رخ می‌دهد مطرح شد و برای مثال، زمان لازم تا وقوع یک پیشامد، یک متغیر تصادفی با توزیع نمایی در نظر گرفته می‌شود.

در این فصل، در بخش دوم به تعاریف پایه‌ای مانند تابع چگالی گاما، تابع چگالی نمایی و تابع چگالی بتا، تابع چگالی وایبل، تابع چگالی پاراتو می‌پردازیم و برخی خواص این توزیع‌ها را نیز به صورت مختصر بیان می‌کنیم. در بخش سوم مفاهیم مورد نیاز در روش بیز مانند توزیع‌های پسین و پیشین و در بخش چهارم برآوردگرهای بیز و را تعریف می‌کنیم.

۱-۲ تعاریف و نمادها

۱-۲-۱ تابع چگالی نمایی

---

[۲] Mudholkar and Srivastava
[۳] Gupta and Kunda

متغیر تصادفی X دارای تابع چگالی نمایی با پارامتر λ است اگر تابع چگالی آن به صورت زیر باشد:

$$f(x;\beta)=\beta e^{-\beta x} \qquad 0<x<\infty \quad , \quad \beta>0.$$

شاخص های میانگین، واریانس، ضریب تغییر، تابع مولد گشتاورها و تابع مشخصه آن برابرند با :

$$E(X)=\frac{1}{\beta} \qquad M_x(t)=\left(\frac{\beta}{\beta-t}\right) \quad ,t<\beta$$

$$var(X)=\frac{1}{\beta^2} \qquad \psi_x(t)=\frac{\beta}{\beta-\text{Ln}\, t} \quad ,t<e^{\beta}$$

$$cv(X)=1$$

۱-۲- ۲ تابع چگالی گاما

متغیر تصادفی X دارای تابع چگالی گاما با پارامترهای α, β است اگر تابع چگالی آن به صورت زیر باشد:

$$f(x;\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} \qquad 0<x<\infty \ ,\alpha>0,\beta>0$$

حالت‌های خاص:

اگر $\alpha=1$، یعنی $\Gamma(1,\beta)$، همان توزیع نمایی با پارامتر β است.

اگر $\alpha=n$، $n\in N$، یعنی $\Gamma(n,\beta)$، به توزیع ارلنگ معروف است.

اگر $\alpha=\frac{n}{2}, \beta=\frac{1}{2}$، یعنی $\Gamma(\frac{n}{2},\frac{1}{2})$، به توزیع کای- دو با n درجه آزادی معروف است و معمولاً آن را با نماد $\chi^2_{(n)}$ نمایش می دهند.

شاخص میانگین، واریانس، ضریب تغییر، تابع مولد گشتاورها و تابع مشخصه آن برابرند با :

$$E(X)=\frac{\alpha}{\beta} \qquad M_x(t)=\left(\frac{\beta}{\beta-t}\right)^{\alpha} \quad ,\, t<\beta$$

$$\mathrm{var}(X)=\frac{\alpha}{\beta^{۲}} \qquad \psi_x(t)=\left(\frac{\beta}{\beta-\mathrm{Ln}\, t}\right)^{\alpha} \quad ,\, t<e^{\beta}$$

$$cv(\alpha,\beta)=\frac{۱}{\sqrt{\alpha}}$$

۱-۲-۳ تابع چگالی بتا

متغییر تصادفی X دارای تابع چگالی بتا با پارامتر های α , β است اگر تابع چگالی آن به صورت زیر باشد:

$$f_X(x;\alpha,\beta)=\frac{۱}{B(\alpha,\beta)}x^{\alpha-۱}(۱-x)^{\beta-۱} \qquad ۰<x<۱\ ,\ \alpha>۰\ ,\beta>۰$$

ارتباط بین توابع گاما وبتا به صورت زیر است::

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

امید ریاضی، واریانس و ضریب تغییر آن به ترتیب برابرند با :

$$E(X)=\frac{\alpha}{\alpha+\beta}$$

$$\mathrm{var}(X)=\frac{\alpha\beta}{(\alpha+\beta+۱)(\alpha+\beta)^{۲}}$$

$$cv(\alpha,\beta)=\left[\frac{\beta}{\alpha(\alpha+\beta+۱)}\right]^{\frac{۱}{۲}}$$

۱-۲-۴ تابع چگالی وایبل

متغیر تصادفی X دارای تابع چگالی وایبل با پارامتر های α , β است اگر تابع چگالی آن به صورت زیر باشد:

$$f_X(x;\alpha,\beta) = \alpha\beta\, x^{\alpha-۱}\, e^{-\beta x^{\alpha}} \qquad x > ۰, \alpha > ۰, \beta > ۰$$

و آن را با $X \sim W(\alpha,\beta)$ نشان می‌دهیم. این تابع چگالی به‌طور چشمگیری در نظریه‌ی قابلیت اعتماد قابل استفاده است.

حالت خاص:

اگر $\alpha = ۲$، یعنی $W(۲,\beta)$، به توزیع رای لی معروف است و معمولاً آن را با نماد $R(\beta)$ نشان می‌دهند.

اگر $\alpha = ۱$، یعنی $W(۱,\beta)$، همان توزیع نمایی با پارامتر β است.

امید ریاضی، واریانس، ضریب تغییر و تابع مشخصه آن به ترتیب برابرند با :

$$E(X) = \frac{\Gamma(۱+\frac{۱}{\alpha})}{\beta^{\frac{۱}{\alpha}}}$$

$$var(X) = \frac{\Gamma(۱+\frac{۲}{\alpha}) - \Gamma^{۲}(۱+\frac{۱}{\alpha})}{\beta^{\frac{۲}{\alpha}}}$$

$$cv(\alpha,\beta) = \left\{ \frac{\Gamma(۱+\frac{۲}{\alpha})}{\Gamma^{۲}(۱+\frac{۱}{\alpha})} - ۱ \right\}^{\frac{۱}{۲}}$$

$$\psi_x(t) = \frac{\Gamma(۱+\frac{t}{\alpha})}{\beta^{\frac{t}{\alpha}}}$$

۱-۲-۵ تابع چگالی پاراتو

متغیر تصادفی X دارای تابع چگالی پاراتو با پارامترهای α, β است اگر تابع چگالی آن به صورت زیر باشد:

$$f_X(x;\alpha,\beta) = \frac{\alpha\beta^{\alpha}}{x^{\alpha+۱}} \qquad x \ge \beta \;,\; \alpha > ۰, \beta > ۰$$

امید ریاضی، واریانس و ضریب تغییرآن برابرند با :

$$E(X)=\frac{\alpha\beta}{\alpha-۱} \quad , \quad \alpha>۱$$

$$\operatorname{var}(X)=\frac{\alpha\beta^{۲}}{(\alpha-۱)^{۲}(\alpha-۲)} \quad , \quad \alpha>۲$$

$$cv(\alpha,\beta)=\left[\frac{۱}{\alpha(\alpha-۲)}\right]^{\frac{۱}{۲}} \quad , \quad \alpha>۲$$

ضمناً برای توزیع تابع مولد گشتاور وجود ندارد.

توزیع پاراتو در الگو بندی مسایلی که شامل توزیع درآمدها، وقتی که درآمدها از حد معین α فراتر روند کاربرد فراوان دارد.

۱-۲-۶ تابع چگالی نرمال

متغیر تصادفی X دارای تابع چگالی نرمال با پارامترهای μ, σ است اگر تابع چگالی آن به صورت زیر باشد:

$$f_X(x;\mu,\sigma)=\frac{۱}{\sqrt{۲\pi}\sigma}e^{-\frac{۱}{۲\sigma^{۲}}(x-\mu)^{۲}} \qquad , \ -\infty<\mu<+\infty \ , \ \sigma>۰, \quad x\in\mathbb{R}$$

حالت خاص:

اگر $\mu=۰$, $\sigma^{۲}=۱$ به نرمال استاندارد معروف است و معمولاً متغیر تصادفی نرمال استاندارد را با Z، تابع چگالی را با ϕ و تابع توزیع آن را با Φ نشان می‌دهند.

شاخص‌های میانگین، واریانس و ضریب تغییر، تابع مولدگشتاور آن برابرند با:

$$E(X) = \mu$$
$$\mathrm{var}(X) = \sigma^{۲}$$
$$\mathrm{cv}(\alpha,\beta) = \frac{\sigma}{\mu}$$
$$M_x(t) = \exp\left\{ \mu t + \frac{۱}{۲}\sigma^{۲}t^{۲} \right\}$$

۱-۳ مفاهیم و اصطلاحات روش بیز

فرمول اساسی که برای دخالت دادن اطلاعات گذشته در تحلیل آماری بکار برده می‌شود توسط توماس بیز[۴] کشیش انگلیسی در سال ۱۷۶۰ ارائه شد. توماس بیز به خاطر دانش ریاضی‌اش به عنوان عضو انجمن سلطنتی انتخاب شده بود. پس از مرگش، مقاله‌ای به نام قضیه بیز که اکنون مشهور است، در بین متعلقاتش پیدا شد. شاید اولین کاربرد آن برای مسائل رگرسیون توسط هارلد جفری[5] در سال ۱۹۳۹ باشد.

در مسائل استنباط آماری به روش بیز بر اساس مشاهداتی که از خانواده توزیع‌ها اختیار می‌شود پارامتر θ دارای یک مقدار نامعلوم است. به عبارت دیگر، θ به عنوان یک متغیر تصادفی در نظر گرفته می‌شود که مقادیر ممکن آن فضای پارامتر $\Theta = \{\theta | \theta \in \mathbb{B}\}$ است ومعمولا دارای تابع احتمال یا تابع چگالی $\pi(\theta)$ است و از آن به عنوان تابع احتمال پیشین یا تابع چگالی پیشین یاد می‌کنند. در واقع توزیع پیشین[۶] تبلور تحلیلگر آمار از خلاصه اطلاعات و دانسته‌های او درباره θ است که احتمال قرار داشتن θ در چه بخش‌هایی از Θ بیش از همه است.

با توجه به آنچه بیان شد، فرض کنید مدل احتمال آزمایش بستگی به پارامتر θ دارد و دارای خانواده

---

[4] T, Bayes
[5] H, Jeffreys
[6] Prior Distribution

چگالی‌های $\{f_\theta;\theta\in\Theta\}$ است. فرض کنید $X=(X_1,X_2,...,X_n)$ با مقادیر مشاهده شده‌ی $x=(x_1,x_2,...,x_n)$ باشند و $\theta=(\theta_1,\theta_2,...)\in\Theta$ بردار پارامتر مجهول باشد. در این‌صورت توزیع توام $\underline{x}$ و θ برابر است با:

$$f(\underline{x}\ ;\theta)=f(x|\theta)\ \pi(\theta)$$

که در آن $f(x|\theta)$ تابع درستنمایی مشاهدات و $\pi(\theta)$ توزیع پیشین می‌باشد. با فرض اینکه مشاهدات x را در دسترس باشند می‌توان به کمک قضیه بیز توزیع θ به شرط x را به صورت زیر بدست آورد.

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

که به آن توزیع پسین[۷] می گویند.

در استنباط بیزی تنها توزیع پسین پایه استنباط درباره پارامترها است. در واقع هر خلاصه آماری از θ مانند میانگین، واریانس، چندک‌ها و ناحیه با بیشترین چگالی پسین را می‌توان امید ریاضی تابعی از θ مانند $k(\theta)$ در نظر گرفت که به صورت زیر می‌توان محاسبه کرد:

$$E\left[k(\theta)|X\right]=\frac{\int k(\theta)\,\pi(\theta)\ \pi(X|\theta)\,d\theta}{\int \pi(\theta)\,\pi(X|\theta)\,d\theta}$$

محاسبه انتگرال فوق تا این اواخر یکی از عمده‌ترین مشکلات استنباط بیزی بوده است. در بسیاری ازکاربردها، بخصوص در مواقعی که بعد θ زیاد است ارزیابی تحلیلی $E\left[k(\theta)|X\right]$ امکان پذیر نیست و برای محاسبه آن روش‌های عددی مانند MCMC استفاده می‌کنیم. برای توضیح بیشتر به پیوست «ب» مراجعه شود.

ممکن است علاقمند به استفاده از اطلاعات مربوط به θ از منابع دیگر هم باشیم. اگر مجموعه همه این اطلاعات موجود در نمونه تصادفی، در قالب توزیع احتمال برای θ باشد (یعنی θ به عنوان یک متغیر تصادفی

[3] Posterior Distribution

در نظر گرفته شود.) آنگاه این اطلاعات را می‌توان با بکارگیری قضیه بیز با داده‌های موجود در هم آمیخت. بنابراین می‌توان مجموع همه این اطلاعات را در قالب یک توزیع پیشین و اطلاعات آمیخته با داده‌ها از توزیع پسین استفاده کرد.

در روش بیز باید توزیعی برای θ نیز مشخص کرد. قبل از مشاهده داده ها باید ارزیابی کنیم که درباره مقادیر احتمالی پارامترها چه می‌دانیم یا باور داریم و این معلومات یا باور را به صورت توزیع احتمال θ بیان کرد. این توزیع را توزیع پیشین θ می‌نامند.

مثال ۱-۳ فرض کنید که متغیر تصادفی X دارای توزیع دوجمله‌ای با پارامترهای n, θباشد و $\theta \in (0,1)$ دارای توزیع پیشین بتا با پارامترهای α, β باشد. در این‌صورت توزیع پسین به صورت زیر بدست می‌آید:

$$f_\theta(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \qquad x = 0,1,2,3,\dots,n$$

$$\pi(\theta) = \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \qquad 0 < \theta < 1$$

بنابراین

$$\begin{aligned} f(x,\theta) &= f_\theta(x).\pi(\theta) \\ &= \frac{1}{B(\alpha,\beta)} \binom{n}{x} \theta^{x+\alpha-1} (1-\theta)^{n+\beta-x-1} \ d\theta \\ f_X(x) &= \int_0^1 f(x,\theta)\, d\theta \\ &= \int_0^1 \frac{1}{B(\alpha,\beta)} \binom{n}{x} \theta^{x+\alpha-1} (1-\theta)^{n+\beta-x-1} \ d\theta \\ &= \binom{n}{x} \frac{B(\alpha+x, n+\beta-x)}{B(\alpha,\beta)} \end{aligned}$$

در نتیجه

$$\begin{aligned}\pi(\theta|x) &= \frac{f_\theta(x)\pi(\theta)}{\int f_\theta(x)\pi(\theta)d\theta} \\ &= \frac{۱}{B(\alpha+x, n+\beta-x)}\theta^{\alpha+x-۱}(۱-\theta)^{n+\beta-x-۱}\end{aligned}$$

رابطه اخیر توزیع پسین برای θ است و دارای توزیع بتا با پارامترهای $(\alpha+x, n+\beta-x)$ می‌باشد.

بنابراین

$$\pi(\theta|x) \sim \text{Beta}(\alpha+x, n+\beta-x)$$

مثال ۲–۳ فرض کنید $X \sim N(\theta,۱)$ و $\theta \sim N(\mu,\tau^۲)$ باشد می‌توان برای یک نمونه تصادفی $x_۱, x_۲, ..., x_n$ نشان داد که:

$$\pi(\theta|\underline{x}) = \frac{۱}{\sqrt{۲\pi\frac{۱}{۱+\tau^۲}}}\exp\left\{-\frac{۱}{۲}(۱+\frac{۱}{\tau^۲})\left[\theta-(x+\frac{\mu}{\tau^۲})/(۱+\frac{۱}{\tau^۲})\right]^۲\right\}$$

رابطه اخیر توزیع پسین برای θ است و دارای توزیع نرمال است. بنابراین

$$\pi(\theta|x) \sim N\left(\frac{x+\frac{\mu}{\tau^۲}}{۱+\frac{۱}{\tau^۲}}, \frac{۱}{۱+\frac{۱}{\tau^۲}}\right)$$

بخش مهم آنالیز داده‌های بیزی انتخاب توزیع پیشین است. پس در آنالیز بیزی ممکن است نگران باشیم که با چه قطعیتی نتایج به توزیع پیشین خاصی که انتخاب شده بود وابسته است. دو نوع توزیع پیشین در زیر ارائه می‌شود:

۱. توزیع پیشین فاقد اطلاع[8]

۲. توزیع‌های پسین و پیشین مزدوج

۱-۳-۱ توزیع پیشین فاقد اطلاع

توزیع پیشین فاقد اطلاع توسط لی[9] (۱۹۸۹، ص۴۶) " مرجع پیشین " نامیده شده است. برآوردهای بیزی توزیع‌های پیشین فاقد اطلاع، همان میانگین پارامتر توزیع پسین است. می‌توان نشان داد که این برآوردها، دقیقاً مانند برآوردهای کمترین مربعات می‌باشند. اطلاعات قبلی یا باور یک تحلیلگر درباره پارامتر θ بستگی به او دارد. اگر متغیر تصادفی θ دارای تابع چگالی زیر باشد.

$$f(\theta)=\frac{۱}{C} \qquad ۰<\theta<C$$

تابع چگالی بالا دقیقاً یک تابع چگالی احتمال نیست، زیرا انتگرال آن در طول دامنه پارامترها بی‌نهایت است. بنابراین رابطه فوق توزیع احتمال معتبری را برای پارامترها مشخص نمی‌کند. با این

وجود، اگر این تابع را در قضیه بیز استفاده کنیم یک توزیع پسین معتبری را ارائه می دهد.

۱-۳-۲ توزیع های پسین وپیشین مزدوج

اغلب خانواده‌هایی از توزیع‌ها برای θ وجود دارند که اگر توزیع پیشین متعلق به این خانواده باشد آنگاه توزیع پسین نیز متعلق به این خانواده خواهد بود، اصطلاحاً این چنین خانواده از توزیع‌ها را خانواده توزیع‌های مزدوج[10] می‌نامند.

---

[8] Nonin Formative Prior Distribution
[9] Lee
[10] Conjugate Distribution Family

به عبارت دیگر توزیع پیشین مزدوج توزیعی است که با ترکیب آن با توزیع بردار داده‌ها توزیع پسینی بدست می‌آید که دارای همان شکلی است که توزیع پیشین دارد. با به کار بردن توزیع پیشین مزدوج قادر خواهیم بود که نسبت به توزیع پسین فرمول صریحی برای برآورد بیزی پارامترها (مانند میانگین پارامترها) بدست آوریم.

یک روش برای انتخاب توزیع پسین مزدوج، بر پایه ساختار توزیع توام $\underline{X}=(X_1,X_2,...,X_n)$ برحسب تابعی از آماره بسنده است. برای روشن شدن این موضوع، می‌دانیم اگر $\underline{x}=(x_1,x_2,...,x_n)$ یک نمونه تصادفی از تابع چگالی $f_\theta(x)$ باشد و T آماره بسنده باشد، آنگاه بنا به قضیه دسته بندی نیمن

$$\begin{aligned} f_\theta(x_1,x_2,...,x_n) &= \prod_{i=1}^{n} f_\theta(x_i) \\ &= h(\theta,T(x)).k(x) \\ &\propto h(\theta,T(x)) \end{aligned}$$

بنابراین

$$\pi(\theta|x) \propto h\left(\theta,T(x)\right)\pi(\theta)$$

یعنی درصورت وجود آماره بسنده، توزیع پسین θ بر پایه مشاهدات $X=(X_1,X_2,...,X_n)$ از طریق آماره بسنده بستگی به x دارد. بنابراین براساس ساختار $h\left(\theta,T(x)\right)$ اگر $\pi(\theta)$ طوری انتخاب شود که توزیع پسین نیز متعلق به خانواده توزیع پیشین شود، چنین انتخابی از توزیع پیشین، از خانواده توزیع‌های مزدوج خواهد بود.

۴-۱ برآوردگر بیز[11]

تعریف : برآوردگر بیز پارامتر $\gamma(\theta)$ با تابع چگالی پیشین g و تابع زیان L[12] که آن را با δ_g نمایش

[11] Bayes Estimator

می‌دهیم، برآوردگری است که دارای مینیمم مخاطره بیزی باشد، یعنی

$$\delta_g = \min_\delta \ r(g,\delta)$$

که در آن

$$r(g,\delta) = \int L\left(\theta,\delta(x)\right)\, g(\theta|x)\, d\nu(\theta)$$

توجه داشته باشید که تعریف فوق روش پیدا کردن برآوردگر بیز $\gamma(\theta)$ را ارائه نمی‌دهد. برای پیدا کردن برآوردگر بیز پارامتر $\gamma(\theta)$، با توجه به تعریف فوق، می‌توانیم به صورت زیر عمل کنیم:

با توجه به اینکه به دنبال برآوردگری نظیر $\delta_g(X)$ هستیم که $r(g,\delta)$ را به ازای هر δ مینیمم کند

پس

$$\begin{aligned} r(g,s) &= \int_\theta \{ \int_X L(\theta,\delta(x))\, f_\theta(x) d\mu(x) \}\, g(\theta|x)\ d\nu(\theta) \\ &= \int_X \{ \int_\Theta L(\theta,\delta(x))\, g(\theta|x)\, d\nu(\theta) \}\, f_X(x)\, d\mu(x) \end{aligned}$$

و چون انتگرال دوگانه اخیر غیرمنفی است، بنابرین انتگرال را می‌توان مینیمم کرد، اگر بتوانیم عبارت داخل آکولاد، یعنی مخاطره پسین را برای هر x داده شده‌ای مینیمم کنیم و آن را با $\delta_g(x)$ نشان می‌دهیم. اگر $\delta_g(x)$ منحصر به فرد باشد آنگاه برآوردگر بیز نیز منحصر به فرد خواهد بود. در ادامه برآوردگر بیز را تحت توابع زیان زیر بدست می‌آوریم:

۱.تابع زیان مربع خطا [۱۳]

۲.تابع زیان مربع خطای وزنی [۱۴]

۳.تابع زیان قدر مطلق خطا [۱۵]

---

[12] Loss Function
[13] Error Square
[۱٤] Weighted Error Square

الف. اگر تابع زیان مربع خطا باشد

$$L(\theta,\delta)=\left(\delta(x)-\gamma(\theta)\right)^{۲}$$

آنگاه برآوردگر بیز به‌صورت زیر خواهد بود:

$$\delta_g(X)=E\left(\gamma(\theta)\middle|X\right)$$

و این یعنی برآوردگر بیز پارامتر $\gamma(\theta)$ نسبت به چگالی پیشین g و تابع زیان مربع خطا، میانگین $\gamma(\theta)$ تحت چگالی پسین است.

ب. اگر تابع زیان مربع خطای وزنی با وزن $w(\theta)$ باشد:

$$L(\theta,\delta)=w(\theta)(\delta(X)-\gamma(\theta))^{۲}$$

آنگاه برآوردگر بیز به‌صورت زیر خواهد بود:

$$\delta_g(X)=\frac{E[w(\theta)\gamma(\theta)|X]}{E[w(\theta)|X]}$$

و این یعنی برآوردگر بیز پارامتر $\gamma(\theta)$ نسبت به چگالی پیشین g و تابع زیان مربع خطای وزنی با وزن $w(\theta)$ نسبت میانگین $w(\theta)\gamma(\theta)$ به $w(\theta)$ تحت چگالی پسین است.

ج. اگر تابع زیان قدر مطلق خطا باشد :

$$L(\theta,\delta)=\left|\delta(X)-\gamma(\theta)\right|$$

آنگاه برآوردگر بیز به‌صورت زیر خواهد بود:

$$\delta_g(X)=\text{میانه توزیع}\quad \gamma(\theta)|X$$

---

[15] Absolute Error