بسم الله الرحمن الرحيم

پایان نامه
برای دریافت درجه کارشناسی ارشد
در رشته آمار ریاضی

دانشکده علوم پایه

گروه علمی آمار

عنوان پایان نامه:

برآورد بیز روی توزیع تعمیم یافته نمایی در داده‌های گروه بندی شده

استاد راهنما:
دکتر پرویز نصیری

استاد مشاور:
دکتر مسعود یار محمدی

نگارش:

شهریورماه ۱۳۸۹

تقدیم به:
اراده پولادینی که
تیشه زد، تا ریشه ام دواند
حک زد، محک هستی ام
تا بدین قلم، و زین روزم
پدر، آیینه همیشه گویا

و تقدیم به:
ز بود و نیستی ام
سوخته دل و زجر راه خورده گام های زندگی ام
مادر، بود و نبود هستی ام

و تقدیم تو را ای آینده:
که می دانم دلم را سوزانده، سفر رانده می خوانی
تو که نامت را فقط می نویسم ...

تقدیر و سپاسگزاری

سپاس پروردگار جهانیان را که به من موهبت تعلم را ارزانی داشت تا با بهره گیری از فروغ حکمت بزرگان، میدان دانش خویش را روشنی بخشم و در پرتو اندیشه‌های والای فرزانگان قدمی هرچند کوتاه پیش نهم.

اکنون که توانسته‌ام پایان نامه ام را به اتمام برسانم، بر خود لازم می‌دانم که از زحمات بی شائبه و هدایتگرانه استاد راهنما، جناب آقای دکتر نصیری که سهم بسزایی در ارشاد و راهنمایی اینجانب بر عهده داشته‌اند و در تمامی مراحل این طرح از ابتدا تا انتها با دانش عمیق علمی خود به طرح مباحث پرداخته و با نکته سنجی‌های دقیق خود موجب غنا بخشیدن و ارتقاء کیفی رساله گردیدند، تشکر و سپاسگزاری فراوان نمایم.

از جناب آقای دکتر یارمحمدی که در سمت استاد مشاور نقش سازنده‌ای در مشاوره و راهنمایی‌ام داشته و ضعف‌ها و نارسایی‌ها را یادآوری کرده و در جهت حل آنها پیشنهادات ارزنده‌ای ارایه نمودند، سپاسگزاری کنم.

همچنین از آقای دکتر رضایی داور محترم پایان نامه که صبورانه مطالب حاضر را مورد مطالعه قرار دادند و در رفع نواقص و اشکالات کمال توجه را مبذول داشتند، قدردانی می کنم. از استاد فرزانه وگرانقدر، نماینده محترم تحصیلات تکمیلی، خانم دکتر سلطانیان کمال تشکر را دارم.

با ابراز ستوده ترین سپاس های قلبی از محبت های بی دریغ و همکاری های بی شائبه دوستان و شرکت کنندگان در جلسه دفاعیه و تمامی همکاران و دوستان عزیزی که انجام این پایان نامه همواره یار، یاور و مشوق من بوده اند، کمال تشکر را دارم. سلامتی، عزت، توفیق و سربلندی را برای کلیه این عزیزان در تمام مراحل زندگی از درگاه ایزد متعال مسئلت می نمایم.

فهرست مطالب

فصل ۱

فصل ۲: توزیع تعمیم یافته نمایی

فصل ۳: **برآوردبیز پارامترهای توزیع تعمیم یافته نمایی**

فصل ۴: داده های گروه بندی شده

فصل ۵: شبیه سازی

پیوست

چکیده

اخیراً توزیع جدیدی به نام توزیع تعمیم یافته نمایی توسط محققین آمار مورد توجه قرار گرفته است که مزیت‌های بیشتری نسبت به توزیع‌های قبلی از جمله توزیع گاما و وایبل دارد، در این پایان نامه به برآورد پارامترهای توزیع تعمیم یافته نمایی با روش بیز پرداخته ایم و همچنین در داده های گروه بندی شده این توزیع را مورد بررسی قرار داده و پارامترها را تحت گروه بندی و با روش بیز بدست آورده ایم. در نهایت به این نتیجه می رسیم که میانگین مربعات خطا در داده های گروه بندی کوچکتر از داده های گروه بندی نشده و معمولی است.

کلمات کلیدی: توزیع تعمیم یافته نمایی، تابع توزیع پیشین، تابع توزیع پسین، تابع زیان، برآوردگر بیز.

مقدمه و تعاریف

بعد از توزیع گاما و وایبل اخیراً توزیع جدیدی به نام توزیع تعمیم یافته نمایی توسط محققین آمار مورد بررسی قرار گرفته است که دراین پایان نامه به تفضیل به آن پرداخته ایم.

در فصل اول در مورد تابع زیان و تابع توزیع پیشین و پسین، در فصل دوم به معرفی تابع توزیع تعمیم یافته نمایی و ویژگی‌های آن، استنباط آماری و بیزی و بالاخره به تفاوت بین توزیع گاما و توزیع تعمیم یافته نمایی پرداخته‌ایم، در فصل سوم برآورد پارامترهای این توزیع مورد بررسی قرار داده و در فصل چهارم داده های گروه بندی شده را که شامل داده های گردشده وفاصله های سانسور شده است را در توزیع تعمیم یافته نمایی مورد بحث قرار دادیم، در انتها در فصل پنجم شبیه سازی انجام شده است .

آمار کلاسیک و بیز در برآورد نقطه‌ای میانگین یا واریانس استفاده می‌شود. در آمار بیز تمرکز روی تولید توزیع پسین بر پارامتر نامعلوم توسط داده‌ها و تابع چگالی پیشین می‌باشد.

همچنین آمار بیز اطلاعات بیشتری از تمام پارامترهای نامعلوم آماده می‌کند و اثر پارامترهای مزاحم (پارامترهای نامربوط به مبحث مورد بررسی) را از بین می‌برد، به مقاله لی[1] سال۱۹۹۷ و دراپر[2] سال ۲۰۰۰ ارجاع می‌دهیم و تفاوت‌هایی در هسته اصلی آمار بیز و آمار کلاسیک به صورت کاملا عمیق وجود دارد. در روش بیزاز ابزار قوی آماری استفاده می‌شود.

---

[1] Lee
[2] Drapper

فرض کنید متغیر تصادفی y را مشاهده کردیم و می خواهیم در مورد متغیر θ قضاوت کنیم که اگر $f(y|\theta)$ و $f(y)$ و $g(\theta)$ به ترتیب توابع چگالی توام و حاشیه‌ای y و تابع چگالی θ باشند آنگاه:

$$f(y, \theta) = f(y|\theta) . g(\theta) \qquad (۱-۱)$$

که در این صورت تابع چگالی حاشیه ای y برابر با

$$\int f(y, \theta)\, d\theta = \int f(y|\theta) . g(\theta) \qquad (۱-۲)$$

اگر $\Pi(\theta | y)$ توزیع پسین برای θ باشد آنگاه:

$$\Pi(\theta | y) = \frac{f(y|\theta).g(\theta)}{\int_0^\infty f(y|\theta).g(\theta)d\theta} \qquad (۱-۳)$$

در صورتی که θ دارای تابع چگالی گسسته باشد، داریم:

$$\Pi(\theta | y) = \frac{f(y|\theta).g(\theta)}{\sum_0^\infty f(y|\theta).g(\theta)}$$

و اگر θ به صورت $(\theta_1, \theta_2, \ldots\ldots, \theta_n)$ آنگاه:

$$\Pi(\theta_i | y) = \frac{f(y|\theta_i).g(\theta_i)}{\int_0^\infty f(y|\theta_i).g(\theta_i)d\theta_i} \qquad (۱-۴)$$

که در واقع $g(\theta)$ یک توزیع پیشین پارامتر θ است و $f(\theta| y)$ یک توزیع پسین θ وقتی داده‌های y مشاهده شده اند، می باشند.

توماس بیز اولین کسی است که در این مورد کار کرده است اما در تمام مدت عمرش مقاله‌ای از او منتشر نشده است و بالاخره پس از مرگ وی دوستش ریچارد[1] در سال ۱۷۶۴ نوشته‌ای

[1]Richard

از او می‌خواند و بعد از او استیگلر[1] در سال ۱۹۸۳ توانست این تئوری را عرضه کند و او کسی است که در سن ۲۹ سالگی به درجه استادی از دانشگاه کمبریج نائل گردید.

برای در ک بهتر از احتمال بیز مثال زیر را ارائه می کنیم.

مثال۱ – فرض کنید از هر ۱۰۰۰ خانواده یک خانواده دارای مشکل ژنتیکی باشد که فقط دارای فرزندان دختر می‌شوند بنابراین برای هرخانواده می‌توان θ را این گونه تعریف کرد:

$$\text{متغیرتصادفی } \theta = \begin{cases} \text{۰} & \text{خانواده نرمال} \\ \text{۱} & \text{خانواده دارای مشکل ژنتیکی} \end{cases}$$

حال فرض می‌کنیم که ۵ فرزند دختر بدون فرزند پسر مشاهده شده است احتمال اینکه این خانواده دارای مشکل ژنتیکی باشد چقدر است ؟

با توجه به اطلاعات اولیه و قبلی، احتمال اینکه خانواده‌ای به طور تصادفی دارای مشکل ژنتیکی باشد، ۰/۰۰۱ است بنابراین $p(\theta) =$ ۰/۰۰۱ و نیز می دانیم که

$p($خانواده مشکل ژنتیک داردا ۵ فرزند دختر باشد$) = 1 = p(y|\theta)$

حال احتمال کل را برای حالتی که ۵ فرزند دختر باشند را به صورت زیرحساب می کنیم.

$$P_r(y) = p(y|\theta=0)\,.\,p(\theta=0) + p(y|\theta=1).\,p(\theta=1)$$

$$= ۰/۹۹۹ \times (۰/۵)^{۵} + ۰/۰۰۱ \times ۱ = ۰/۰۳۲۲$$

$$P_r(\theta=1|\,y=5) = p(y|\,\theta=1)\;.\;p(\theta=1)/\,P_r(y)$$

$$= ۰/۰۳۲$$

پس از هر ۱۰۰۰ خانواده که دارای ۵ فرزند دختر هستند در واقع ۳۲ خانواده مشکل ژنتیکی دارند.

و در نهایت تابع چگالی بیز زمانی که θ به صورت یک بردار به ابعاد k باشد برابر است با

[1] Stigler

$$P(\theta | y) = \frac{p(y|\theta).p(\theta)}{\int_0^{\infty} p(y,\theta)\, d\,\theta} \qquad (۵-۱)$$

و $\theta=(\theta^{(1)},\dots\dots\theta^{(k)})$ یک بردار k بعدی از متغیرهای پیوسته است.

۱- ۱ معرفی برخی از توابع توزیع

در این بخش به معرفی تابع توزیع‌هایی می‌پردازیم که دانستن آنها در این فصل ضروری می‌باشد.

۱- توزیع نمایی: اگر متغیر تصادفی X دارای تابع چگالی احتمالی زیر باشد

$$f_{\lambda}(x) = \lambda e^{-\lambda x} \qquad x>0 \;\; , \;\; \lambda>0$$

گوییم X دارای توزیع نمایی با پارامتر λ است و آن را با نماد $x \sim E(\lambda)$ نشان می‌دهیم. که در این صورت

$$E(X) = \lambda$$

$$\operatorname{var}(X) = \frac{1}{\lambda^2}$$

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

۲- توزیع گاما: اگرمتغیر تصادفی X دارای تابع چگالی احتمال زیر باشد.

$$f_{\alpha,\lambda}(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)}.x^{\alpha-1}e^{-\lambda x} \qquad x>0\,,\,\alpha,\lambda>0$$

گوییم X دارای توزیع گاما با پارامتر α و λ است و آن را با نماد $X \sim \Gamma(\alpha,\lambda)$ نشان می‌دهیم.

$$E(X) = \frac{\alpha}{\lambda}$$

$$\operatorname{var}(X) = \frac{\alpha}{\lambda^2}$$

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}, t<\lambda$$

۳- توزیع وایبل: اگر متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد

$$f_{\alpha,\lambda}(x) = \alpha\lambda x^{\alpha-۱}e^{-\lambda x^{\alpha}} \qquad x>۰\,,\,\alpha,\lambda>۰$$

گوییم X دارای توزیع وایبل با پارامترهای α و λ است و آن را با نماد $X \sim \omega(\alpha,\lambda)$ نشان می‌دهیم که

$$E(X)=\frac{\Gamma(1+\frac{1}{\alpha})}{\lambda^{\frac{1}{\alpha}}}$$

$$\operatorname{var}(X)=\frac{\Gamma(1+\frac{2}{\alpha})-\Gamma^2(1+\frac{2}{\alpha})}{\lambda^{\frac{2}{\alpha}}}$$

۴- توزیع نرمال: اگر متغیر تصادفی X دارای چگالی احتمال زیر باشد

$$f_{\mu,\sigma^۲}(x)=\frac{۱}{\sqrt{۲\pi\sigma^۲}}\exp^{-\frac{۱}{۲\sigma^۲}(x-\mu)^۲} \qquad x\in R,\mu\in R,\sigma>۰$$

گوییم که X دارای توزیع نرمال با پارامترهای μ و σ^2 است و آن را با نماد $X \sim N(\mu,\sigma^۲)$ نشان می‌دهیم. حالت خاص $\mu = ۰$ و $\sigma^۲ = ۱$ به نرمال استاندارد معروف است و معمولاً تابع چگالی احتمال آن را با ϕ نشان می‌دهیم و در توزیع نرمال داریم

$$E(X)=\mu$$

$$\operatorname{var}(X)=\sigma^2$$

$$M_X(t)=e^{\mu t+\frac{\sigma^۲ t^۲}{۲}}$$

۵- توزیع لگ نرمال: اگر متغیر تصادفی X دارای تابع چگالی احتمال زیر باشد.

$$f_{\mu,\sigma^۲}(x)=\frac{۱}{x\sqrt{۲\pi\sigma^۲}}\exp^{-\frac{۱}{۲}\{(\ln x-\mu^۲\}} \qquad x>۰,\mu\in R,\sigma>۰$$

$$E(X)=\exp(\mu+\frac{1}{2}\sigma^2)$$

$$\operatorname{var}(X)=\omega(1-\omega)e^{2\mu} \qquad \omega=e^{\sigma^2}$$

۱-۲ تابع زیان و تابع مخاطره

شاید اساسی ترین انتظار از یک برآوردگر "خوب" برآوردی است که به مقداری واقعی به پارامتر نامعلوم Θ نزدیکتر باشد. به بیان دیگر، یک برآوردگر "خوب" برآوردگری است که میزان خطا در آن یعنی $\delta(x)-\theta$ با احتمال یک نزدیک به صفر باشد. اگر D نمایانگر کلاس تمام برآوردگرها باشد آنگاه برای هر $\delta(x) \in D$ فرض کنید که $\delta(x) \in \Theta$ باشد یعنی مقدار مشاهده شده $\delta(x)$ بر اساس یافته $X=x$ متعلق به Θ است. برای هر $\theta\in\Theta$ و هر برآورد ممکن $\delta(x)$ که متعلق به Θ است، مقدار عدد $L(\theta\ \ \delta(x))$ ای وجود دارد که میزان زیان آماردان را در برآورد پارامتر θ بر پایه $\delta(x)$ اندازه می‌گیرد. در حقیقت تابع زیان که آن را با L نشان دادیم، تابعی دو متغیره از $\Theta\times D$ به زیر مجموعه ای از اعداد حقیقی غیر منفی است، بنابراین تابع $L(\theta , \delta(X))$ خود یک متغیر تصادفی است.

در آمار، توابع زیان متعددی با توجه به نوع مسئله به کار گرفته می شود. توابع زیان معروف عبارت است از:

الف) تابع زیان مربع خطا : $L(\theta, \delta(x)) = (\delta - \theta)^2$

ب) تابع زیان وزنی مربع خطا : $L(\theta, \delta(x)) = w(\theta)\ (\delta - \theta)$

ج : تابع زیان خطی :

$$L(\theta,\delta(x)) = \begin{cases} K_0\ (\delta-\theta)\quad \theta-\delta>۰ \\ \\ K_1\ (\theta-\delta)\quad \theta-\delta<۰ \end{cases}$$

که k_0 و k_1 اعداد صحیح می باشند

د)تابع زیان قدر مطلق خطا: در حالت خاص با انتخاب $k1=k0=1$ در بالا تابع زیان قدر مطلق خطا را خواهیم داشت.

$L(\theta, \delta(x)) = | \delta - \theta |$

در برآورد پارامتر$\gamma(\theta)$، بر اساس برآوردگر$\delta(x)$، میزان دقت یا بهتر بگوییم عدم دقت برآوردگر با تابع مخاطره، یعنی

$$R(\theta, \delta) = E_\theta \{ L(\theta, \delta(x))\} \qquad (۱-۲-۱)$$

اندازه گرفته می شود. در مسائل برآوردیابی علاقه‌مند به دستیابی برآوردگری مانند δ هستیم که برای هر $\theta \in \Theta$، رابطه (۱-۲-۱) را مینیمم کند در حالت کلی مسئله دارای جواب نیست نتیجه منطقی این موضوع ایجاد محدودیت در کلاس برآوردگرها برای دستیابی به برآوردگرهای بهینه است، یک روش برای پیدا کردن برآوردگرهای بهینه، برقراری یک نوع رابطه ترتیبی بین برآوردگرهاست. دو روش اساسی در برقراری رابطه ترتیبی بین برآوردگرها، اصل مینیماکس و اصل بیز است که در این پایان نامه به بررسی اصل بیز می پردازیم. اصل بیز، برآوردگری از θ را بررسی می کند که برای یک تابع وزنی نظیر G، مقدار $\int R(\theta,\delta)dG(\theta)$ حداقل شود، یعنی بر اساس تابع وزنی G و تابع زیان L کلیه مقادیر $\int R(\theta,\delta)dG(\theta)$ را برای $\delta \in D$ از کوچک به بزرگ مرتب و برآوردگر δ ای را انتخاب می کند که مقدار این انتگرال از همه کوچکتر باشد و به انتخاب برآوردگر بیز $\delta_B(x)$ می‌انجامد به طوریکه

$$\int R(\theta,\delta)dG(\theta) \leq \int R(\theta,\delta_B)dG(\theta)$$

۱-۳ توزیع پشین و پسین

با توجه به ماهیت اصل بیز، در مسائل استنباط آماری به روش بیز بر اساس مشاهداتی که از خانواده توزیع‌ها اختیار می شود، پارامترθ دارای یک مقدار نامعلوم است. بعبارت دیگر، در حقیقت θ به عنوان یک مقدار متغیر تصادفی W در نظر گرفته می شود که مقادیر ممکن آن فضای پارامتر Θ را تشکیل می دهد و دارای توزیع$G(\theta)$ یا تابع چگالی احتمال (تابع احتمال) $g(\theta)$ است و از آن به عنوان توزیع پیشین یا تابع احتمال پیشین یاد می‌کنیم. در واقع توزیع پیشین تبلور کاربر آمار از خلاصه اطلاعات و دانسته های او در این باره است که احتمال قرار داشتن θ در چه زیر فضایی از Θ بیشتر است.

در این حالت تابع مخاطره بیزی که با نماد $r(G، \delta)$ نمایش می دهیم و انتخاب δ به عنوان یک برآوردگر با توزیع پیشین G و تابع زیان L که نقش مهمی دارد را به صورت زیر نشان می دهیم:

$$r(G, \delta) = E\{E(L(W, \delta(X))\}$$
$$= E[R(W, \delta)]$$
$$= \int R(\theta,\delta)dG(\theta) \qquad (۱-۳-۱)$$

و به دنبال برآوردگر δ_B هستیم که عبارت (۱-۳-۱) را مینیمم کند در این حالت از δ_B به عنوان برآوردگر بیز $\gamma(\theta)$ نسبت به توزیع پیشینG تحت تابع زیان L یاد می کنیم .

روش ماکسیمم درستنمایی و تحلیل بیز تا حدی به یکدیگر مربوط و وابسته‌اند. فرض کنید $L(\theta|x)$ یک تابع درستنمایی تحت برآورد ماکزیمم درستنمایی باشد که ما بزرگترین مقادیر L را تحت تابعی از متغیرθ و داده های x محاسبه می کنیم و همچنین از تابع درستنمایی برای ساختن فاصله اطمینان استفاده می‌کنیم همچنین برآورد ماکزیمم درستنمایی به تعداد نمونه‌های بزرگ (تعداد نمونه های زیاد) اتکا دارد و وقتی نمونه‌ها به اندازه کافی بزرگ باشد؛ آزمون میانگین و آزمون نسبت درستنمایی آن از توزیع χ^2 پیروی می‌کند اما این ویژگی لزوماً در مورد نمونه‌های کوچک صدق نمی‌کند. یک روش دیگر برای زمانیکه اطلاعات اولیه را داریم این است که یک توزیع با پارامترهای نامعلوم را حدس بزنیم مانند $p(\theta)$ و در توزیع بیز با توجه داده‌های درستنمایی و توزیع پیشین، توزیع پسین را تولید می‌کنیم.

$$\Pi(\theta \mid x) = (1/p(x)).\ p(x \mid \theta).p(\theta) \qquad (۲-۳-۱)$$

$$= (\text{توزیع احتمال پیشین}).(\text{تابع درستنمایی}).(\text{مقدار ثابت})$$

و می دانیم که $\Pi(x \mid \theta) = L(\theta \mid x)$ یعنی $\Pi(\theta \mid x)$ همان تابع درستنمایی است و $1/p(x)$ مقداری ثابت است. حال توزیع پسین را می توان به صورت زیر با تابع درستنمایی متناسب کرد.

$$\Pi(\theta \mid x) \propto L(\theta \mid x).p(\theta) \qquad (۳-۳-۱)$$

اگر بین توزیع پسین با توزیع پیشین داده‌ها وابستگی زیادی باشد به معنی آن است که داده‌ها اطلاعات کمی دارند و هرگاه این ارتباط کم باشد و درجه وابستگی بین پسین و پیشین کمتر می‌شود و تقریباً وابستگی وجود ندارد آنگاه داده‌ها دارای اطلاعات بیشتر و مفیدتری می‌باشد. حال اگر از رابطه (۱-۳-۳) لگاریتم بگیریم، خواهیم داشت:

$$\log(\Pi(\theta \mid x)) = \log(p(\theta \mid x)) + \log(L(\theta \mid x)) \qquad (۴-۳-۱)$$

۱- ۴ توزیع پسین حاشیه‌ای

غالبا فقط یک زیر مجموعه از پارامترهای نامعلوم به ما مربوط می‌شوند و بقیه پارامترهای مزاحم به کار ما وابسته نیستند. البته یک ویژگی بارز تحلیل بیز آن است که می توانیم اثر پارامترهای مزاحم را به وسیله انتگرال گیری از توزیع پسین حول پارامتری که مزاحم است از بین ببریم که به این انتگرال گیری، حاشیه‌ای پسین می‌گوییم. به طورمثال فرض کنید که میانگین و واریانس داده‌ها از توزیع نرمال با پارامترهای نامعلوم باشند؛ حال اگر پارامتر مورد نظر، واریانس باشد و میانگین یک پارامتر مزاحم باشد، در اینجا برآورد کردن میانگین کاری اضافه و بیهوده است. در حالیکه در آمار کلاسیکی، برآورد میانگین کاری استاندارد بوده اما در تحلیل بیز توسط توزیع حاشیه‌ای پسین برای σ^2 یعنی:

$$P(\sigma^2 \mid x) = \int p(\mu, \sigma^2 \mid x)\, d\mu \qquad (۱-۴-۱)$$

توزیع حاشیه‌ای پسین از یک توزیع توام بدست می‌آید که چندین پارامتر را درگیر می‌کند.

یک ویژگی بارز برای آنالیز بیزی، انتخاب یک توزیع پیشین است. اگر داده‌ها علائم خاص و نشانه‌های کافی داشته و دارای اطلاعات باشند حتی اگر یک پیشین بد داشته باشیم، اثر بزرگی روی پسین نخواهد داشت و اگر پسین نسبتاً استوار باشد، به این معنی است که اطلاعات داده‌ها، کافی و

کامل است. پارامترهای مکانی و مقیاسی از توزیع پیشین هایی می باشند که نسبت به پیشین‌هایی که دارای شکل انحنایی هستند، دارای ویژگی بهتری هستند و همچنین اغلب از پیشین‌هایی استفاده می‌شود که از تابع درستنمایی گرفته شده و در این صورت توزیع پسین همان توزیعی را خواهد داشت که توزیع پیشین دارد که به این نوع انتخاب، پیشین مزدوج می‌گویند.

از توزیع های معمول و مورد استفاده به عنوان توزیع پیشین می توان به توزیع یکنواخت و یا توزیع فاقد اطلاع اشاره کرد در واقع به عبارت ساده‌تر پیشین را به یک ثابت تبدیل می کند.

$$g(\theta) = k = 1/(b-a) \qquad a \le \theta \le b \qquad (۱-۴-۲)$$

در واقع با یک پیشین یکنواخت توزیع پسین با تابع درستنمایی که در یک مقدار ثابتی ضرب شده است، برابری می‌کند.

$$\Pi(\theta|x) = c\, L(\theta|x) \qquad (۱-۴-۳)$$

$$\Pi(\theta|x) \propto L(\theta|x) \qquad (۱-۴-۴)$$

در اکثر موارد صحبت‌های کلاسیک‌ها از آماره های فراوانی از تحلیل بیزی که دارای پیشین یکنواخت است، منشا می‌گیرد و اگر متغیر مورد نظر در بازه $(0, \infty)$ یا $(-\infty, \infty)$ قرار بگیرد؛ توزیع بدون اطلاع یا تابع پیشین یکنواخت وجود ندارد.

اگر در مواقعی حتی نتوانیم با درستنمایی تبدیل شده به راحتی مقیاس طبیعی را پیدا کنیم و یا درستنمایی قابل تجزیه نباشد، می‌توانیم از پیشنهادی که جفری[1] در سال ۱۹۶۱ ارائه کرده است استفاده کنیم که در واقع یک پیشین عمومی می‌دهد که در هر موردی کاربرد دارد. می دانیم که:

$$I(\theta|x) = -E(\partial^2 \ln(L(\theta|x)/\partial\theta^2) \qquad (۱-۴-۵)$$

$$P(\theta) \propto I(\theta|x)^{1/2} \qquad (۱-۴-۶)$$

مثال۲- اگر X متغیر تصادفی مستقل داشته باشیم که درستنمایی آن از توزیع دوجمله‌ای آمده باشد آنگاه داریم:

$l(\theta | x) = c\theta^{x} (1-\theta)^{n-x}$

و c مقدار ثابتی است که به پارامتر مربوط نیست و با گرفتن لگاریتم از طرفین داریم :

$L(\theta | x) = \ln(l(\theta | x)) = \ln c + x \ln\theta + (n-x)\ln(1-\theta)$

بنابراین $\partial L(\theta | x)/\partial\theta = x/\theta - (n-x)/(1-\theta)$

و داریم $\partial^{2} l(\theta | x)/\partial\theta^{2} = -(x/\theta^{2} + n-x/(1-\theta)^{2})$

و از آن جایی که $E(X) = n\theta$

$-E_x(\partial^2 \ln(L(\theta | x)/\partial\theta^2) = n\,\theta^{-1}(1-\theta)^{-1}$

$$P(\theta) \propto \theta^{-1/2}(1-\theta)^{-1/2} \qquad (۷-۴-۱)$$

که در واقع دارای توزیع بتا $Beta(1/2, 1/2)$ می باشد.

به طورمثال فرض می کنیم $x_1, \ldots\ldots\ldots, x_n$ داده ها یی از توزیع نرمال باشند و تابع درستنمایی برای i امین مشاهده x_i عبارت است از:

$L(\mu, \sigma^2 | x_i) = 1\times (2\pi\sigma^2)^{-1/2}\exp(-(x_i-\mu)^2/2\sigma^2)$

و برای n مشاهده داریم:

$L(\mu | x) = 1\times (2\pi\sigma^2)^{-1/2}\exp(-\sum(x_i-\mu)/2\sigma^2)$

واریانس معلوم و میانگین نامعلوم

فرض می کنیم وقتی واریانس معلوم و میانگین مجهول است، توزیع پیشین برای μ یعنی $p(\mu)$ با فرض تابع پیشین گوسین، به صورت زیر است:

$\mu \sim N(\mu_0, \sigma_0^2)$

$P(\mu) = 1\times (2\pi\sigma^2)^{-1/2}\exp(-(\mu-\mu_0)^2/2\sigma_0^2)$

و نکته مهم این است که هر گاه بخواهیم توزیع پسینی را به گونه بالا با پارامتر نامعلومی بدست آوریم به صورت زیر عمل می کنیم. به طور مثال فرض می کنیم در توزیع داده های x ، θ_1 برداری با پارامترهای معلوم و θ_2 برداری از پارامترهای نامعلوم و مجهول مدل باشد می توانیم توزیع پسین را به صورت زیر بنویسیم:

$$P(\Theta_2 | x,\Theta_1) = f(x,\Theta_1).g(x,\Theta_1,\Theta_2)$$

$$P(\Theta_2 | x,\Theta_1) \propto g(x,\Theta_1,\Theta_2)$$

در واقع تابعی از داده‌ها و پارامتر معلوم را به صورت یک ثابت در نظر می گیریم.

پس در بررسی داده های توزیع نرمال با میانگین نامعلوم داریم:

$$P(\mu | x) \propto l(\mu | x) . p(\mu)$$

$$\propto \exp(-(\mu-\mu_0)^2/2\sigma_0^2 - 1/2\sigma^2[\sum x^2_i - 2\mu nx+n\mu^2])$$

$$P(\mu| x) \propto \exp(-\mu^2/2\sigma_0^2+\mu\mu_0/ \sigma_0^2 +\mu nx/\sigma^2 - n\mu^2/2\sigma^2)$$

و در صورتی که از فاکتور بگیریم، خواهیم داشت:

$$-\mu^2/2(1/\sigma_0^2 +n/\sigma^2) +\mu(\mu_0/\sigma_0^2+nx/\sigma^2)$$

$$\sigma_*^2 =(1/\sigma_0^2 +n/\sigma^2)^{-1} \qquad \mu_* = \sigma_*^2(\mu_0/\sigma_0^2 +nx/\sigma^2)$$

$$= -\mu^2/\sigma_*^2 +2\mu\mu_*/2\sigma_*^2$$

و در نهایت خواهیم داشت:

$$P(\mu| x) \propto \exp(- (\mu - \mu_*)^2/2\sigma_*^2 +f(x, \mu_0 , \sigma^2,\sigma_0^2))$$

$$P(\mu| x) \propto \exp(-(\mu - \mu_*)^2/2\sigma_*^2)$$