دانشگاه پیام نور

پایان نامه

برای دریافت درجه کارشناسی ارشد

در رشته آمار ریاضی

دانشگاه پیام نور تهران

گروه علمی آمار

عنوان پایان نامه :

مقایسه روش‌های بیزی و احتمالی

استاد راهنما :

دکتر نرگس عباسی

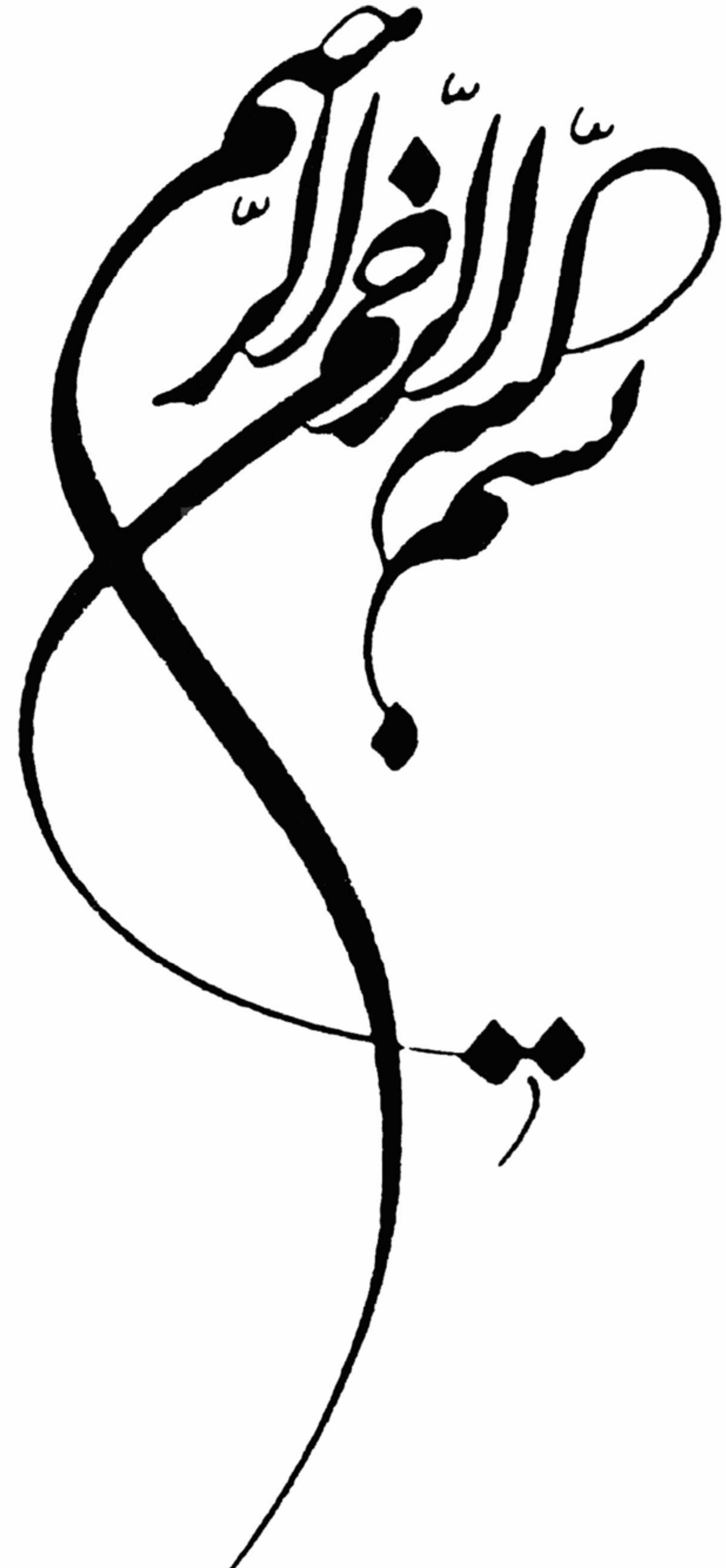
استاد مشاور :

دکتر مسعود یار محمدی

نگارش:

فهیمه السادات اعظمی‌پور

تابستان ۸۹


بسم الله الرحمن الرحیم

تقدیم به

اولین معلمان و فرشتگان زندگیم؛

مادر عزیز و فداکارم:

که نیمی از وجودش ایثار و گذشت کامل است

و نیمی دیگر عشق و محبت

پدر عزیز و بزرگوارم:

که دل دریایی اش آموزگار چگونه زیستنم بود.

فهرست

چکیده

مقایسه بین روش‌های برآورد بیزی و احتمالی از موضوعات جالب و چالش‌زای علم آمار به شمار می‌رود. در این پایان‌نامه برآوردها و پیش‌بینی کننده‌های بیز از یک توزیع نرمال گرفته می‌شوند. متداول‌ترین پیش‌بینی کننده‌ی احتمالی برآورد ماکسیمم درست‌نمایی دارای یک روند درونی است، می‌توان ذکر کرد که با جایگزینی برآورد ماکسیمم درستنمایی برای میانگین جامعه، توزیع پیش‌بینی کننده به‌دست می‌آید. این پایان‌نامه مشتمل بر چهار فصل است. فصل اول و فصل دوم را مفاهیم و برخی قضایای مفیدی از نظریه تصمیم را در بردارد. فصل سوم مسئله‌ی برآوردیابی میانگین جامعه‌ی نرمال در حالت یک متغیره و در فصل چهارم مسئله در جامعه‌های نرمال چندمتغیره، تحت موضوع مدل‌های خطی بررسی می‌کنیم.

واژه‌های کلیدی:

برآورد بیز، چگالی پیشین جفریز، تابع زیان، برآوردگر ماکسیمم درستنمایی، تابع مخاطره.

پیشگفتار

در سال‌های آخر نیمه‌ی اول سده‌ی بیستم، ابراهام والد[1]، اصول تصمیم‌گیری را در چهارچوبی ریاضی ارائه داد و کوشید که مباحث گوناگون آماری، مانند برآوردیابی، آزمون‌ها و فاصله‌ی اطمینان را به هم پیوند دهد و از این راه، نظریه‌ی آمار را پربارتر سازد. در آمار کلاسیک به کمک نمونه‌های آماری و داده‌ها درباره‌ی پارامتر یا توزیع مجهول، به دستیابی می‌پردازیم و نتیجه‌گیری می‌کنیم، اما در نظریه‌ی تصمیم، به کمک نمونه‌های آماری و دیگر جوانب، برای نمونه، سود و زیان، می‌کوشیم تا به بهترین تصمیم‌گیری دست یابیم (۱).

طی بیست سال گذشته موضوع مقایسه میان رویکردهای بیزی و احتمالی تحت توابع زیان هموار با مقادیر معلوم، مانند تابع زیان مربع خطا، توجه زیادی را به خود جلب کرده است. به‌کارگیری توابع زیان متقارن، مانند زیان مربع خطا، برای بسیاری از مسائل عملی مؤثر است. این توابع در بیشتر موارد، گزینه‌هایی منطقی به شمار می‌روند. مطالعات برگر[2] ۱۹۸۰، مارتز و والر[3] ۱۹۸۲، شینها و کیل[4] ۱۹۸۰ را ذکر نمود. با این وجود ممکن است برای برخی از مسائل ارزیابی و پیش‌بینی مناسب نباشند. در این موارد مشاهده شده است که توابع نامتقارن مؤثرتر بوده‌اند. برای مطالعه‌ی بیشتر می‌توان از تحقیقاتی که توسط واریان[5](۱۹۷۵)، زلنر[6](۱۹۸۶)، مورهد[7] و یو[8] (۱۹۹۸)، اسپرینگ[9] ویونگ[10] (۱۹۹۸)، چاندرا[11](۲۰۰۱)، بهره جست.

[1] Abraham Wald, 1902-1950
[2] Berger
[3] Martz and Waller
[4] Shinha and Kale
[5] Varian
[6] Zellner
[7] Moorhead
[8] Wu
[9] Spiring
[10] Yeung
[11] Chandra

برای مثال جوزف[1](۲۰۰۴)برای ضخامت پوشش لحیم، صفحه‌ی مدارهای چاپی را در نظر گرفت و کشف کرد که تابع متقارن زیان جهت اندازه‌گیری زیان کلی برای مصرف‌کنندگان و تولیدکنندگان مناسب نیست. تا به حال تصور بر این بود که روش‌های بیز همواره و در هر شرایطی خوب عمل می‌کنند اما این تصور منطقی و صحیح نیست.

در اینجا پیش‌بینی بیز[2] را تحت تابع زیان $\alpha-$ خطای مطلق، تابع زیان لینکس، تابع زیان آنتروپی، به عنوان یک مورد خاص از زیان‌ها، $\alpha-$ خطای مطلق را مورد بررسی قرار می‌دهیم. چنانچه واریانس نامعلوم باشد، از مقدار ترکیبی مشترک قبلی جهت برآورد نامعلوم برای زیان‌های $\alpha-$ خطای مطلق استفاده می‌شود.

سنجری (۱۳۷۳)، برآورد پارامتر مکانی، با تکیه بر قید پایایی، تحت تابع زیان لینکس بررسی نمود. وی با تشریح مسئله‌ی برآورد میانگین جمعیت‌های انتخابی، مجموعه‌ای از برآوردگرها، با استفاده از تابع زیان لینکس برای دو جمعیت نرمال پیشنهاد نمود که همگی مکان–پایا بودند. این برآوردگرها از نقطه نظر اریبی و ریسک مقایسه شدند.

به این ترتیب که واریانس نمونه برای زیان‌های لینکس جایگزین واریانس می‌شود. برآوردهای بیز نیز به ترکیبات خطی ضریب‌های رگرسیون، قابل گسترش است. در شرایط فرضی خاص در ماتریس طرح، زیان‌های مجانب مورد انتظار استخراج (مشاهده) می‌شوند.

چنانچه مقادیر پیشین مناسب باشد، ارزیابی بیز و عامل پیش‌بینی کننده بهتر ازبرآورد تنهایی (برآورد ماکسیمم درستنمایی) عمل می‌کنند. در شرایط تابع زیان لینکس برآورد بیز با مقدار پیشین جفریز نسبت به برآورد ماکسیمم درستنمایی ارجحیت دارد. با این وجود، روشن نیست که کدامیک از موارد بیز یا حداکثر درستنمایی در زمینه پیش‌بینی بهتر عمل می‌کنند. در برخی شرایط حتی وقتی که زیان از نوع تابع زیان حقیقی باشد ارزیابی بیز در شرایط زیان دیگر بهتر از شرایط زیان حقیقی عمل می‌کند. این نکته هشداری برای طرفداران بی‌تجربه بیز است که تصور می‌کنند روش‌های بیز همواره و در همه‌ی شرایط بهتر از روش‌های دیگر عمل می‌کنند.

---

[1] Joseph

[2] Bayes

فصل اول
معرفی توابع زیان و
اصطلاحات نظریه‌ی تصمیم

۱–۱ مقدمه

نظریه‌ی تصمیم، همانگونه که از نام آن برمی‌آید، درباره‌ی تصمیم‌گیری؛ یعنی یکی از کارهای عمده‌ی روزانه‌ی انسان‌ها است. انسان‌های مسئول و منطقی، آنها که دارای عقل سلیم هستند، با علم و آگاهی بسنده، با در نظرگرفتن احتمال پیشامدها، با توجه به سود و زیان و یا شهرت و رسوایی، تصمیم می‌گیرند و کاری را انجام می‌دهند. در این فصل، برخی از این واژه‌ها را به عنوان اصطلاحات نظریه‌ی تصمیم، شرح می‌دهیم.

۱–۲ تابع تصمیم

اغلب بدون توجه به پارامتر مجهول، θ، تصمیم گیری می‌کنیم. اما عقل سلیم به ما حکم می‌کند که در صورت امکان، نخست بررسی آماری انجام داده تا بتوانیم درباره‌ی θ آگاهی بیابیم و از این راه تصمیم گیری نماییم.

به طور کلی، نتیجه‌ی هر بررسی آماری را که در نظر می‌گیریم با بردار تصادفی

$$\mathbf{X} = (X_۱, X_۲, \ldots, X_n)$$

نشان می‌دهیم. چگالی توأم این بردار باید به θ بستگی داشته باشد تا بتوان درباره‌ی θ آگاهی یافت. فرض می‌کنند $X_۱, X_۲, \ldots, X_n$ مستقل و هم‌توزیع است.

تعریف (تابع تصمیم). تابعی با دامنه‌ی O_n (فضای یافته‌ها) و برد A (فضای کارها) است. این تابع را با $d(\mathbf{x})$ نشان می‌دهیم. بنابراین، برای هر $\mathbf{x} \in O_n$ به یک کار a می‌پردازیم. یعنی داریم:

$$d(\mathbf{x}) = a \in A.$$

از نظر آمارریاضی، $d(\mathbf{x})$ یافته‌ی آماره‌ی $d(\mathbf{X})$ است. هنگامی که θ پارامتر یک توزیع و $d(\mathbf{x})$، برآورد آن باشد، کار a می‌شود گزینش $d(\mathbf{x})$ به عنوان برآورد θ.

۱–۳ انواع توابع زیان

هنگامی که با مشاهده‌ی $\mathbf{X}$ برای انجام کار a تصمیم می‌گیریم، ممکن است با زیان روبرو شویم. این زیان رابا یک تابع دو متغیری غیرمنفی از θ و $a = d(\mathbf{x})$ بیان می‌شود. چنین تابعی را که می‌تواند به صورت‌های گوناگون باشد، تابع زیان می‌نامند.

تعریف (تابع زیان). تابعی دومتغیری، کراندار و نامنفی با دامنه‌ی $T \times A$ (حاصلضرب دکارتی فضای اوضاع طبیعت و فضای کارها) است که آن را با $L(\theta, a)$ نشان می‌دهند.

طبیعی است که اگر دو نفر درباره‌ی امری، دو تابع زیان متفاوت داشته باشند، ممکن است به تصمیم گیری‌های متفاوت بپردازند؛ در حالی که هر دو، خردمندانه تصمیم گیری می‌کنند.

هنگامی که پارامتر مجهولی را برآورد می‌کنیم؛ معمولاً فضای پارامتر و فضای کارها را مساوی برمی‌گزینیم؛ یعنی $T = A$. در این مورد، $a \in A$، به عنوان برآورد $\theta \in T$ است. هرچه a و θ به هم نزدیکتر باشند، زیان ناشی از انجام کار a، کمتر خواهد بود. البته تابع زیانی مناسب باید این امر را بنمایاند. برای نمونه، در توزیع دوجمله‌ای θ را احتمال پیروزی فرض می‌کنیم. در این مثال داریم:

$$T = A = (0\,,1)$$

اینک تابع زیان

$$L(\theta\,,a) = |\theta - a|$$

را در نظر می‌گیریم. آشکار است که این تابع، منظور ما را برآورده می‌کند. چنین توابعی بی‌مانند نیستند. اکنون برخی از آنها را در این پایان نامه مطرحند را ارائه می‌دهیم.

تابع زیان درجه‌ی دو

تابع زیان

$$L(\theta\,,a) = (\theta - a)^2$$

را تابع زیان درجه‌ی دوم، یا توان دوم خطا می‌نامند. این تابع برای برآوردیابی، کاربرد فراوان دارد، با اینکه به عنوان تابع زیان ممکن است کراندار نباشد.

تابع زیان α- خطای مطلق

تابع زیان

$$L(\theta\,,a) = |\theta - a|^{\alpha}$$

را تابع زیان α- خطای مطلق می‌نامند. واضح است که برای $\alpha = 2$ تابع زیان با تابع زیان درجه دو برابر خواهد شد.

تابع زیان صفر و یک

فرض کنید که $A=\{a_\circ, a_1\}$، فضای کارها باشد. به عنوان مثال، برای آزمون $H_\circ: \theta=\theta_\circ$ در برابر $H_1:\theta=\theta_1$ ($T=\{\theta_\circ, \theta_1\}$ فضای پارامتر است.)، کار $a_\circ$ را پذیرفتن فرض $H_\circ$ و کار **Error! Objects cannot be created from editing field codes.** را پذیرفتن فرض H_1 می‌گیریم.

تابع زیر را که برای $a_\circ$ و a_1 به صورت زیر تعریف شده است تابع زیان صفر و یک می‌نامند:

$$L(\theta, a_\circ)=\begin{cases}\circ & \theta=\theta_\circ \\ 1 & \theta=\theta_1\end{cases} \qquad\qquad L(\theta, a_1)=\begin{cases}\circ & \theta=\theta_1 \\ 1 & \theta=\theta_\circ\end{cases}$$

با این تابع اگر کار درست انجام دهیم، زیان صفر است و اگر کار نادرست انجام دهیم، زیان یک است. برای نمونه، اگر $a_\circ$ را انجام دهیم. یعنی بپذیریم که $\theta=\theta_\circ$، اما در حقیقت $\theta=\theta_1$، آنگاه زیان یک است.

تابع زیان خطی نمایی (لینکس)

این تابع به صورت زیر است:

$$L(\theta, a)=b[e^{c(a-\theta)}-c(a-\theta)], \qquad b>\circ, \quad c\neq\circ$$

که در آن b پارامتر مقیاس و a پارامتر شکل است. از آنجا که در این تابع سازه‌ی خطی و نمایی هر دو وجود دارند، از این رو آن را تابع زیان لینکس به معنی خطی نمایی می‌نامند. این تابع نسبت به a محدب اکید است. زیان وارد با این تابع برای بیش برآوردی محسوس‌تر از کم برآوردی است، یعنی تابع نسبت به این دو امر نامتقارن است. شاید در برخی موارد بر تابع زیان درجه‌ی دو و تابع خطای مطلق رجحان داشته باشد که هر دو متقارن هستند.

واریان (۱۹۷۵) این تابع زیان مفید را تابع زیان لینکس معرفی کرد، با این ویژگی که برآمده از تابعی بود که تقریباً در یک طرف صفر به صورت نمایی و در طرف دیگر عدد صفر به صورت خطی ظاهر می‌شوند. زلنر(۱۹۸۶) نیز کاربردهای جالبی از توابع زیان لینکس را معرفی کرد.

۴-۱ تابع مخاطره

تابع تصمیم d و تابع زیان $L(\theta, a)$ را در نظر می‌گیریم که هر دو از پیش مشخص شده‌اند. اگر با مشاهده‌ی $\mathbf{x}$، به کار $a=d(\mathbf{x})$ تصمیم بگیریم و وضع طبیعت θ باشد، مقدار زیان عدد $L(\theta, d(\mathbf{x}))$ است. با تبدیل $\mathbf{x}$ به $\mathbf{X}$، این عدد به متغیر تصادفی $L(\theta, d(\mathbf{X}))$ تبدیل می‌شود. در حقیقت پیش از انجام آزمایش، زیان به صورت این متغیر تصادفی است. بنابراین

باید امیدریاضی این متغیر تصادفی را محاسبه کرد، تا زیان به طور متوسط به دست آید. آشکار است که این امید ریاضی فقط به θ بستگی دارد، یعنی تابعی از θ است. چنین تابعی را تابع ریسک و یا تابع مخاطره می‌گویند.

تعریف (تابع مخاطره)

برای هر تصمیم مشخص d، تابع مخاطره عبارت است تابعی نامنفی و پایاندار با دامنه‌ی T و برد R که آن را با $R_d(\theta)$ نشان می‌دهند و به صورت زیر تعریف می‌کنند:

$$R_d(\theta) = E[L(\theta, d(\mathbf{X}))]$$

چگالی مربوط به امید بالا، چگالی بردار $\mathbf{X}$ و یا چگالی آماره‌ی $d(\mathbf{X})$ بوده که معمولاً به پارامتر مجهول θ بستگی دارد.

واضح است تابع $R_d(\theta)$ به تابع زیان و تابع تصمیم بستگی دارد و با تغییر هر کدام تغییر می-کند.

۱-۵ چگالی پسین و پیشین و برآورد بیز

در اینجا پارامتر θ را به چشم متغیری تصادفی با چگالی $\pi(\theta)$، به نام چگالی پیشین، می-نگرند. اطلاعات نهفته در $\mathbf{X}$ و $\pi(\theta)$ استفاده می‌شود تا در چگالی پیشین تجدید نظر شده و نوعی چگالی برای θ به نام چگالی پسین تعیین شود.

فرض کنید $\pi(\theta)$ چگالی پیشین θ، $f(\mathbf{x},\theta)$ چگالی توام $\mathbf{X}$ و θ، $f(\mathbf{x}|\theta)$ چگالی $\mathbf{X}$ به شرط θ، $m(\mathbf{x})$ چگالی کناری $\mathbf{X}$ باشند. چگالی پسین θ عبارت است از چگالی θ به شرط $\mathbf{x}$.

$$\pi(\theta \mid \mathbf{x}) = \frac{f(\mathbf{x}, \theta)}{m(\mathbf{x})}.$$

چگالی کناری $\mathbf{X}$ در حالت پیوسته از انتگرال زیر به دست می‌آید:

$$m(\mathbf{x}) = \int_T f(\mathbf{x}, \theta) d\theta = \int_T f(\mathbf{x} \mid \theta)\, \pi(\theta)\, d\theta.$$

حال از روی چگالی پسین، می‌توان از تابع مخاطره امیدریاضی گرفت و مقادیر دیگری به دست آورد.

تعریف (مخاطره بیزی یا ریسک بیز)

اگر $R_d(\theta)$ تابع مخاطره برای تصمیم d و $\pi(\theta)$ چگالی پیشین θ باشد، آنگاه مخاطره بیز برای d با چگالی پیشین $\pi(\theta)$ عبارت است از:

$$r_{d,\pi} = E[R_d(\theta)].$$

تابع مخاطره در یک عدد به نام مخاطره بیز فشرده شده است. چگالی پیشین در حالی که باور شخصی را بر مبنای احتمال شخصی درباره‌ی طبیعت می‌نمایاند، به خودی خود در نظریه‌ی تصمیم نوعی تابع وزنی برای تعیین ریسک بیز به شمار می‌آید. این تابع ممکن است به برخی θها وزن بیشتری دهد تا به برخی دیگر.

اگر میزان را بر مخاطره بیز استوار کنیم، طبیعی است هر تصمیمی که برای آن مخاطره بیز کمتر باشد، بهتر است. بنابر این تعریف زیر بخردانه به نظر می‌آید:

تعریف (تصمیم بیز)

فرض کنید D مجموعه‌ی تصمیم‌ها و $\pi(\theta)$ چگالی پیشین مشخصی باشد. تصمیم $d \in D$ را یک تصمیم بیز نسبت به $\pi(\theta)$ می‌نامیم، هرگاه برای هر $d \in D$ داشته باشیم:

$$r_{d_b,\pi} < r_{d,\pi}$$

به عبارت دیگر، d_b را تصمیم بیز می‌گوییم هرگاه کمترین ریسک بیز را دارا باشد، یعنی

$$r_{d_b,\pi} = \min_{d \in D} r_{d,\pi}$$

تصمیم بیز ممکن است یکتا نباشد. گذشته از این، در حالتی که D بی‌پایان است ممکن است تصمیم بیز وجود نداشته باشد.

۱–۶ توزیع و توابع زیان مورد مطالعه

از آنجا که توزیع نرمال نقش مهمی درآمار به‌عهده دارند، در این پایان نامه، به مطالعه ارزیابی و پیش‌بینی پارامترهای آن می‌پردازیم.

فرض کنید که $X_n = (X_1, \ldots, X_n)$ یک نمونه تصادفی از $N(\mu, \sigma^2)$ است و $\sigma^2 > 0$ در نظر گرفته می‌شود و $\mu \in R$ یک پارامتر نامعلوم است. در اینجا هدف برآورد میانگین نامعلوم μ است. ابتدا گروهی از توابع زیان α- خطای مطلق را برای برآورد $\delta = \sigma(X_n)$ در نظر می‌گیریم:

$$L_1(\mu, \delta) = |\mu - \delta|^{\alpha} \qquad (۱)$$

در اینجا α یک مقدار ثابت معلوم و مثبت است. برای $\alpha=۲$ تابع فوق به صورت تـابع زیـان مربع خطا می‌شود. و برای $\alpha=۱$، مطابق با تابع زیان خطای مطلق می‌شود.

حال دسته‌ای از توابع زیان لینکس را در نظر می‌گیریم.

$$L_۲(\mu,\delta)=e^{c(\delta-\mu)}-c(\delta-\mu)-۱ \qquad (۲)$$

در اینجا c یک مقدار ثابت غیرصفر است.

یکی دیگر از توابع زیان که کاربرد فراوان دارد، تابع زیان آنتروپی است (رجوع کنید به روبرت[۱](۱۹۹۷)). چنانچه $f(.|\theta)$ و $f(.|\gamma)$ توزیع‌های مرتبط با پارامترحقیقی θ و برآورد γ باشند. آنگاه فاصله‌ی آنتروپی به‌صورت زیر تعریف می‌شود:

$$L_۳(\theta,\gamma)=\int_{-\infty}^{+\infty}\log\left\{\frac{f(x|\theta)}{f(x|\gamma)}\right\}f(x|\theta)dx. \qquad (۳)$$

برای مثال در توزیع نرمال، به راحتی می‌توان نشان داد که

$$L_۳(\theta,\gamma)=\int_{-\infty}^{+\infty}\log\{\frac{\frac{۱}{\sqrt{۲\pi\sigma^۲}}e^{-\frac{۱}{۲\sigma^۲}(x-\theta)^۲}}{\frac{۱}{\sqrt{۲\pi\sigma^۲}}e^{-\frac{۱}{۲\sigma^۲}(x-\gamma)^۲}}\}\frac{۱}{\sqrt{۲\pi\sigma^۲}}e^{-\frac{۱}{۲\sigma^۲}(x-\theta)^۲}dx$$

$$=E_\theta(\frac{۱}{۲\sigma^۲}(X-\gamma)^۲-\frac{۱}{۲\sigma^۲}(X-\theta)^۲)$$

$$=\frac{(\gamma-\theta)^۲}{۲\sigma^۲}$$

با تغییری جزئی و با معلوم بودن مقدار $\sigma^۲$ تابع آنتروپی به صورت تابع زیان مربع خطا در می‌آید.

$$L_۳(\mu,\gamma)=\frac{(\gamma-\mu)^۲}{\sigma^۲}. \qquad (۴)$$

در این پایان نامه سه تابع زیان در خانواده‌ی توزیع نرمال مورد مطالعه قرار می‌گیرند. که نتایج آن‌ها برخلاف این تصور ظاهر شده است. برای مثال، برآوردگر بیزی $\hat{\mu}_B$ که به عنوان برآوردگر بیز تحت تابع زیان لینکس، $L_۲$، محاسبه شده است با μ_n یا برآوردگر بیزی که

---

[۱]Robert

تحت تابع زیان α – خطای مطلق، L_1، حاصل شده مقایسه می‌شود. در برخی شرایط حتی وقتی L_1 تابع زیان حقیقی است، $\hat{\mu}_B$ بهتر از μ_n عمل می‌کند. بنابراین این نوع نتایج، هشداری برای طرفداران بیز به شمار می‌روند.

فصل دوم

توابع مخاطره

۲-۱ مقدمه

در این فصل، مقدار مخاطره برای برآورد ماکسیمم درستنمایی μ، مقدار مخاطره پیش‌بینی کننده با استفاده از برآوردهای درونی به‌دست می‌آوریم. در L، برآورد μ به روش بیز را با مقدار ترکیبی تواماً (μ,σ^2) به‌دست می‌آوریم. این مرحله معلوم یا نامعلوم بودن σ^2 را نیز در نظر می‌گیریم. همچنین پیش‌بینی توزیع نرمال را نیز که تخمینی از پیش‌بینی بیز در شرایط L است با شرط اینکه مقدار اختلاف σ^2 با مقدار ترکیبی پیشین μ معلوم باشد، به‌دست می-آوریم.

۲-۲ مخاطره تحت تابع زیان α - مطلق خطای

******

۲-۲-۱ برآورد حداکثر درستنمایی

فرض کنید که $X_1,\ldots,X_n$ مشاهدات مستقل و یکسان از توزیع نرمال $N(\mu,\sigma^2)$ با مقدار واریانس معلوم، σ^2، هستند.

برآورد ماکسیمم درستنمایی برای میانگین جامعه‌ی نرمال، میانگین نمونه، $\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$ است. برای محاسبه‌ی مقدار مخاطره‌ی برآوردگر احتمالی برای برآورد ماکسیمم درستنمایی قضیه‌ی زیر آورده شده است.

قضیه ۱

میانگین زیان تحت تابع زیان α- خطای مطلق برابر است با

$$E|\overline{X}-\mu|^{\alpha}=\sqrt{\frac{2^{\alpha}}{\Pi}}.\frac{\sigma^{\alpha}}{n^{\frac{\alpha}{2}}}.\Gamma\left(\frac{\alpha+1}{2}\right). \qquad (5)$$

اثبات:

می‌دانیم که $\overline{X}_n\sim N\left(\mu,\frac{1}{n}\sigma^2\right)$ است بنابراین

$$E|\bar{X}-\mu|^{\alpha} = \int_{-\infty}^{+\infty} |y|^{\alpha} \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}} .\exp\left(\frac{-n}{2\sigma^2} y^2\right) dy$$

$$= 2\int_{0}^{+\infty} y^{\alpha} \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}} .\exp\left(\frac{-n}{2\sigma^2} y^2\right) dy$$

در اینجا برای انجام محاسبه‌ی ساده‌تر تغییر متغیرهای زیر را در نظر می‌گیریم

$$w = \frac{ny^2}{2\sigma^2}$$

$$y = \sqrt{\frac{2\sigma^2 w}{n}}$$

$$dy = \sqrt{\frac{\sigma^2}{2nw}} dw$$

حال مقادیر مناسب را جایگزین می‌کنیم و عباراتی را که به متغیر درون انتگرال مربوط نیستند از انتگرال بیرون می‌آوریم.

$$E|\bar{X}-\mu|^{\alpha} = 2\sqrt{\frac{n}{2\pi\sigma^2}} \int_{0}^{+\infty} \sqrt{\frac{\sigma^2}{2nw}} \left(\frac{2\sigma^2}{n}\right)^{\alpha/2} w^{\frac{\alpha}{2}} e^{-w} dw$$

$$= 2\sqrt{\frac{n}{2\pi\sigma^2}} \sqrt{\frac{\sigma^2}{2n}} \left(\frac{2\sigma^2}{n}\right)^{\alpha/2} \underbrace{\int_{0}^{+\infty} w^{\frac{\alpha}{2}-\frac{1}{2}} e^{-w} dw}_{\Gamma(\frac{\alpha+1}{2})}$$

$$= \sqrt{\frac{2^{\alpha}}{\pi}} \frac{\sigma^{\alpha}}{n^{\frac{\alpha}{2}}} \Gamma\left(\frac{\alpha+1}{2}\right).$$

و اثبات کامل می‌گردد.

پارامتر مهم دیگری که در اکثر مباحث استنباطی مورد توجه قرار می‌گیرد برآورد کمیت $\Phi\left(\frac{x-\mu}{\sigma}\right)$ است. این موضوع را این چنین آغاز می‌کنیم: به توزیع پیش‌بینی کننده یک مشاهده‌ی جدید با نماد X^* توجه می‌کنیم.

$$\eta(x;\mu) \equiv P(X^* \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \qquad x \in R$$

که در آن Φ تابع توزیع نرمال استاندارد است. با توجه به خاصیت پایایی برای به‌دست آوردن برآورد حداکثر درستنمایی $\eta(x;\mu)$، می‌توان برآورد حداکثر درستنمایی μ را جایگزین کرد. این برآورد بر مبنای یک برآوردیابی درونی است که در بسیاری از مسائل به تابع زیان ۱–۰ (صفر–یک) مربوط است.

$$\hat{\eta}_n^M = \Phi\left(\frac{x-\overline{X}_n}{\sigma}\right).$$

قضیه ۲

مقدار مخاطره‌ی پیش بینی شده در استفاده از برآورد ماکسیم درستنمایی برای μ برابر است با

$$E\left|\hat{\eta}_n^M - \eta\right|^\alpha = \sqrt{\frac{2^\alpha}{\pi}}\,\frac{\phi(v)^\alpha}{n^{\frac{\alpha}{2}}}\,\Gamma\left(\frac{\alpha+1}{2}\right)\left[1+\frac{\alpha(\alpha+1)\left\{(3\alpha+1)v^2-4\right\}}{24n}\right] + O\left(n^{-2-\frac{\alpha}{2}}\right),$$

که در آن $v = \frac{x-\mu}{\sigma}$ و Φ تابع چگالی توزیع نرمال استاندارد است.

اثبات این قضیه در ضمیمه ب آورده شده است.

۲–۲–۲ برآورد بیز میانگین

فرض کنید $X_1,\ldots,X_n$ مشاهدات مستقل و هم‌توزیع از توزیع نرمال $N(\mu,\sigma^2)$ با مقدار واریانس معلوم، σ^2، هستند و چگالی پیشین مزدوج برای μ را به کار ببریم یعنی μ دارای توزیع $N(\mu_0,\tau_0^2)$ باشد به‌طوری‌که $\mu_0 \in R$ و $\tau_0^2 > 0$ معلوم هستند، آنگاه μ دارای توزیع پسین $N(\mu_n,\sigma_n^2)$ است که در آن

(۷)

$$\mu_n = \frac{\frac{n\overline{X}_n}{\sigma^2}+\frac{\mu_0}{\tau_0^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau_0^2}} = \frac{n\overline{X}_n\tau_0^2+\sigma^2\mu_0}{n\tau_0^2+\sigma^2} = \left(1+\frac{\sigma^2}{n\tau_0^2}\right)^{-1}\left[\overline{X}_n+\frac{\sigma^2\mu_0}{n\tau_0^2}\right],$$