بسم الله الرحمن الرحيم

٢٤٨١٥

مرکز اطلاعات مدارک علمی ایران
ثبت مدارک

۱۳۷۸ /۴/ ۲۰

دانشگاه فردوسی مشهد

دانشکده علوم

پایان نامه

" برای اخذ درجه کارشناسی ارشد آمار ریاضی "

عنوان

برآورد و بررسی مجاز بودن برآورد تحت زیان نامتقارن Linex

استاد راهنما

آقای دکتر ناصر رضا ارقامی - دانشیار گروه آمار دانشگاه مشهد

استاد مشاور

خانم دکتر ناهید سنجری - استادیار گروه آمار دانشگاه شیراز

استاد داور:

آقای دکتر ابوالقاسم بزرگ نیا - استاد گروه آمار دانشگاه مشهد

نگارش:

عباس رسولی

زمستان ۱۳۷۴

۲۴۸۱۰

1819/2

شماره
تاریخ
پیوست

دانشگاه فردوسی مشهد

دانشکده علوم

دانشگاه فردوسی " مشهد"

دانشکده علوم

گروه آمار

صورتجلسه دفاع رساله کارشناسی ارشد آمارریاضی

درتاریخ ۷۴/۱۰/۷ خانم / آقای عباس رسولی از رسالهٔ کارشناسی ارشد خود

تحت عنوان :

" برآورد و بررسی مجازبودن برآوردها تحت زیان نامتقارن Linex "

با بیان خلاصه ای از کار انجام شده و پاسخ به سئوالات داوران دفاع نمودند و

این رساله با نمره ۱۹/۵ معادل عالی ~~قبول~~ شد.

۱ ـ استادراهنما آقای دکتر ناصررضا ارقامی

۲ ـ اعضاء هیئت داوران ۱ ـ خانم دکتر ناهید سنجری (استاد مشاور) عضوهیأت علمی دانشگاه شیراز

۲ ـ آقای دکتر ابوالقاسم بزرگ نیا (استاد داور)

۳ ـ

مدیر گروه آمار

دکترغلامحسین شاهکار

معاون آموزشی دانشکده

«تقدیم به پدر گرامی و مادر مهربانم»

پدرم که چون او پدری را در جهان سراغ ندارم، پدری که مظهر محبت و عشق است

مادرم که چشمان مهربان او آشیانه آسایشم و یاد او امیدبخش زندگیم است

و

یگانه خواهرم که تشویق او پشتوانه ادامه راهم بود

«فهرست مطالب»

عنوان **صفحه**

عنوان صفحه

عنوان	صفحه

«مقدمه»

یکی از دیدگاههای آمار استنباطی رهیافت نظریه تصمیم است به این معنی به هر یک از شاخه‌های استنباط به صورت یک مسئله تصمیم نگریسته می‌شود. معمولاً زمانی که برآورد یک پارامتر مجهول مطرح است تابع زیانی که تابعی است از تفاوت برآورد و مقدار واقعی پارامتر مورد استفاده قرار می‌گیرد و زمانی که آزمون فرضهای آماری مطرح است تابع زیانهای دیگر از قبیل تابع زیان "$\kappa - \circ$" مورد استفاده قرار می‌گیرد. بنابراین یک ابزار مهم انتخاب تابع زیان مناسب است. از معروف‌ترین توابع زیان که هنگام برآورد کردن در نظریه تصمیم آماری مورد استفاده قرار می‌گیرند می‌توان " زیان میانگین مربع خطا "، " زیان قدر مطلق خطا " را نام برد.

گرچه این توابع زیان به‌طور نسبی از مفاهیم ریاضی ساده‌ای برخوردارند و برای عملیات دستی مفید و آسان می‌باشند ولی در اغلب موارد نامناسب به نظر می‌رسند. برای مثال زیان مربع خطا نتیجه می‌دهد که زیان مربوط به دوری برآورد از هدف واقعی تنها به توان دوّم خطای برآورد مربوط است و نه به انحراف‌های مثبت و منفی برآوردگر و ...

به عنوان مثال در یک مدار بسته تولید کالا که هدف برآورد نسبت خطا می‌باشد اگر برآورد

بزرگتر از مقدار واقعیش باشد، پروسه متحمل هزینه بالایی برای بهبودی تولید خواهد شد، برعکس اگر برآورد کمتر از مقدار واقعیتش باشد فروشندگان سهم بازار را از دست خواهند داد و این بدان علت است که درجه اعتماد کالا دورتر از مقداریست که ادعا شده است و گاهی ممکن است همین علت باعث نابودی شرکت گردد. پس ملاحضه می‌گردد که در این موارد تنها قدر مطلق خطا مهم نیست و باید علاوه بر مقدار مطلق خطا به علامت آن نیز توجه داشت. به همین علت کلاس توابع زیان نامتقارن مطرح می‌گردند که یکی از آنها تابع زیان نامتقارن *LINEX* است. این تابع نخستین بار توسط *Varian (1975)* معرفی گردید. اما *Zellner (1986)* به بحث و تحقیق مفصل و مبسوطی درباره آن پرداخت. فرم کلی این تابع زیان به صورت

$$L(\Delta)=b\, e^{a\Delta} - c\Delta^{-b} \quad a,c \neq 0,\ \Delta=\delta-\theta, b > 0$$

است که Δ نشان دهنده میزان خطای برآورد θ با استفاده از برآوردگر δ می‌باشد. بسادگی می توان دید که $L(0) = 0$ و برای اینکه در $\Delta = 0$ تابع دارای مینیمم باشد باید داشته باشیم $ab = c$. در این صورت فرم استاندارد تابع زیان *LINEX* به صورت

$$L(\Delta) = b\ \{e^{a\Delta}-a\Delta-1\} \quad a \neq 0, b > 0$$

در می‌آید. b پارامتری است که مقدار زیان را تغییر می‌دهد و a پارامتر تعیین شکل است.

Klebanov (1976) توابع‌زیان‌متقارن‌و نامتقارن را از لحاظ نامساوی رائو ـ بکلول مقایسه کرد. *Rojo(1987)*، *J.A.Giles&D.E.A.Giles(1993)*، *Sadooghi alvand(1979)*، *Kuo&Dey(1990)*، *Parsian(1990a,b)*، *Sadooghi.(1990)A*، *Sadooghi parsian (1992)*، *Parsian - N.Sanjari(1993)*، *Parsian - N. Sanjari(1996)* در بخش استنباط آماری تحت این تابع زیان مقالاتی را به رشته تحریر در آورده‌اند. در این رساله سعی شده است تا آنجا که ممکن است تحقیقاتی که در این مورد انجام شده گردآوری و مورد بحث و بررسی قرارگیرد.

فصل اول شامل مقدماتی از استنباط آماری غیرکلاسیک و نظریه تصمیم می‌باشد. در این فصل، ابتدا

مسائل تصمیم معرفی گردیده‌اند سپس قضیه بیز به عنوان محور اصلی بحث در نظریه تصمیم آمده است. قاعده بیز تحت یک تابع زیان کلی تعریف گشته و قضایای مربوطه مطرح شده‌اند. همچنین مجاز بودن قواعد تصمیم و شرایط آن، مجموعه توابع محدب و قواعد تصمیم بیز تعمیم یافته و مینیماکس معرفی گردیده‌اند. در فصل دوم به معرفی تابع زیان *LINEX* پرداخته و سپس قاعده بیز را تحت این تابع زیان بدست آورده‌ایم. در فصل سوم تحت بعضی توزیع‌های خاص از جمله نرمال در حالتی که میانگین معلوم و واریانس مجهول، میانگین مجهول واریانس مجهول است، پواسان و دوجمله‌ای وقتی حجم نمونه مجهول است، مجاز بودن یا نبودن برآوردگر پارامترهای مجهول را بررسی کرده‌ایم.

در فصل چهارم که فصل اصلی این رساله می‌باشد به بحث در مورد پایایی برآوردگر در خانواده نمائی پرداخته‌ایم. و در غالب چند توزیع خاص از این خانواده بهترین برآوردگرهای مکان ـ پایا، مقیاس ـ پایا و برآوردگر نوع *Pitman* را تحت زیان *LINEX* معرفی کرده و در بخش‌های دیگر این فصل به بررسی مجاز بودن این برآوردگرها برای خانواده‌ای از توزیع‌ها با پارامترهای *Scale* ،*Location* ،*Location - scale* نامعلوم، پرداخته‌ایم. در فصل پنجم مجاز بودن برآوردگر بردار میانگین توزیع چند متغیری بررسی گردیده است. همچنین در یک مدل رگرسیون خطی $y = X\beta + u$ برآورد ضرایب و واریانس خطا تحت حالات با محدودیت $(R\beta = r)$ و بدون محدودیت بدست آمده‌اند و مجاز بودن آنها را نیز تحقیق کرده‌ایم. فصل ششم و هفتم به عنوان یک کار *original* (جدید) در این رساله آمده است. فصل ششم به بررسی توزیع *Parto* و مجاز بودن برآورد یکی از پارامترهای آن تحت زیان *LINEX* و مقایسه آن با زیان *MSE* می‌پردازد و در فصل هفتم روشهای دنباله‌ای بیزی مطرح گردیده و قواعد دنباله‌ای بیز که تحت بعضی توزیعهای خاص مینیماکس نیز می‌باشند، تحت زیان *LINEX* بدست آمده‌اند. علاوه بر دو فصل فوق مواردی که با علامتهای (**) و (*) مشخص گردیده‌اند، بترتیب یا در مقالات نبوده‌اند یا به طور خیلی مختصر مورد بحث قرار گرفته بودند، که ما آنها را شرح و بسط داده‌ایم.

مقدمه

در پایان برخود واجب می‌دانم که مراتب سپاسگزاری خود را از:

جناب آقای دکتر ناصر رضاارقامی دانشیار گروه آمار دانشگاه فردوسی مشهد که متحمل زحمت زیادی در امر تدوین این پایان‌نامه شده و راهنمائی رساله را برعهده داشته‌اند، جناب آقای دکتر ابوالقاسم بزرگ‌نیا استاد گروه آمار دانشگاه فردوسی مشهد که داوری رساله را بر عهده داشته‌اند و از راهنمائی‌های خود بنده را بهره‌مند ساخته‌اند همچنین سرکار خانم دکتر ناهید سنجری فارسی‌پور استادیارگروه آمار دانشگاه شیرازکه دعوت ما را قبول کرده و در امر تدوین و ویرایش این رساله بحق ما را مدیون خود قرار داده‌اند ابراز دارم. تذکرات و پیشنهادات این اساتید محترم بسیار ارزنده و مفید بوده و در بهبود بخشیدن این رساله کاملاً موثر بوده است.

همچنین از جناب آقای دکتر شاهکار مدیر محترم گروه آمار و سایر اعضای گروه و سرکار خانم تحققی که دلسوزانه زحمت تایپ این رساله را بر عهده داشته‌اند و از آقای رسول اتحاد مسئول محترم کتابخانه دانشکده علوم ۲ که نقش اساسی را در تهیه بسیاری از مقالات و منابع داشته‌اند، مسئولین بخش تکثیر و خانم حسینی منشی گروه آمار که اینجانب را در جهت گردآوری و تدوین این رساله بنحوی یاری نموده‌اند بی‌نهایت سپاسگزارم.

عباس رسولی

پائیز ۱۳۷۴

فصل اوّل ـ تئوری تصمیم بیزی

مقدمه:

مسائل تصمیم آماری مسائلی هستند که درآنها داده‌ها یا مشاهدات درباره حالت طبیعت‌اند و امید بر این است که مشاهدات حاوی اطلاعاتی باشند که بتوانند برای اخذ تصمیمی بهتربکار روند.

فرض اصلی در حل مسئله تصمیم که فرض بیزی نامیده می‌شود آن است که باورهای انسان در مورد طبیعت می‌تواند به وسیله یک مدل توزیع را نشان داد. این توزیع، توزیع پیشین نامید. به این معنی که باورهای آماردان را به صورتی که قبل از جمع‌آوری و تحلیل داده‌ها وجود دارند، نشان می‌دهد. حال که تا اندازه‌ای یک مسئله تصمیم ساختار ریاضی به خود گرفته و دارای قانون گشته است بعد از بدست آوردن قواعد تصمیم به مقایسه آنها می‌پردازیم. اما چون توابع مخاطره توابعی از θ هستند مقایسه همیشه به سادگی صورت نمی‌گیرد. بدنبال این موضوع با مجموعه قواعد تصمیم مجاز و غیرمجاز آشنا می‌شویم و خواهیم دید که یک قاعده غیر مجاز عملاً بلا استفاده خواهد بود و همچنین بهترین قاعده ممکن را نیز با در نظر گرفتن اصل‌هایی بدست می‌آوریم. در مقابل بیز مبحث دیگری در

استنباط آماری غیرکلاسیک مطرح می‌گردد که به مینیماکس معروف است و با احتیاط بیشتری به مسئله می‌نگرد. در بخش آخر از این فصل آن را مطرح کرده و به مقایسه آن با بیز خواهیم پرداخت.

۱-۱- استفاده از قضیه بیز

در نظرگرفتن وضع (حالت) طبیعی (*state of nature*) به عنوان یک متغیر تصادفی θ، به این معنی است که تابع احتمال یا تابع چگالی $f(\circ;\theta)$ در واقع یک احتمال یا چگالی شرطی است و شرط بر این است که وضع طبیعی θ باشد. فرض می‌کنیم که برای هر ترکیب از وضع θ و عمل a، یک زیان $L(\theta,a)$ وجود دارد که مطلوبیت منفی نتایج حاصل از انتخاب عمل a را وقتی حالت طبیعی در وضع θ است، بدست می‌دهد. چون زیان، به صورت مطلوبیت منفی در نظر گرفته می‌شود کمیتی که باید در تصمیم‌گیری‌های مربوط به زیان تصادفی ملاک مقایسه قرارگیرد مقدار امید این متغیر است.

$$R(\theta,d) = E[L(\theta,d(z))] \qquad (۱-۱-۱)$$

این کمیت به حالت θ و قاعده تصمیم d بستگی دارد و تابع مخاطره نامیده می‌شود. اگر برای Z توزیعی فرض کنیم که با $P_\theta(Z=z) = f(z;\theta)$ تعریف شده است.

تابع مخاطره به صورت

$$R(\theta,d)= \sum_i L(\theta,d(z_i))\, f(z_i,\theta)$$

محاسبه می‌شود.

همان طور که اشاره شد پاورهای انسان در مورد طبیعت می‌تواند به صورت یک مدل احتمال نشان داده شود، که این احتمال نظری یا مدل احتمالی را توزیع پیشین طبیعت می‌نامند. با استفاده از نماد $\pi(\theta)$ برای توزیع پیشین و از روی توزیع شرطی Z می‌توان توزیع توام (Z,θ) را بدست آورد:

$$f(z,\theta)=f(z\,|\,\theta)\pi(\theta)$$

و توزیع مطلق یا حاشیه‌ای متناظر Z، با چگالی احتمال

$$f_z(z) = E_\pi f(z|\oplus) = \begin{cases} \int f(z \mid \theta)\, \pi(\theta)\, d\theta & \Theta \text{ پیوسته} \\ \sum f(z|\theta_i)\, \pi(\theta_i) & \Theta \text{ گسسته} \end{cases}$$

رامحاسبه کرد حال بااستفاده از قضیه بیز توزیع شرطی وضع طبیعی با فرض داده‌های $Z = z$ را بدست می‌آوریم:

$$h(\theta|z) = f(z|\theta)\pi(\theta) / f_z(z)$$

توزیعی از Θ که به صورت فوق تعریف می‌گردد توزیع پیشین نامیده می‌شود و باورهای (جدید) آماردان را نسبت به وضع طبیعی نشان می‌دهد. فکر بکار بردن قضیه بیز برای تعدیل باورهای پیشین به کمک اطلاعات حاصل از نمونه، اساس " آماری بیزی " را تشکیل می‌دهد. لازم به یادآوری است که چگالی پسین (یا تابع احتمال، اگر Θ گسسته باشد) اساساً عبارت است از حاصلضرب تابع درستنمائی مقدار مشاهده شده z در تابع چگالی (یا تابع احتمال) پیشین. مخرج کسر در فرمول مربوط به h یک ثابت تناسب است و گاهی بکار نمی‌رود. همچنین ذکر این نکته ضروری است که توزیع پسین را می‌توان به عنوان وسیله‌ای برای محاسبه قواعد تصمیم بیز بکار برد. روش کار از این قرار است که زیان پسین مورد انتظار را محاسبه می‌کنیم. مثلاً

$$E_h[L(\Theta,a)] = \int l(\theta,a)\, h(\theta|z)\, dz$$

و سپس عملی مانند a را انتخاب می‌کنیم که این کمیت را مینیمم کند.

۱-۱-۱- خانواده‌های مزدوج

در اغلب موارد، خانواده‌ای از توزیعها برای Θ وجود دارد که به ازای حجم نمونه و هر مقدار خاص $X = x$، اگر توزیع پیشین متعلق به این خانواده باشد، توزیع پسین نیز به آن تعلق خواهد داشت. چنین خانواده‌ای، خانواده توزیعهای مزدوج نامیده می‌شود.

در محاسبات توزیع‌های پسین اغلب از مفهوم بسندگی استفاده فراوان می‌شود به این ترتیب که اگر T یک آماره بسنده برای θ باشد و چگالی آن را با $g\ (t|\theta)$ نشان دهیم نتیجه زیر حاصل می‌گردد:

لم ۱-۱-۱- فرض کنید $m\ (t)$ (چگالی حاشیه‌ای t) از صفر بزرگتر و قضیه تجزیه به عوامل نیز برقرار باشد. سپس اگر $T\ (x)=t$ داریم:

$$\pi(\theta|x)=\pi(\theta|t)=\frac{\pi(\theta)\ g\ (t|\theta)}{m(t)}$$

برهان : فرض کنیم θ پیوسته باشد. (اثبات حالت گسسته مشابه است)

با توجه به قضیه تجزیه به عوامل $f\ (x|\theta)\ =\ g\ (t|\theta)\ h\ (x)$ داریم:

$$\pi\ (\ \theta\ |\ x\ )\ =\ \frac{\pi(\theta)\ f\ (x\ |\ \theta)}{\int \pi(\theta)\ f\ (x\ |\ \theta)\ d\theta}\ =\ \frac{\pi(\theta)\ g\ (t\ |\ \theta)\ h\ (x)}{\int\ \pi(\theta)\ g\ (t\ |\ \theta)\ h(x)\ d\theta}$$

$$=\ \frac{\pi(\theta)g(t|\theta)h(x)}{(\int\pi(\theta)g(t|\theta)d\theta)\ h\ (x)}\ =\ \frac{\pi\ (\theta)\ g\ (t\ |\ \theta)}{\int\ \pi(\theta)g(t\ |\ \theta)d\theta}$$

مثال ۱-۱- فرض کنید $X \sim N\ (\ \theta\, \sigma^2\ )$ که θ مجهول و σ^2 معلوم است. فرض کنید $\pi\ (\theta)$ نرمال با میانگین μ و واریانس τ^2 باشد که هر دو معلوم‌اند. حال نشان می‌دهیم که $\pi\ (\theta|x)$ نرمال با میانگین $\mu(x)$ و واریانس ρ^{-1} است که

$$\mu(x)=\frac{1}{\rho}\ (\mu/\tau^2+x/\sigma^2)\ =\ \frac{\tau^2}{\sigma^2+\tau^2}\,x+\frac{\sigma^2}{\sigma^2\ +\ \tau^2}\,\mu \qquad ,\ \rho^{-1}\ =\ \frac{\tau^2\ \sigma^2}{\tau^2\ +\ \tau^2}$$

حل:

$$\pi\ (\theta\ |\ x)=\frac{\pi(\theta)\ f\ (x\ |\ \theta)}{\int\ \pi\ (\theta)\ f(\ x\ |\ \theta)\ d\theta}=\frac{\frac{1}{\sqrt{2n\pi}}\,e^{-\frac{(\theta-\mu)^2}{2\tau^2}}\quad \frac{1}{\sqrt{2\ \pi\ \sigma}}\,e^{-\frac{(x-\theta)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi\ \tau}}\cdot\ \frac{1}{\sqrt{2\pi\ \sigma}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}[\frac{(\theta-\mu)^2}{\tau^2}+\frac{(x-\theta)^2}{\sigma^2}]}d\theta}$$