پایان نامه

برای دریافت درجه کارشناسی ارشد

آمار

موضوع :

مدلهای پویای بیزی برای جداول گسسته

استاد راهنما :

آقای دکتر سیامک نوربلچی

اساتید مشاور :

آقای دکتر علی عمیدی
آقای دکتر جلال داودزاده
آقای دکتر محمدرضا مشکانی

نگارش :

فرزاد اسکندری

شهریور ماه ۱۳۷۲

فهرست مطالب

عنوان | صفحه

۱۷۴۸۹

عنوان	صفحه

عنوان | صفحه

تقــــدیـــم بـــه :

پـدرم، بـهخـا طـرمرا رتـها یشـا ن

مـا درم ،بـهخا طــــررنجها یشـــا ن

خوا هرا نم ، بـهخا طـرمهربا نیها یشا ن

برا درا نم ، بـهخا طرتشویقها یشـا ن

شا یدکهبتوا نم یک ا زهزا ررنجی را کهبهخا طرمــن تحمل

کردها ندجبرا ن کردهبا شـــم .

موضوع: مدلهای پویای بیزی برای جداول گسسته

نگارش : فرزاد اسکندری

دوره آموزش : کارشناسی ارشد آمار دانشگاه شهید بهشتی

تاریخ دفاع : شهریور ۱۳۷۲

استاد راهنما : دکتر سیامک نوربلچی

اساتید مشاور:

۱- دکتر علی عمیدی

۲- دکتر جلال داودزاده

۳- دکتر محمدرضا مشکانی

خلاصه :

آمار به عنوان یک روش علمی امروزه در سطح بسیار وسیعی توسط دانشمندان علوم مختلف مورد استفاده قرار می گیرد ، با به کارگیری روشهای مختلف آماری پژوهشگران سعی می کنند استنتاج های استقرائی نظریه شان را در مورد نتایج علمی به صورت دقیق بیان کنند . همین امر باعث می شود تا آمارشناسان در تلاش باشند تئوریهای آماری خود را با دقتی هرچه کاملتر عرضه نمایند . با توجه به متفاوت بودن علوم مختلف طبیعتاً " نوع به کارگیری

روشهای مختلف آماری نیزبا یکدیگرتفاوت می کنند .

یکی ازمسائل مهمی که معمولا" برای کارشناسان دارای اهمیت است مسئله پیش بینی وضع آینده سیستم یا فرایندهائی است که با زمان درحال تغییرند . درآماریکی ازروشهائی که می تواند به پرسشهائی درباره این فرایندها پاسخ دهد سریهای زمانی است ، که عمدتا " توسط باکس وجنکینز (۱۹۷۲) به صورت کلاسیک صورت بندی شده است .

زلنر (۱۹۷۵) با به کاربستن قضیه معروف بیز درسریهای زمانی سعی کرده است مسئله پیش بینی را به طریقه بیزی نیزحل کند . اما درهردو دیدگاه فرض اساسی ، عدم تغییرپارامترهای سیستم نسبت به زمان می باشد .

با حذف این فرض اخیرکالمن (۱۹۶۰) ایده سیستم های پویا وآنچه معروف به صافی خطی کالمن است را معرفی کرد .مین هولدوسینگپوروالا (۱۹۸۳) با استفاده ازایده صافی کالمن که درآن فرض شده است پارامترهای سیستم نسبت به زمان درحال تغییرند ، مسئله اخیررا به صورت بیزی ، برای حالتی که جامعه های آماری گسترده درزمان نرمالند مطالعه کردند .درسال ـ های اخیرهریسون ، وست ، پل ومیگن (۱۹۸۵) تحلیل بیزی مساله صافی کالمن را به صورت وسیعی موردمطالعه قراردادندونظریه متمایزمدل های

پویای بیزی را عرضه‌نموده‌اند.

درفصل اول این رساله‌سعی کرده‌ایم مدل های پویا را دردوحالت خطی وغیرخطی معرفی وخواص عمومی آنها را بیان کنیم.

درفصل دوم چگونگی به‌کارگیری صافی کالمن بیزی درمدلهای پویا توسط مین هولدوسینگپوروالا وهمچنین به‌هنگام سازی پارامترهای یک سیستم پویا با استفاده‌ازشبکه‌بیزی که‌توسط نورمندوتریچلر (۱۹۹۰) بیان شده‌است را به‌تفضیل شرح داده‌ایم.

ابتدای فصل سوم این رساله‌به‌تعمیم مفهوم به‌هنگام کردن پارامترهای یک سیستم با استفاده‌ازصافی کالمن که‌توسط هریسون، وست، پل ومیگن (۱۹۸۵) ارائه‌شده‌است، اختصاص داده‌ایم وسپس برای خانواده‌ٔ پواسون روش ایشان رابکاربرده‌ایم ونحوه‌ٔ استنباط و به‌هنگام سازی سری زمانی که‌داده‌های آن دارای توزیع پواسون هستندرا بررسی کرده‌ایم.

فصل چهارم به‌مساله‌به‌هنگام کردن پارامترهای یک سیستم با استفاده‌ازروشهای خطی پویا جداول چندبعدی گسسته‌می پردازدودرپایان نیزبرای کاربردعملی نظریه‌گفته‌شده‌دراین فصل یک برنامه‌ٔ نرم‌افزاری، همراه‌باخروجی های کامپیوتری مربوط به‌داده‌های یک جدول (۲× ۲× ۲) ارائه‌شده‌است.

فصـــل اول

مقدمـــه

یکی از مسائل مهمی که همواره ذهن پژوهشگران را به خود مشغول کرده است، پاسخ دادن به این سوءال مهم می باشد که چگونه می توان وضع سیستمی کـــه وابسته به یک سری عوامل تصادفی وغیرتصادفی که نسبت به زمان درحال تغییـــر می باشند را تحت کنترل دربیاورند و برای زمانهای آینده سیستم وضع آن را پیش بینی بکنند؟ این ایده برای اولین بار توسط " باکس" وجنکینز "[۱] (۱۹۷۲) به صورت کلاسیک فرمول بندی گردید. ولی تئوری آنها برای حالتی بیان شده بود که فرض کرده بودند، پارامترهای نامشخصی که درسیستم وجوددارند، نسبت بـــه زمان ثابتند. اما از آنجائی که این موضوع یعنی بررسی تغییرات پارامترهای یک سیستم نسبت به زمان نیزدارای اهمیت است بسیاری از آمارشناسان سعـی کردند این مسئله را نیزمدل بندی کنند. درسال (۱۹۸۳) مین هولدوسینگپور- والا[۲] با استفاده ازایدهٔ کالمن[۳] (۱۹۶۰) طی انتشارمقاله‌ای بـرای اولین بار به این مسئله پاسخی مناسب براساس نظریه بیزی آماردادند اساس تئوری آنها همان قضیه معروف بیز می باشد ولی از "مدلهای پویا" نیـــز استفاده کردند. این که ایشان چگونه پاسخ این سئوال یعنی نحوهٔ تغییـــرات پارامترسیستم را داده اند درفصل دوم به طورمفصل موردبررسی قرارخواهـــد

۱- BOX.G.E.P and JENKINS
۲- MEINHOLD. R. and D.SINGPURWALLA. NOZER
۳- KALMAN, R.E.

گرفت .درفصل حاضرسعی کرده‌ایم مدلهای پویا را به‌طوررسمی تعریف کنیـــم و مشخص کنیم این گونه مدلها درچه مسائلی کاربرددارند . ابتدا دربخش بعـــدی مقدمه‌ای راجع به مدلهای پـویا بیان می کنیم .سپس مشخص می کنیم که درچـــه زمینه‌های آماری مدلهای پویا کاربرددارند .

۱ - ۱ - مدلهای پویا

فرض کنیدکه درمکانیزمی خاص متغیرپاسخی مانند Y وجوددارد که به نوعی به یک متغیرورودی به این سیستم مانند X مربوط می شودواین ارتباطی که بین Y و X وجوداردبه صورت :

$$Y = X\theta + \varepsilon \qquad (۱-۱-۱)$$

بیان شده است .دررابطهٔ (۱-۱-۱) θ پارامترنا مشخصی و ε خطای موجوددر سیستم است .همچنین برای این سیستم فرض کنیدشخصی که چنین ارتباطی را بررسی می کندبراین باوربا شدکه پارامترنا مشخص θ خود، دارای تغییراتی تصادفی است که از یک قانون آماری تبعیت می کند .ما این قانون را $P(\theta)$ می نامیم . پس درواقع می توان چنین گفت که مجموعه ای ازمدلهای از نوع (۱-۱-۱) داریم که نسبت به مقادیرمختلف θ بدست آمده اند.

اصولا" طبیعت پویای فرآیندها وسیستم های مختلف طلب می کندکه به طور دقیق نتوانیم درزمان خاص مدل اصلی را تشخیص دهیم واین مطلب را به دنبال خواهدداشت که احساس کنیم ساختار θ به کندی نسبت به زمان درحال تغییر است .درواقع می توان گفت مدل مزبورنسبت به زمان درحال تغییرمی باشد، ودرنتیجه ما به جای اینکه یک مدل داشته باشیم ، دنباله ای ازمدلهای منظم راخواهیم داشت .همچنین باید به این نکته اشاره کردکه هرکدام ازاعضاء این

دنبا له با توزیع احتمالی همراه است. به طور کلی به این مجموعه مدلهای بدست آمده که نسبت به زمان درحال تغییرند "مدلهای پویا" می گویند. برای سادگی، این مجموعه را با $\mathcal{M}$ نمایش می دهیم وتوزیع احتمال بوجود آمده توسط هرکدام از اعضاء این مجموعه را با $p(Y/\mathcal{M})$ نشان می دهیم: البته باید اشاره کرد که Y خود بردار مشاهدات می باشد و توزیع پیش بینی آن به صورت زیر است:

$$p(Y) = \int_{M\in\mathcal{M}} p(Y/M)\, dp(M) \qquad (۱-۱-۲)$$

ذکر این نکته حائز اهمیت است که بخش بزرگی از مدلهای پویائی که در طبیعت موردبررسی قرار می گیرند خطی هستند که از آن جمله می توان مدلهای چندجمله ای مرتبه اول، مدلهای رگرسیونی ومدلهای سریهای زمانی را نام برد و در بخشهای بعدی به طور مفصل درباره آنها سخن خواهیم گفت. همانطورکه درمقدمه بیان کردیم هدف اصلی ای که دراینگونه مدلها دنبال می کنیم پیش بینی وضعیت مدل برای زمانهای آینده است ولی اصولا" در یک مسئله پیش بینی احتیاج به شرایط اولیه ای داریم که ما این شرایط اولیه را با $\mathcal{D}_0$ نمایش می دهیم. درواقع $\mathcal{D}_0$ شامل کلیه اطلاعاتی است که می خواهیم آنها را آغاز گر بررسی تغییرات یک سیستم قرار دهیم وبراساس $\mathcal{D}_0$ توزیع پیش- بینی را مشخص کنیم. که ما این توزیع پیش بینی را براساس اطلاعات $\mathcal{D}_0$ با

$p(Y_1/D_0)$ نمایش می دهیم .

فرض کنیدکه درزمان دلخواه $(t-2)$ می باشیم وکلیه اطلاعات موجود در زمان $(t-2)$ را با D_{t-2} نمایش داده ایم .طبیعتا " توزیع پیش بینـــی در زمان $(t-1)$ به صورت $p(Y_{t-1}/D_{t-2})$ می باشد.حال اگردرزمـــان t ، مشاهده جدیدی به مجموعه مشاهدات دردست افزوده شود ، درواقع اطلاعات جدیدی

بدست آمده است واین اطلاعات به مقداراطلاعات D_{t-2} افزوده

میشود.چنانچه اطلاعات موجوددر Y_t را با I_t نمایش دهیم آنگاه کل اطـــلاعات را می توان به صورت $D_{t-1}=\{I_{t-1}, D_{t-2}\}$ نشان داد.

با توجه به این اطلاعات جدیدبرای توزیع پیش بینی خواهیم داشت :

$$p(Y_t, \theta_t / D_{t-1}) = p(\theta_t / D_{t-1}) \cdot p(Y_t / D_{t-1}, \theta_t)$$

$$p(Y_t / D_{t-1}) = \int_{\theta} p(\theta_t / D_{t-1}) \cdot p(Y_t / D_{t-1}, \theta_t)\, d\theta_t$$

۱ ـ ۲ ـ مدلهای چندجمله‌ای مرتبهٔ اول :

بعد از تعریف مدلهای پویا و به‌طور خاص مدلهای خطی پویا، در این قسمت می‌خواهیم یکی از حالتهای خاص مدلهای خطی پویا یعنی مدلهای چندجمله‌ای مرتبه اول را معرفی کنیم .

یکی از ساده‌ترین حالتهای مدلهای خطی پویا که در طبیعت نیز به وفور دیده می‌شود، مدلهای چندجمله‌ای مرتبه اول می‌باشد. صورت کلی این گونه مدلها به‌شکل (۱ـ۲ـ۱) $Y_t = \mu_t + \nu_t$ می‌باشد که به آن مدل مشاهدات میگویند.

در این مدل μ_t پارامتر موجود در سیستم است که نسبت به زمان در حال تغییر می‌باشد و ν_t خطای موجود در مدل مشاهدات می‌باشد و از یک قانون آماری تبعیت می‌کند. فرض می‌کنیم ν_t دارای توزیع نرمال با میانگین صفر و واریانس V_t است . هرگاه مشاهدهٔ جدیدی در زمان دلخواه t به سیستم افزوده شود، اطلاعات موجود در آن به کل اطلاعات موجود در سیستم یعنی D_{t-1} افزوده می‌شود. همین امر باعث می‌شود که μ_t یعنی پارامتر موجود در سیستم در زمان t ، دستخوش این تغییرات شده و به نوعی با پارامترهای موجود در زمان قبل در ارتباط باشند. فرض می‌کنیم این ارتباط به صورت (۱ـ۲ـ۲) $\mu_t = \mu_{t-1} + \omega_t$ است. در رابطهٔ (۱ـ۲ـ۲)، μ_{t-1} پارامتر سیستم در زمان $(t-1)$ و ω_t خطای موجود در سیستم است که فرض می‌کنیم از

توزیع نرمال با میانگین ۰ وواریانس W_t تبعیت می‌کند.

درمدلهای (۱-۲-۱)و(۱-۲-۲) دونوع خطا وجوددارد که فرض می کنیم برای تمام زمانهای t و s که $t>s$ می باشد ν_t و ν_s و همچنین ω_t و ω_s ازیکدیگر مستقلند. برای زمان دلخواه t نیز ν_t و ω_t ازیکدیگرمستقل می باشند. ماترسهای واریانس W_t و V_t دراین مدل ابتدا فرض می کنیم معلوم هستند. از آنجائی که فرض کرده‌ایم دونوع خطای موجود، ازقانون نرمال تبعیت می کنندپس خواهیم داشت:

$$
(۱-۲-۳) \qquad t=1,2,\ldots \qquad \begin{cases} (Y_t/\mu_t) \sim N(\mu_t, V_t) \\ (\mu_t/\mu_{t-1}) \sim N(\mu_{t-1}, W_t) \end{cases}
$$

براساس روابط (۱-۲-۳) چنانچه بخواهیم برای K امین مشاهده بعد از مشاهده t عمل پیش بینی راانجام بدهیم، خواهیم داشت:

$$
(۱-۲-۴) \qquad E[Y_{t+K}/\mu_t] = E[\mu_{t+K} + \nu_{t+K}/\mu_t] = E[\mu_{t+K}/\mu_t] = \mu_t
$$

وتابع پیش بینی کننده که براساس اطلاعات موجودرزمان t، یعنی D_t بدست می آیدبه‌صورت زیراست:

$$
(۱-۲-۵) \qquad f_t(K) = E[Y_{t+K}/D_t] = E[\mu_{t+K}/D_t] = m_t \qquad \text{برای تمام } K>0
$$

البته باید توجه داشت ازآنجائیکه برای t های مختلف خطاهای موجوددر مدل مشاهدات ازیکدیگرمستقل می باشندپس می توان گفت: $E(\nu_{t+K}/D_t) = E(\nu_t)$

بعدازمعرفی مدلهای چند جمله‌ای مرتبه اول، درزیراین