بررسی ناایستایی واریانس سری زمانی

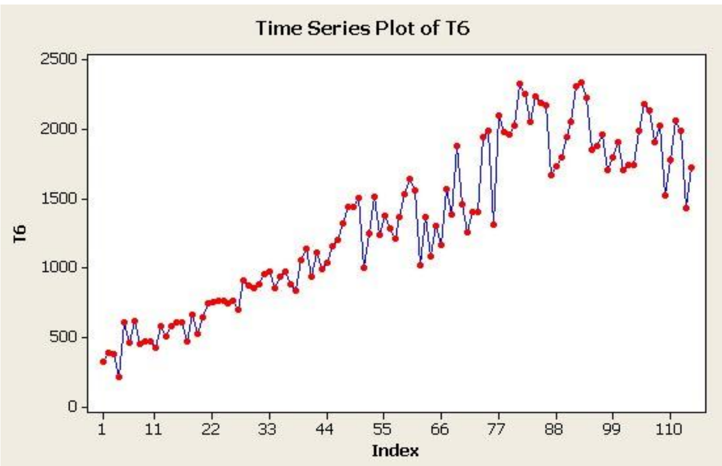
مقدمه

اگر به مرور زمان تغییر پذیری یک سری زمانی افزایش یابد بدین معنی است که سری مزبور نسبت به واریانسش ناایستا است. با توجه به آنکه بسیاری از مدل های احتمال سری های زمانی بر مبنای ایستایی سری استوار می باشند، لذا در صورت ناایستایی سری در میانگین یا در واریانس می توان با استفاده از تکنیکهایی نسبت به ایستا سازی سری زمانی اقدام کرد. ناایستایی در میانگین در مقالات قبلی مورد بررسی قرار گرفت. در این مقاله موضوع ناایستایی سری در واریانس را مورد بررسی قرار خواهیم داد.

۱- ناایستایی در واریانس

چنانچه یک سری زمانی در میانگین ایستا باشد ولی در واریانس ناایستا باشد باید بااستفاده از تبدیل مناسب اقدام به ایستا سازی واریانس نمود. ممکن است یک سری زمانی هم در میانگین و هم در واریانس ناایستا باشد، در این صورت ابتدا باید واریانس آن را ایستا نمود.

نمودار زیر یک سری زمانی را نشان می دهد که هم در میانگین و هم در واریانس ناایستا می باشد. برای ایستاسازی این سری همانطورکه قبلا گفته شد، ابتدا باید واریانس آنرا ایستا نمود. یعنی با استفاده از یک تبدیل مناسب داده ها را طوری تغییرداد که دارای واریانس ثابت باشند. بعد از تثبیت واریانس، با استفاده از تفاضلی کردن می توان داده ها را نسبت به میانگین نیز ایستا نمود.



نمودار یک سری زمانی ناایستا در میانگین و واریانس

تغییر واریانس یک فرآیند ناایستا، وقتی سطح آن تغییر می کند، بسیار متداول است. بنابراین برای ثابت c و تابع f داریم :

$$\operatorname{var}(x_t) = cf(\mu_t)$$

یعنی واریانس فرآیند تابعی از میانگین آن فرآیند می باشد. جهت ایستاسازی واریانس باید یک تابع T را بگونه ای پیدا کنیم که سری تبدیل شده $T(x_t)$ یک واریانس ثابت داشته باشد. برای تشریح روش، تابع مورد نظر را با یک سری تیلور مرتبه اول در حول نقطه μ_t تقریب می زنیم. فرض کنید:

$$T(x_t) \approx T(\mu_t) + T'(\mu_t)(x_t - \mu_t)$$

که $T'(\mu_t)$ مشتق اول $T(x_t)$ در μ_t است. اکنون داریم:

$$\operatorname{var}(T(x_t)) = [T'(\mu_t)]^2 \operatorname{var}(x_t) = c[T'(\mu_t)]^2 f(\mu_t)$$

بنابراین برای اینکه واریانس $T(x_t)$ ثابت شود، باید تبدیل پایداری واریانس $T(x_t)$ را چنان انتخاب کنیم که:

$$T'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}}$$

از معادله فوق نتیجه می شود:

$$T(\mu_t) = \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t$$

به عنوان مثال اگر انحراف معیار یک سری متناسب با سطح آن باشد بطوریکه $\operatorname{var}(x_t) = c^2 \mu_t^2$ آنگاه داریم :

$$T(\mu_t) = \int \frac{1}{\sqrt{\mu_t^2}} d\mu_t = \ln(\mu_t)$$

بنابراین یک تبدیل لگاریتمی(مبنای آن مهم نیست) بشکل $\ln(x_t)$ باعث ثابت شدن واریانس خواهد شد.

به عنوان مثال دیگرفرضکنید واریانس سری متناسب باسطح سری باشد، یعنی $\operatorname{var}(x_t) = c\mu_t$، آنگاه داریم:

$$T(\mu_t) = \int \frac{1}{\sqrt{\mu_t}} d\mu_t = 2\sqrt{\mu_t}$$

بنابراین یک تبدبل ریشه دوم سری بشکل $\sqrt{x_t}$ باعث ثابت شدن واریانس می شود.

۲- تبدیلات باکس-کاکس

بطورکلی برای تبدیل واریانس، از تبدیل توانی زیرکه بوسیله باکس و کاکس (۱۹۶۴) معرفی شده است، استفاده می کنیم :

$$T(x_t) = x_t^{(\lambda)} = \frac{x_t^{\lambda} - 1}{\lambda}$$

λرا پارامتر تبدیل می نامند. تبدیلات مربوط به چند مقدار λ که معمولا مورد استفاده قرار می گیرد، بصورت زیر می باشد. لازم به ذکر است که چنانچه مقدار λ برابر یک شود، نیازی به تبدیل نیست.

تبدیل مناسب	مقدار λ
$1/x_t$	-۱
$1/\sqrt{x_t}$	-۰.۵
$\ln(x_t)$	۰
$\sqrt{x_t}$	۰.۵
نیازی به تبدیل نیست.	۱

جدول تبدیلات توانی باکس- کاکس

چند نکته

۱- تبدیلات پایداری واریانس، فقط برای سریهای مثبت بکار می رود. با وجود این آنطور که به نظر می رسد، محدودیتی وجود ندارد. زیرا همیشه می توان مقدار ثابتی را به سری افزود، بدون اینکه ساختار همبستگی سری تغییر کند.

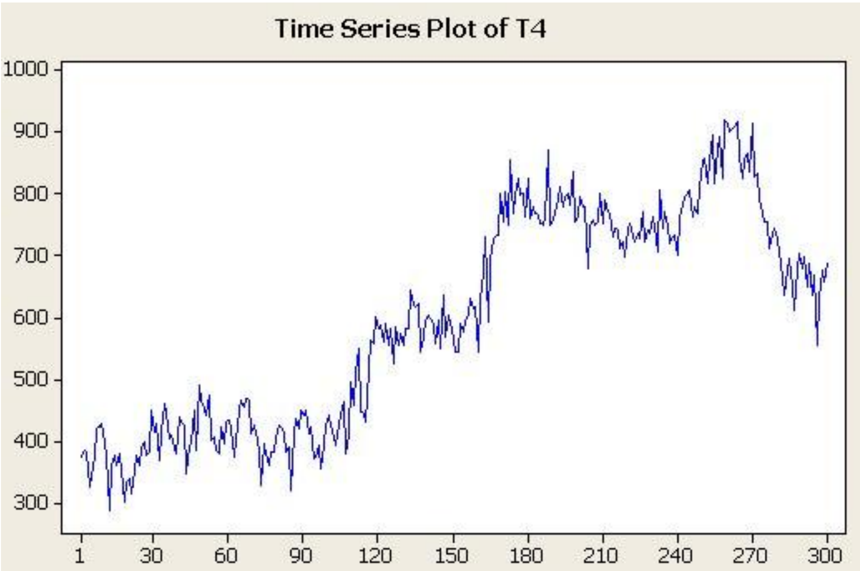
۲- اگر یک تبدیل پایداری واریانس لازم باشد، باید قبل از هر گونه تحلیلیمانند تفاضلی کردن، به اجرا درآید.

۳- غالبا تبدیل فقط برای پایداری واریانس نیست بلکه تقریب برای نرمال بودن را نیز بهتر می کند.

۴- در تبدیل توانی می توان λ را به عنوان یک پارامتر که از سری مشاهده شده برآورد می گردد، در نظر گرفت. برآورد درست نمائی ماکزیمم λ آن است که مجموع مربعات باقیمانده را می نیمم می کند. برای هر مقدار λ مجموع مربعات از الگوی برازش شده، محاسبه می شود. برآورد درستنمائی ماکزیمم λ آناست که در بین مقادیر دیگر کوچکترین مجموع مربعات باقیمانده را می دهد.

مثال

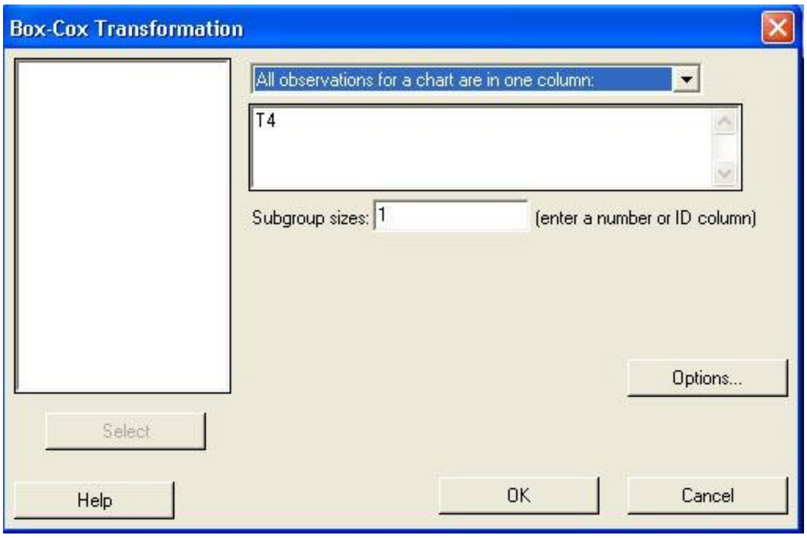
برای روشن شدن مطلب، سری زمانی زیر را در نظر می گیریم و لزوم استفاده از تبدیلات توانی باکس-کاکس را برای تثبیت واریانس بررسی می کنیم. همانطور که ملاحظه می شود این سری در میانگین ناایستا است.یعنی سطح سری با گذشت زمان تغییر می کند. اما با توجه به نمودار سری می توان گفت که واریانس این سری تقریبا ثابت است و نیازی به تبدیل ندارد، یعنی تغییر پذیری سری با گذشت زمان تقریبا ثابت باقی می ماند.



نمودار سری زمانی T۴

برای بررسی بیشتر این مطلب از تبدیل باکس-کاکس استفاده می کنیم. چنانچه بهترین مقدار پیشنهادی برای λ (پارامتر تبدیل) عدد یک باشد، با توجه به جدول تبدیلات توانی، می توان گفت واریانس داده ها ثابت است و نیازی به تبدیل داده ها نداریم.

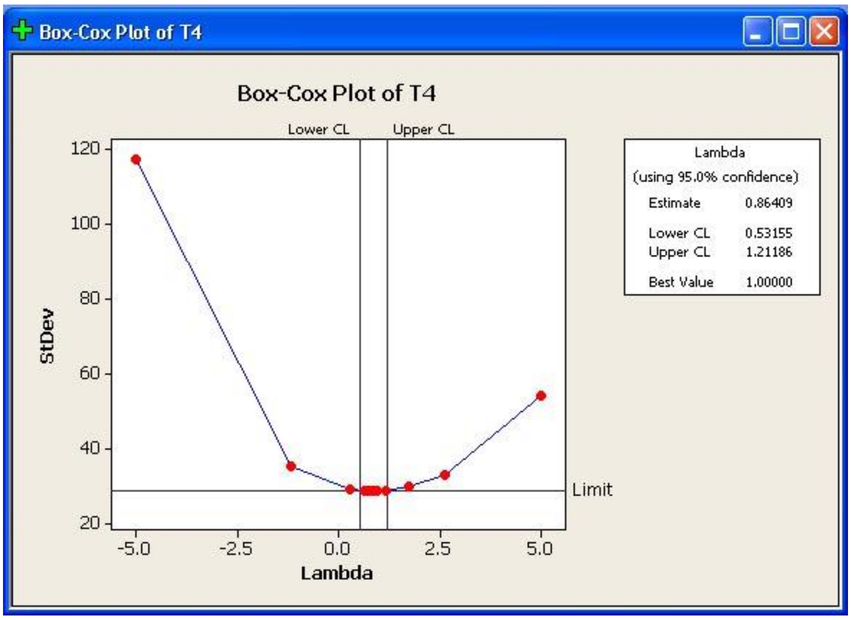
حال تبدیل باکس-کاکس را درمینی تب اجرا می کنیم. برای تثبیت واریانس با استفاده از تبدیلات باکس-کاکس در مینی تب، از منوی Stat گزینه Control Charts را انتخاب نموده و سپس از منوی ظاهر شده گزینه Box-Cox Transformation را انتخاب می کنیم تا پنجره زیر باز شود.



پنجره تبدیلات توانی باکس-کاکس

در پنجره ظاهر شده در کادر بالای صفحه گزینه پیش فرض مینی تب را که می گوید، همه مشاهدات در یک ستون قرار دارند، می پذیریم. سپس با دو بار کلیک کردن برروی سری مورد نظر، آنرا به کادر وسط صفحه منتقل می کنیم.

در قسمت Subgroup sizes نیز عدد یک را وارد می کنیم. (در مبحث کنترل کیفیت، حجم زیر گروه ها معمولا ۴ یا ۵ می باشد اما در سری های زمانی برای استفاده از تبدیلات باکس-کاکس، حجم زیر گروه را همیشه ۱ در نظر می گیریم). با فشردن دکمه *ok* نتیجه به شکل زیر خواهد بود:



نمودار باکس-کاکس برای سری T۴

همانطور که ملاحظه می شود درکادر سمت راست برآورد λ وحدود اطمینان ۰.۹۵ وبهترین مقدارپیشنهادی برای λ، داده شده است. با توجه به اینکه بهترین مقدار پیشنهادی برای λ عدد یک می باشد و با توجه به جدول تبدیلات توانی باکس- کاکس به این نتیجه می رسیم که هیچ تبدیلی برای پایایی واریانس این سری لازم نیست.

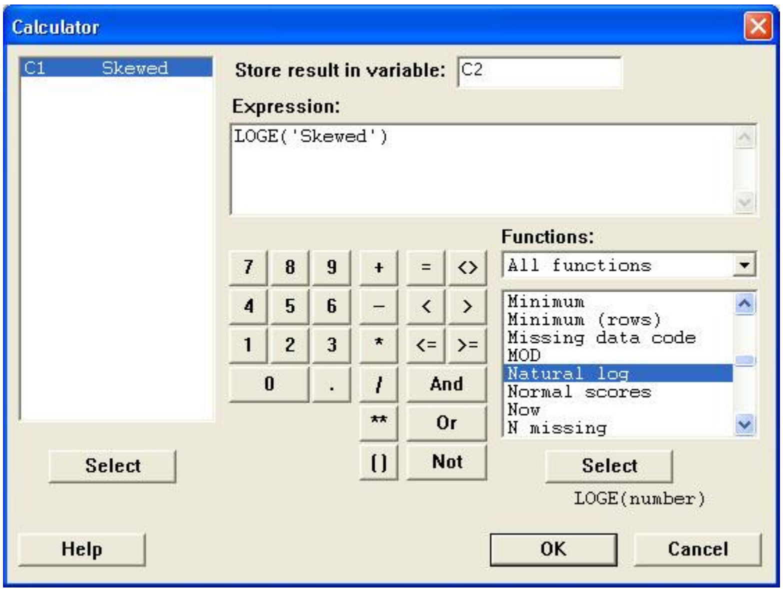
توجه

حدود اطمینان ۰.۹۵ شامل همه مقادیر λ می باشد که انحراف استانداردشان کمتر یا مساوی با مقداری است که خط افقی نشان می دهد. بنابراین هرمقداری از λ که انحراف استانداردی نزدیک به خط افقی داشته باشد، نیز یک مقدار قابل قبول و منطقی برای استفاده جهت تبدیل داده ها خواهد بود.

۳- تبدیل لگاریتمی در MINITAB

ممکن است رویه باکس- کاکس، تبدیل لگاریتمی داده ها را پیشنهاد نماید. برای انجام این کار از منوی Calc گزینه Calculator را انتخاب می کنیم. سپس در قسمتFunctions تابع مورد نظر را که دراینجا تابع لگاریتم طبیعی(Natural log) می باشد برمی گزینیم. با دوبارکلیک کردن برروی تابع موردنظر،این تابع به کادر Expression منتقل می شود. حال از کادر سمت چپ، بر روی سری زمانی مورد نظر (در شکل زیر Skewed) دو بار کلیک می کنیم تا تبدیل مورد نظر برروی این سری انجام شود. همچنین در کادر Store result in variable

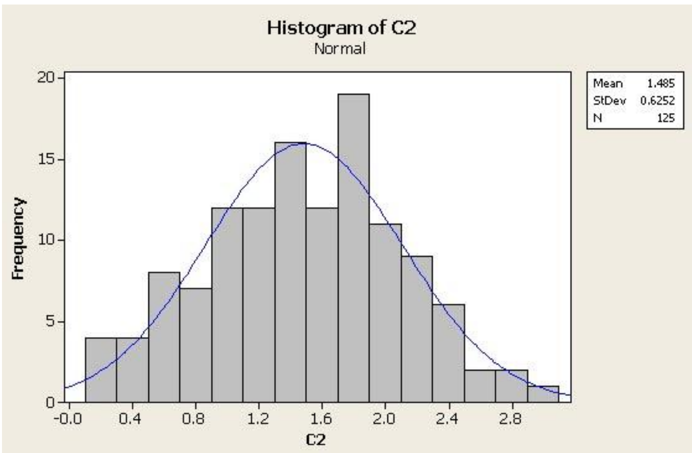
می توان یکی از ستونهای مینی تب را برای ذخیره کردن مقادیر تبدیل شده در نظر گرفت. ما در اینجا مقادیر تبدیل شده را در ستون C۲ ذخیره می کنیم.



تبدیل لگاریتمی متغیر Skewed

بررسی اثر تبدیل لگاریتمی داده ها

چنانچه هیستوگرام داده های تبدیل شده را رسم نماییم، می توانیم اثر تبدیل انجام شده بر روی توزیع داده ها را ببینیم.



هیستوگرام داده ها پس از تبدیل

همانطور که ملاحظه می شود، تبدیل باکس-کاکس تأثیر زیادی بر نزدیک کردن توزیع داده ها به توزیع نرمال داشته است. تبدیل علاوه بر پایدار کردن واریانس، تقریب برای نرمال بودن را نیز بهتر می کند.

پایان.

توضیحات:

مطالب فوق بخشی از کتاب " تجزیه و تحلیل سریهای زمانی با نرم افزار مینی تب" اثر مصطفی خرمی و دکتر ابوالقاسم بزرگنیا می باشد .علاقه مندان به یادگیری تکنیکها و روشهای تحلیلی و پیش بینی سریهای زمانی و آموزش عملی با نرم افزار مینی تب می توانند نسخه الکترونیک این کتاب را به راحتی از فروشگاه اینترنتی شرکت داده پردازی آماری اطمینان شرق به نشانی:

http://spss-iran.ir/eshop.php، دریافت نمایند.

این کتاب دارای ۳۵۰ صفحه می باشد و مبحث سریهای زمانی را با جزئیات کامل در قالب حل مثالهای واقعی و متنوع در نرم افزار مینی تب توضیح می دهد. برای آشنایی بیشتر با این کتاب، فصول و فهرست مطالب و صفحات اول آنرا می توانید بصورت رایگان از لینک زیر دانلود نمائید.(کافیست در کیبرد سیستم خود کلید ctrl را فشار داده و روی لینک زیر کلیک نمائید و پیغام نمایش داده شده را تائید کنید.)

دانلود فهرست مطالب و نام فصول کتاب : تجزیه و تحلیل سریهای زمانی با نرم افزار مینی تب