



دانشگاه گیلان

گروه آمار

گرایش آمار ریاضی

عنوان :

مدل جدیدی در گارچ نامتقارن، گارچ گویا و آزمونی تجربی برای بازده سهام

از :

امین منصوری کوهستانی

استاد راهنما :

دکتر حسین صمیمی حق گزار

اسفند ۱۳۹۷



باسمه تعالی

صور تجلسه دفاع از پایان نامه کارشناسی ارشد

جلسه دفاع از پایان نامه کارشناسی ارشد آقای امین منصوری کوهستانی در رشته: آمار گرایش: آمار ریاضی با عنوان: مدل جدیدی در گارچ نامتقارن، گارچ گویا و آزمونی تجربی برای بازده سهام به ارزش ۶ واحد، در ساعت ۱۷:۳۰ روز یکشنبه مورخ ۹۷/۱۲/۱۹ در محل دانشکده علوم ریاضی دانشگاه گیلان تشکیل شد. هیأت داوران پس از بررسی، نتیجه را به شرح زیر اعلام می‌دارند:

□ پایان نامه با نمره..... و با امتیاز عالی □، بسیار خوب □، خوب □، قابل قبول □ مورد تأیید قرار گرفت.

□ پایان نامه در وضع فعلی با اصلاحات جزئی و نمره ۱۷.۵ و امتیاز عالی □، بسیار خوب □، خوب □، قابل قبول □ پذیرفته شد.

□ پایان نامه به شکل فعلی، مورد تأیید قرار نگرفت و پیشنهاد شد

امضا	محل کار عضو هیات داوران	تخصص	مرتبه دانشگاهی	اعضای هیات داوران
	دانشگاه گیلان دانشگاه گیلان	نظریه احتمال	استادیار	استاد راهنما ۱- دکتر حسین صمیمی حق گزار
	دانشگاه گیلان			استاد(ان) مشاور -۱
	دانشگاه گیلان دانشگاه گیلان	آمار ریاضی مالی	استاد دانشیار	داوران ۱- دکتر بهروز فتحی ۲- دکتر فرشید مهر دوست

امضا	گروه آموزشی/پژوهشی	مرتبه دانشگاهی	نماینده تحصیلات تکمیلی ^۱
	گروه علوم کامپیوتر	استادیار	دکتر مریم سیفا الدینی

نمونه ۳



دانشگاه گیلان

حوزه معاونت آموزشی و تحصیلات تکمیلی دانشگاه

تعهذنامه اصالت پایان نامه

اینجانب.....امین منصوری کوهستانی.....دانش آموخته مقطع کارشناسی ارشد ناپیوسته در رشتهآمار.....گرایشآمار ریاضی.....که در تاریخ۱۳۹۷/۱۲/۱۹..... از پایان نامه خود با عنوانمدل جدیدی در گارچ نامتقارن، گارچ گویا و آزمون تجربی برای بازده سهام..... با کسب نمره۱۷/۵۰..... و درجهبسیار خوب.....دفاع کرده ام، اظهار می کنم که:

۱) این پایان نامه حاصل تحقیق و پژوهش اینجانب بوده و در مواردی که از دستاوردهای علمی و پژوهشی دیگران (اعم از پایان نامه، کتاب، مقاله و ...) استفاده کرده ام، مطابق ضوابط موجود، نام منبع مورد استفاده و سایر مشخصات آن را در فهرست منابع ذکر و درج نموده ام.

۲) این پایان نامه قبلاً برای دریافت هیچ مدرک تحصیلی (هم سطح، پایین تر یا بالاتر) در سایر دانشگاهها و مؤسسات آموزش عالی داخلی و خارجی ارائه نشده است.

ضمناً متعهد می شوم:

۳) چنانچه بعد از دانش آموختگی، قصد استفاده و هر گونه بهره برداری اعم از چاپ مقاله، کتاب، ثبت اختراع و ... از این پایان نامه یا رساله را داشته باشم، از استاد محترم راهنما و گروه آموزشی مربوطه مجوزهای لازم را اخذ نمایم.

۴) چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود، عواقب ناشی از آن را بپذیرم و دانشگاه گیلان مجاز است با اینجانب مطابق ضوابط و مقررات رفتار نموده و در صورت ابطال مدرک تحصیلی ام، هیچگونه ادعایی نخواهم داشت.

نام و نام خانوادگی: امین منصوری کوهستانی

تاریخ و امضاء: ۱۳۹۸، ۶، ۱۹

تقدیم به

به پاس تعبیر عظیم و انسانی شان از کلمه‌ی ایثار و از خودگذشتگی؛
به پاس عاطفه‌ی سرشار و گرمای امید، بخش وجودشان که در این سردترین روزگار ان بهترین پشتیبان است؛
به پاس قلب‌های بزرگشان که فریادس است و سرگردانی و ترس در پناهشان به شجاعت می‌کراید؛
و به پاس محبت‌های بی‌دینشان که هرگز فروکش نمی‌کند؛
این مجموعه را به پدر و مادر عزیزم تقدیم می‌کنم.

تقدیر و شکر

سپاس ایزدمنان که به من این فرصت را داد تا به این مرحله از علم رسیده و از هیچ محبتی دریغ نکرد و در تمام مراحل زندگی تا کنون مرا قوت قلب بوده.

بر حسب وظیفه و مصداق بیان گهربار مولای متقیان که فرمودند:

((هر کس کلمه ای به من بیاموزد مرا بنده خویش ساخته است.)) خود را مدیون زحمات استاد با کمالات و شایسته ام، جناب آقای دکتر حسین صمیمی حق گزار می دانم. ایشان حتی مضاعف بر کردن من داشته اند و با پیشنهادها و افتخارهای مفید خود را به منیانی مراد مراحل تهیه این پایان نامه عهده دار بوده اند. لطفشان را منت دارم و تلاششان را سپاس گزارم.

فهرست مطالب

د	فهرست جداول
ه	فهرست تصاویر
و	چکیده فارسی
۱	پیش‌گفتار
۲	۱ مطالب و مفاهیم مورد نیاز
۳	۱-۱ مقدمه‌ای بر سری‌زمانی.....
۳	۲-۱ فرایندهای تصادفی.....
۶	۳-۱ الگوهای فرایندهای تصادفی ایستا.....
۶	۱-۳-۱ فرایندهای میانگین متحرک (MA).....
۷	۲-۳-۱ فرایندهای خودبرگشت (AR).....
۸	۳-۳-۱ فرایندهای آمیخته خودبرگشت-میانگین متحرک ($ARMA$).....
۹	۴-۱ توزیع نرمال.....
۱۲	۵-۱ توزیع گاما.....
۱۳	۶-۱ توزیع تی.....
۱۵	۷-۱ بسط تیلور.....
۱۶	۱-۷-۱ قضیه رول.....
۱۶	۲-۷-۱ قضیه تیلور.....
۱۷	۸-۱ مدل‌های مبتنی بر واریانس ناهمسان شرطی.....
۱۷	۱-۸-۱ واریانس شرطی.....
۱۷	۲-۸-۱ مدل واریانس ناهمسان شرطی خودبرگشت ($ARCH$).....
۱۸	۳-۸-۱ مدل واریانس ناهمسان شرطی خودبرگشت تعمیم یافته ($GARCH$).....
۱۹	۴-۸-۱ مدل واریانس ناهمسان شرطی آستانه‌ای ($GJR - GARCH$).....
۲۰	۵-۸-۱ مدل واریانس ناهمسان شرطی نمایی ($EGARCH$).....

۲ برآورد پارامترها و معیارهای ارزیابی مدل‌های پیشنهادی ۲۱

۲۲	۱-۲	برآورد پارامترهای الگوهای سری زمانی
۲۲	۱-۱-۲	استنتاج بیزی
۲۲	۲-۱-۲	تابع زیان
۲۳	۳-۱-۲	مخاطره و تصمیم بیزی
۲۴	۲-۲	زنجیر مارکف
۲۵	۱-۲-۲	ماتریس احتمال انتقال زنجیر مارکف
۲۷	۲-۲-۲	روش زنجیر مارکف مونت کارلو
۲۷	۳-۲-۲	الگوریتم متروپلیس-هستینگز (MH)
۲۸	۴-۲-۲	مراحل الگوریتم متروپلیس-هستینگز
۲۹	۳-۲	مفاهیم مربوط به ارزیابی مدل‌های سری زمانی
۲۹	۱-۳-۲	عامل بیز
۳۰	۲-۳-۲	اطلاع فشر
۳۲	۳-۳-۲	معیار اطلاعات آکائیک (AIC)
۳۳	۴-۳-۲	کمیت اطلاعات کولبک-لیبلر
۳۴	۵-۳-۲	برخی خواص کمیت کولبک-لیبلر
۳۴	۶-۳-۲	معیار انحراف اطلاعات (DIC)

۳ مدل گارچ گویا و آزمونی تجربی برای بازده سهام ۳۶

۳۷	۱-۳	مقدمه
۳۷	۲-۳	تقریب گویا
۳۸	۳-۳	تقریب گویا برای مدل‌های ایستای میانگین
۴۱	۴-۳	مدل گارچ گویا (RGARCH)
۴۲	۱-۴-۳	مدل گارچ و تبدیل‌های نامتقارن آن
۴۳	۲-۴-۳	استنتاج بیزی
۴۵	۳-۴-۳	شبیه‌سازی الگوریتم متروپلیس-هستینگز
۴۶	۵-۳	نتایج تجربی

- ۶-۳ معیار انحراف اطلاعات (*DIC*) برای آزمون نیکویی برازش هر مدل ۵۲
- ۷-۳ منحنی تأثیر اخبار (*NI*) ۵۳
- ۸-۳ نتیجه‌گیری ۵۶
- ۹-۳ پیشنهاد برای ادامه‌ی کار ۵۶

دستورات نرم‌افزار *R*

۵۷

منابع و مآخذ

۶۲

واژه‌نامه فارسی به انگلیسی

۶۶

فهرست جداول

۴۰	۱-۳	ورودی‌های جدول C برای تعیین رتبه‌ی مدل $ARMA$
۴۲	۲-۳	جدول C متناظر با مدل $ARMA(1,1)$
	نتایج برآورد پارامتر مدل‌های خانوادگی گارچ و مدل گارچ گویا با توزیع خطای نرمال	۳-۳	
۵۲		
۵۴	۴-۳	مقادیر DIC برای مدل‌های خانوادگی گارچ براساس توزیع خطای نرمال
۵۴	۵-۳	رتبه‌بندی مدل‌ها براساس معیار انحراف اطلاعات

فهرست تصاویر

- ۱-۱ ACF و $PACF$ یک فرایند تصادفی محض: $Z_t = \mu + X_t$ ۶
- ۲-۱ نمایش تابع چگالی نرمال دو متغیره در دستگاه مختصات ۱۰
- ۳-۱ نمودارهای توزیع گاما با مقادیر پارامترهای معلوم ۱۲
- ۴-۱ اثر اهرمی در نوسانات بازدهی سهام ۱۹
- ۱-۲ نمونه‌ای از زنجیر مارکف با سه حالت ۲۵
- ۱-۳ نمودارهای تغییرات روند سری‌های زمانی S_1, S_2, S_3 ۴۸
- ۲-۳ نمودارهای بازده‌های S_1, S_2, S_3 بعد از اعمال تبدیل لگاریتم تفاضلی ۴۹
- ۳-۳ نمودارهای خودهمبستگی و خودهمبستگی جزئی بازده‌های مانای سری‌های زمانی S_1, S_2 و S_3 ۵۰
- ۴-۳ نوسانات برآوردشده سری زمانی S_2 از مدل‌های خانواده گارچ و مدل گارچ گویا ... ۵۳
- ۵-۳ نمایش منحنی‌های NI مدل‌های خانواده‌ی گارچ و مدل گارچ گویا برای سری زمانی S_2 ۵۳
- ۵۶ ۵۶

چکیده:

مدل جدیدی در گارچ نامتقارن، گارچ گویا و آزمونی تجربی برای بازده سهام

امین منصوری کوهستانی

در این پایان نامه مدل جدیدی از خانواده‌ی مدل‌های گارچ با عنوان مدل گارچ گویا با جملات خطای نرمال تعریف می‌شود. براساس قضیه‌ی بیز و برآورد پارامتر بیزی، پارامترهای مدل جدید به همراه مدل‌های دیگر تحت مقایسه به دست می‌آید که برای آن از الگوریتم متروپلیس-هستینگز از روش‌های عددی زنجیر مارکوف مونت کارلو کمک گرفته شده است. با استفاده از معیار انحراف اطلاعات و منحنی‌های تأثیر اخبار، کارایی هر مدل مورد بررسی قرار می‌گیرد که براساس آن رتبه‌بندی هر مدل برای ثبت نوسانات فرایند تصادفی و پیش‌بینی تلاطم آینده، انجام می‌گیرد.

کلید واژه‌ها: مدل گارچ گویا، مدل‌های گارچ نامتقارن، اثر اهرم، تقریب‌های گویا، معیار انحراف اطلاعات، منحنی تأثیر اخبار

پیش‌گفتار

در داده‌های مالی، نوسانات بازده‌های سهام یک معیار مهم ریسک محسوب می‌شود. این نوسانات به‌طور مستقیم در بازارهای مالی قابل مشاهده نیستند و می‌بایست از کمیت‌های قابل مشاهده‌ای هم‌چون قیمت دارایی سهام برآورد شوند. استفاده از یک مدل که خواص بازده‌های دارایی را ثبت می‌کند یک روش گسترده و کارآمد برای برآورد نوسانات است. یکی از خاصیت‌های قابل توجه بازده‌های دارایی، تمایل نوسانات بزرگ به خوشه‌ها است که به آن نوسانات خوشه‌ای گفته می‌شود.

مدل‌های متعددی برای برآورد نوسانات بازده‌ها معرفی شده‌اند که مدل گارچ نامتقارن یکی از محبوب‌ترین این مدل‌ها است. تعمیم جدیدی از این مدل را با عنوان مدل گارچ گویا که از توابع گویا برای تقریب فرایند نوسانی استفاده می‌کند، معرفی کرده‌ایم و آن را با سایر مدل‌های نامتقارن خانواده $ARCH^1$ مقایسه می‌کنیم. این مقایسه را با استفاده از معیار انحراف اطلاعات و منحنی‌های تأثیر اخبار برای هرمدل انجام داده‌ایم و در نهایت نشان خواهیم داد که مدل جدید کارایی بهتری را نسبت به سایر مدل‌ها نشان می‌دهد.

معرفی فصل‌ها

- در فصل اول مفاهیم اولیه‌ی موردنیاز برای فصل‌های بعدی را تعریف می‌کنیم.
- در فصل دوم روش‌های برآورد پارامتر مدل جدید و معیارهای اطلاعات برای مقایسه با سایر مدل‌ها آورده می‌شود.
- در فصل سوم مدل جدید گارچ گویا را معرفی می‌کنیم و بعد از برآورد پارامترهای آن به روش عددی زنجیر مارکف مونت‌کارلو، با سایر مدل‌های مطرح شده مقایسه می‌کنیم.

فصل اوّل

مطالب و مفاهيم مورد نیاز

در این فصل می‌خواهیم برخی از مفاهیم موردنیاز در فصل‌های بعدی را بیان کنیم.

۱-۱ مقدمه‌ای بر سری‌زمانی

تعریف ۱-۱-۱. یک سری‌زمانی دسته‌ای از مشاهدات به شکل X_1, X_2, \dots, X_T می‌باشد که برحسب زمان T مرتب شده است. مثال‌های متعددی در زمینه‌های مختلف که حوزه‌ی تغییر آن‌ها از اقتصاد تا مهندسی است پیش می‌آید که تجزیه و تحلیل آن‌ها قسمت مهمی از آمار است.

تعریف ۲-۱-۱. سری‌زمانی $X_t : t = 1, \dots, T$ را پیوسته گوئیم هرگاه مشاهدات اجزای سری به‌طور پیوسته در زمان ایجاد شوند و درمقابل یک سری‌زمانی گسسته است هرگاه مشاهدات تنها در زمان‌های معینی که به‌طور معمول در فاصله‌های مساوی از یکدیگر قرار دارند، ایجاد شوند.

۲-۱ فرایندهای تصادفی

تعریف ۱-۲-۱. یک فرایند تصادفی مجموعه‌ای از متغیرهای تصادفی با اندیس مرحله یا زمان است که وضعیت یک پدیده یا آزمایش تصادفی را درطول یک دوره نشان می‌دهد. با درنظر گرفتن t به‌عنوان متغیر زمان و X_t متغیر تصادفی متناسب با آن، مجموعه‌ی $\{X_t : t \in T\}$ که در آن X_t به ازای هر $t \in T$ ، یک متغیر تصادفی است را یک فرایند تصادفی می‌نامند. درحالت کلی فضای حالت فرایند تصادفی (T) ، می‌تواند مقدارهای گسسته یا پیوسته را اختیار کند [۱].

تعریف ۲-۲-۱. برای یک فرایند تصادفی $\{X_t : t \in T\}$ ، تابع میانگین (دنباله‌ی میانگین) به‌صورت زیر تعریف می‌شود

$$\mu_t = E[X_t].$$

μ_t مقدار موردانتظار فرایند در زمان t است که درحالت کلی می‌تواند در هر نقطه از فضای حالت متفاوت باشد.

تابع واریانس فرایند به‌صورت زیر تعریف می‌شود

$$\sigma_t^2 = \text{Var}(X_t) = E(X_t - \mu_t)^2.$$

تابع اتوکواریانس یک فرایند برای هر تأخیر زمانی k به‌شکل زیر است

$$Y_{t,k} = \text{Cov}(X_t, X_{t+k}) = E((X_t - \mu_t)(X_{t+k} - \mu_{t+k})) = E(X_t X_{t+k}) - \mu_t \mu_{t+k}.$$

تابع خودهمبستگی^۲ (ACF) یک فرایند را برای هر تأخیر زمانی k تعریف می‌کنیم

$$\rho_{t,k} = \text{Corr}(X_t, X_{t+k}) = \frac{\text{Cov}(X_t, X_{t+k})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t+k})}} = \frac{\gamma_{t,k}}{\sqrt{\gamma_{t,t} \gamma_{k,k}}}.$$

با در نظر گرفتن X'_t و X'_{t+k} به عنوان بهترین برآوردگرهای خطی X_t و X_{t+k} ، برای تابع خودهمبستگی جزئی^۳ ($PACF$) یک فرایند تصادفی داریم

$$\phi_{t,k} = \frac{\text{Cov}[(X_t - X'_t), (X_{t+k} - X'_{t+k})]}{\sqrt{\text{Var}(X_t - X'_t)} \sqrt{\text{Var}(X_{t+k} - X'_{t+k})}}.$$

نکته ۱-۲-۳. تابع خودهمبستگی برای همه‌ی فرایندها خوش تعریف نیست چون ممکن است واریانس برابر صفر (برای یک فرایند ثابت) یا بی‌نهایت باشد. اگر تابع خودهمبستگی خوش تعریف باشد مقدار آن در محدوده‌ی $[-1, 1]$ قرار می‌گیرد، که مقدار یک، نشان‌دهنده‌ی همبستگی کامل بین X_t و X_{t+k} و منفی یک، نشان‌دهنده‌ی ضد همبستگی کامل است. اگر مقدار تابع صفر گردد، X_t و X_{t+k} را ناهمبسته گوئیم و اگر X_t و X_{t+k} مستقل باشند آنگاه خودهمبستگی آن‌ها صفر خواهد بود [۲].

تعریف ۱-۲-۴. به‌طور کلی مانایی به این معنا است که قوانین احتمالی حاکم بر فرایند تصادفی با زمان تغییر نمی‌کند یعنی فرایند در تعادل آماری است. یک فرایند را مانای اکید گوئیم هرگاه به‌ازای همه‌ی مقادیر t_1, t_2, \dots, t_n و برای هر تأخیر زمانی h ، توزیع توأم متغیرهای X_{t_1}, \dots, X_{t_n} مانند توزیع توأم $X_{t_1+h}, \dots, X_{t_n+h}$ باشد. به‌بیان دیگر با تغییر زمان اولیه به اندازه‌ی h توزیع توأم متغیرها تغییری نداشته باشد.

تعریف ۱-۲-۵. یک فرایند تصادفی مانای ضعیف خواهد بود اگر دارای میانگین و واریانس ثابت باشد و کواریانس آن تنها تابعی از اختلاف زمانی متغیرهای تصادفی باشد، یعنی داشته باشیم

$$E[X_t] = \mu \quad \text{ثابت}$$

$$\text{Var}(X_t) = E[(X_t - \mu_t)^2] = \sigma^2 < \infty$$

$$\gamma_{t,s} = \text{Cov}(X_t, X_s) = f(t - s) \quad t < s.$$

Autocorrelation Function^۲

Partial Autocorrelation Function^۳

این سه شرط بیان می‌کند که یک فرایند مانا باید دارای میانگین، واریانس و ساختار اتوکواریانس ثابت باشد.

نکته ۱-۲-۶. اگر فرایند تصادفی مانا باشد، تابع خودهمبستگی $\rho_{t,k} = \text{Corr}(X_t, X_{t+k})$ فقط به تأخیر زمانی k بستگی دارد. بنابراین می‌توان آن را با ρ_k نشان داد.

نکته ۱-۲-۷. در یک فرایند تصادفی، تابع خودهمبستگی جزئی در تأخیر k ام در واقع برابر

$$\phi_{t,k} = \text{Corr}(X_t, X_{t+k} \mid X_{t+1}, X_{t+2}, \dots, X_{t+k-1})$$

است که در شرایط مانایی، فقط به k بستگی دارد و آن را با ϕ_{kk} نشان داده و خواهیم داشت

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}.$$

تعریف ۱-۲-۸. فرایند $\{X_t: t \in T\}$ را تصادفی محض (نوفه سفید)^۴ و یا اغتشاش خالص می‌نامند هرگاه دنباله‌ای از متغیرهای تصادفی ناهمبسته از توزیعی ثابت با میانگین ثابت $E[X_t] = \mu$ ، که معمولاً صفر فرض می‌شود، واریانس ثابت $\text{Var}(X_t) = \sigma^2$ و برای هر $k \neq 0$ ،

$$\gamma_{t,k} = \text{Cov}(X_t, X_{t+k}) = 0.$$

با این تعریف می‌توان این چنین عنوان کرد که فرایند تصادفی محض $\{X_t: t \in T\}$ ، ماناست و داریم

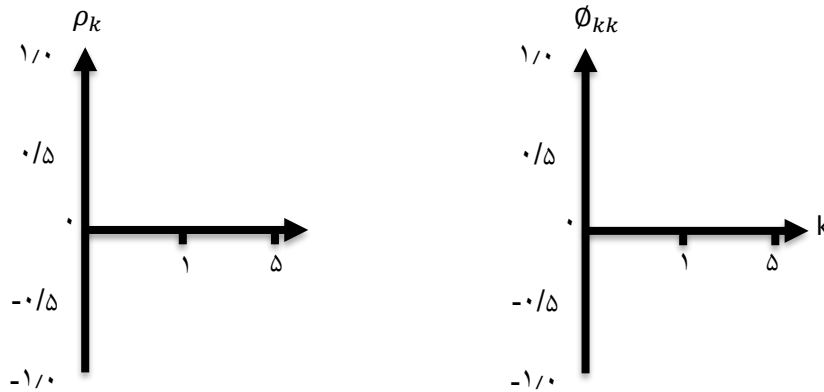
$$\gamma_k = \begin{cases} \sigma^2 & k = 0 \\ 0 & k \neq 0 \end{cases}, \quad \rho_k = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}, \quad \phi_{kk} = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases}.$$

مشخصه‌ی اصلی فرایندهای تصادفی محض این است که ACF و $PACF$ آن برابر صفر است.

شکل (۱-۱)، ACF و $PACF$ یک فرایند تصادفی محض را نشان می‌دهد که در وقفه‌های زمانی k ، مقادیر

خودهمبستگی و خودهمبستگی جزئی برابر صفر مشاهده می‌شود و مقدار یک را در $k = 0$ دارد [۳].

^۴ White Noise

شکل ۱-۱: ACF و PACF یک فرایند تصادفی محض: $Z_t = \mu + X_t$

یک فرایند تصادفی محض گاوسی است هرگاه توزیع توأم آن، نرمال باشد.

۳-۱ الگوهای فرایندهای تصادفی مانا

به منظور پیش‌بینی روند یک فرایند تصادفی با تعداد متناهی از مشاهدات از یک الگوی پارامتری از مرتبه‌ی متناهی برای بیان رفتار آن، استفاده می‌کنیم.

۱-۳-۱ فرایندهای میانگین متحرک (MA)

اسلاتسکی (۱۹۷۲)^۶ و والد (۱۹۳۸)^۷ برای اولین بار فرایندهای تصادفی میانگین متحرک را برای الگودهی به سری‌های زمانی ارائه کردند. براین اساس فرایند $\{Z_t\}$ ، میانگین متحرک از مرتبه‌ی q می‌باشد اگر داشته باشیم

$$Z_t = \mu + X_t + \theta_1 X_{t-1} + \dots + \theta_q X_{t-q}.$$

که در آن $\{X_t\}$ ، فرایند تصادفی محض با میانگین صفر و واریانس σ_x^2 است. براساس نمایش فرایند تصادفی میانگین متحرک با به‌کارگیری عملگر پسر و L به صورت

$$L^i Z_t = Z_{t-i}$$

می‌توان معادله‌ی فرایند را به شکل زیر بازنویسی کرد

$$Z_t = \mu + \sum_{i=1}^q \theta_i L^i X_t + X_t \quad \text{یا} \quad Z_t = \mu + \theta(L) X_t$$

^۵ Moving Average

^۶ Eugen Slutsky

^۷ Abraham Wald

که در آن

$$\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q .$$

چون $1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2 < \infty$ ، لذا یک فرایند میانگین متحرک، همواره مانا است و نیاز به هیچ محدودیتی روی ضرایب $\{\theta_i\}$ ندارد اما شرط لازم برای وارون پذیری این است که تمامی ریشه‌های معادله $\theta(L) = 0$ خارج دایره‌ی واحد، واقع شوند [۴]، [۵].

برای برخی از خواص فرایند میانگین متحرک از مرتبه‌ی q داریم

$$E(Z_t) = \mu$$

$$Var(Z_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2$$

$$Cov(Z_t, Z_s) = \gamma_{t,s} = \begin{cases} (\theta_s + \theta_{s+1}\theta_1 + \theta_{s+2}\theta_2 + \dots + \theta_q\theta_{q-s}) \sigma^2 & s = 1, 2, \dots, q \\ 0 & s > q \end{cases}$$

۲-۳-۱ فرایندهای خودبرگشت (AR)^۱

اگر $\{X_t\}$ را یک فرایند تصادفی محض با میانگین صفر و واریانس محدود σ_x^2 در نظر بگیریم، $\{Z_t\}$ یک فرایند خودبرگشت از مرتبه‌ی p نام دارد، اگر

$$Z_t = a_1 Z_{t-1} + a_2 Z_{t-2} + \dots + a_p Z_{t-p} + \mu + X_t .$$

نمایش کلی فرایند تصادفی خودبرگشت به صورت $Z_t = \mu + \sum_{i=1}^p a_i Z_{t-i} + X_t$ است که با استفاده از عملگر پسر و آن را می‌توان به صورت $Z_t = \mu + \sum_{i=1}^p a_i C^i Z_t + X_t$ نوشت. با فرض صفر بودن میانگین فرایند $\{Z_t\}$ ، می‌توان این رابطه را به صورت $a(L)Z_t = X_t$ بازنویسی کرد که در آن خواهیم داشت

$$a(L) = (1 - a_1 L - a_2 L^2 - \dots - a_p L^p) .$$

فرایندهای تصادفی خودبرگشت از مرتبه‌ی p را می‌توان با به‌کارگرفتن عملگر پسر و به صورت فرایندهای میانگین متحرک از مرتبه‌ی بی‌نهایت نیز تبدیل کرد

$$a(L)Z_t = X_t \Rightarrow (1 - a_1 L - a_2 L^2 - \dots - a_p L^p)Z_t = X_t$$

^۱ Autoregressive

$$\Rightarrow Z_t = X_t / (1 - a_1 L - a_2 L^2 - \dots - a_p L^p) = f(L) X_t .$$

که در آن

$$f(L) = (1 - a_1 L - a_2 L^2 - \dots - a_p L^p)^{-1} = (1 + b_1 L + b_2 L^2 + \dots) .$$

بنابراین اگر فرایند مانا باشد، ضرایب $MA(\infty)$ در نهایت همراه با طول وقفه کاهش می یابد.

چون $\sum_{j=1}^p |a_j| < \infty$ ، لذا فرایند همواره وارون پذیر است. برای مانایی، باید ریشه های معادله ی

$$a_p(L) = 0 \text{ خارج دایره ی واحد، واقع شوند [۴].}$$

۳-۳-۱ فرایندهای آمیخته خودبرگشت-میانگین متحرک (ARMA)

مشکل فرایندهای میانگین متحرک و خودبرگشت این است که در برخی مواقع تعداد پارامترهای زیادی را دربر می گیرد. این موضوع حتی برای یک میانگین متحرک و یا خودبرگشت از مرتبه ی متناهی نیز درست است، زیرا اغلب یک الگو با مرتبه ی بالاتر را برای یک تقریب خوب در نظر می گیرد، در نتیجه کارایی در برآورد کردن پارامترها کاهش می یابد. بنابراین، در ساختن یک الگو، لحاظ کردن جملات خودبرگشت و میانگین متحرک ممکن است لازم شود که به فرایند مرکب خودبرگشت-میانگین متحرک (ARMA) به شکل زیر منتهی می شود

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + X_t - \theta_1 X_{t-1} - \dots - \theta_q X_{t-q} . \quad (1-1)$$

با استفاده از عملگر پسرو می توان مدل $ARMA(p, q)$ را دوباره نویسی کرد

$$\phi_p(B) Z_t = \theta_q(B) X_t$$

که در آن

$$\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q .$$

برای وارون پذیری فرایند، باید ریشه های معادله ی $\theta_q(B) = 0$ خارج دایره ی واحد، واقع شوند. هم چنین

برای مانایی، لازم است ریشه های معادله ی $\phi_p(B) = 0$ خارج دایره ی واحد قرار گیرد. لازم به ذکر است

که معمولاً فرض می شود $\phi_p(B) = 0$ و $\theta_q(B) = 0$ ریشه ی مشترک ندارند [۴].

تعریف ۱-۳-۱. (تابع خودهمبستگی فرایند آمیخته). با ضرب کردن طرفین معادله‌ی (۱-۱) در عبارت

Z_{t-k} و محاسبه‌ی امیدریاضی آن، تابع اتوکواریانس در معادله‌ی تفاضلی زیر به دست می‌آید

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + \gamma_{zx}(k) - \theta_1 \gamma_{zx}(k-1) - \dots - \theta_q \gamma_{zx}(k-q). \quad (2-1)$$

$\gamma_{zx}(k)$ ، تابع کواریانس متقابل بین Z و X است و به وسیله‌ی $\gamma_{zx} = E[Z_{t-k} \cdot X_t]$ تعریف می‌شود.

چون Z_{t-k} فقط به ضرباتی که تا زمان $t-k$ رخ داده‌اند بستگی دارد، داریم

$$\begin{cases} \gamma_{zx} = 0 & k > 0 \\ \gamma_{zx} = 0 & k \leq 0 \end{cases}$$

ملاحظه می‌کنیم که از (۲-۱) چنین برمی‌آید که

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} \dots + \phi_p \gamma_{k-p} \quad k \geq q+1$$

و بنابراین

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \dots + \phi_p \rho_{k-p} \quad k \geq q+1$$

یا

$$\phi_p(B) \rho_k = 0.$$

پس، برای فرایند $ARMA(p, q)$ به تعداد q خودهمبستگی وجود خواهند داشت که مقادیر آنها، از طریق رابطه‌ی (۲-۱) به صورت مستقیم به انتخاب q پارامتر مربوط به میانگین متحرک θ و نیز به p پارامتر مربوط به خودبرگشت ϕ بستگی دارند.

۴-۱ توزیع نرمال

تعریف ۱-۴-۱. بردار تصادفی $X = [x_1, \dots, x_n]^T$ دارای توزیع نرمال چندمتغیره (گاوسی) با پارامتر مکان $\mu \in \mathbb{R}^n$ و پارامتر مقیاس $\Sigma \in S_{++}^n$ که در آن فضای ماتریس‌های $n \times n$ معین مثبت و متقارن به صورت

$$S_{++}^n = \left\{ A \in \mathbb{R}^{n \times n}; A = A^T \text{ \& } X^T A X > 0; X \neq 0 \right\}$$

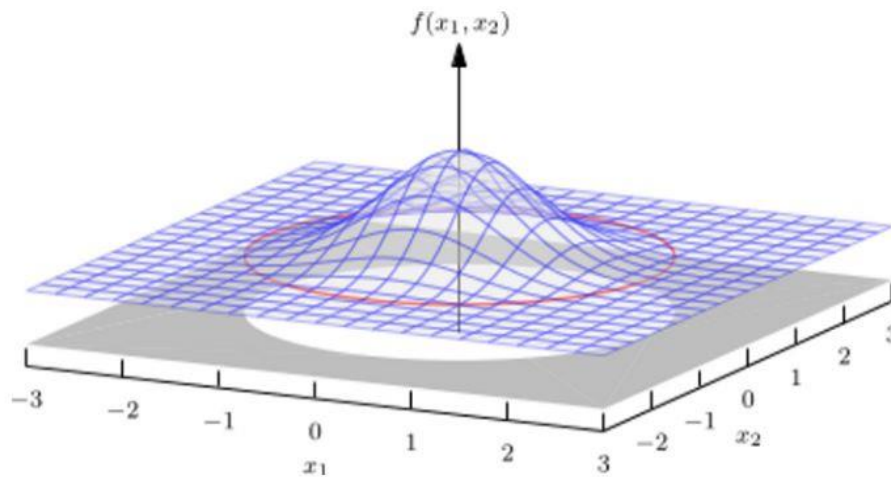
است، دارای تابع چگالی احتمال به شکل زیر می‌باشد

$$P(X; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (X - \mu)^T \Sigma^{-1} (X - \mu)\right).$$

ضریب $\frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}}$ به بردار تصادفی X بستگی ندارد از این رو می‌تواند به‌عنوان یک عامل نرمال برای پوشش کل فضای احتمال در نظر گرفته شود [۶].

$$\frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2} (X - \mu)^T \Sigma^{-1} (X - \mu)\right) dx_1 dx_2 \dots dx_n = 1.$$

شکل (۲-۱) چگالی توزیع نرمال معیار دومتغیره را با میانگین صفر در مبدأ دستگاه مختصات نشان می‌دهد.



شکل ۲-۱: نمایش تابع چگالی نرمال دومتغیره در دستگاه مختصات

قضیه ۲-۴-۱. برای هر بردار تصادفی X با میانگین μ و ماتریس کواریانس Σ داریم

$$\Sigma = E[(X - \mu)(X - \mu)^T] = E[XX^T] - \mu\mu^T.$$

برهان. ماتریس کواریانس Σ در تعریف تابع چگالی احتمال چند متغیره باید معین مثبت و متقارن باشد.
 $(\Sigma \in S_{++}^n)$

$$\begin{aligned} \Sigma &= \begin{bmatrix} \text{Cov}(x_1, x_1) & \dots & \text{Cov}(x_1, x_n) \\ \vdots & \dots & \vdots \\ \text{Cov}(x_n, x_1) & \dots & \text{Cov}(x_n, x_n) \end{bmatrix} \\ &= \begin{bmatrix} E[(x_1 - \mu_1)^2] & \dots & E[(x_1 - \mu_1)(x_n - \mu_n)] \\ \vdots & \dots & \vdots \\ E[(x_n - \mu_n)(x_1 - \mu_1)] & \dots & E[(x_n - \mu_n)^2] \end{bmatrix} \\ &= E \begin{bmatrix} (x_1 - \mu_1)^2 & \dots & (x_1 - \mu_1)(x_n - \mu_n) \\ \vdots & \dots & \vdots \\ (x_n - \mu_n)(x_1 - \mu_1) & \dots & (x_n - \mu_n)^2 \end{bmatrix} \end{aligned}$$

$$= E \left[\begin{bmatrix} x_1 - \mu_1 \\ \vdots \\ x_n - \mu_n \end{bmatrix} [x_1 - \mu_1 \dots x_n - \mu_n] \right]$$

$$= E[(X - \mu)(X - \mu)^T].$$

تعریف ۳-۴-۱. برای یک تابع $f: \mathbb{R}^n \rightarrow \mathbb{R}$ یک منحنی سطح از تابع f یک مجموعه به شکل $\{X \in \mathbb{R}^n; f(x) = C\}$ برای برخی $C \in \mathbb{R}$ است. در یک حالت $n = 2$ و ماتریس کواریانس قطری، توزیع گاوسی چند متغیره دارای چگالی به صورت زیر است

$$P(X; \mu, \Sigma) = \frac{1}{\sqrt{2\pi}\sigma_1\sigma_2} \exp\left(-\frac{1}{2\sigma_1^2}(x_1 - \mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2 - \mu_2)^2\right).$$

در این صورت منحنی سطح شامل تمامی نقاط $x_1, x_2 \in \mathbb{R}$ خواهد بود که برای برخی ثابت‌های $C \in \mathbb{R}$ ، $P(X; \mu, \Sigma) = C$ از این رو خواهیم داشت

$$C = \frac{1}{\sqrt{2\pi}\sigma_1\sigma_2} \exp\left(-\frac{1}{2\sigma_1^2}(x_1 - \mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2 - \mu_2)^2\right)$$

$$\sqrt{2\pi}C\sigma_1\sigma_2 = \exp\left(-\frac{1}{2\sigma_1^2}(x_1 - \mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2 - \mu_2)^2\right)$$

$$\text{Log}(\sqrt{2\pi}C\sigma_1\sigma_2) = -\frac{1}{2\sigma_1^2}(x_1 - \mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2 - \mu_2)^2$$

$$\text{Log}\left(\frac{1}{\sqrt{2\pi}C\sigma_1\sigma_2}\right) = \frac{1}{2\sigma_1^2}(x_1 - \mu_1)^2 + \frac{1}{2\sigma_2^2}(x_2 - \mu_2)^2.$$

تعریف می‌کنیم

$$r_2 = \sqrt{2\sigma_2^2 \text{Log}\left(\frac{1}{\sqrt{2\pi}C\sigma_1\sigma_2}\right)}, \quad r_1 = \sqrt{2\sigma_1^2 \text{Log}\left(\frac{1}{\sqrt{2\pi}C\sigma_1\sigma_2}\right)}$$

در این صورت داریم

$$1 = \left(\frac{x_1 - \mu_1}{r_1}\right)^2 + \left(\frac{x_2 - \mu_2}{r_2}\right)^2.$$

این معادله، منحنی سطح توزیع نرمال در حالت $n = 2$ را نشان می‌دهد، که معادله‌ی یک بیضی با مرکز (μ_1, μ_2) است و در آن x_1 دارای طول r_1 و x_2 دارای طول r_2 خواهد بود [۷].

۵-۱ توزیع گاما

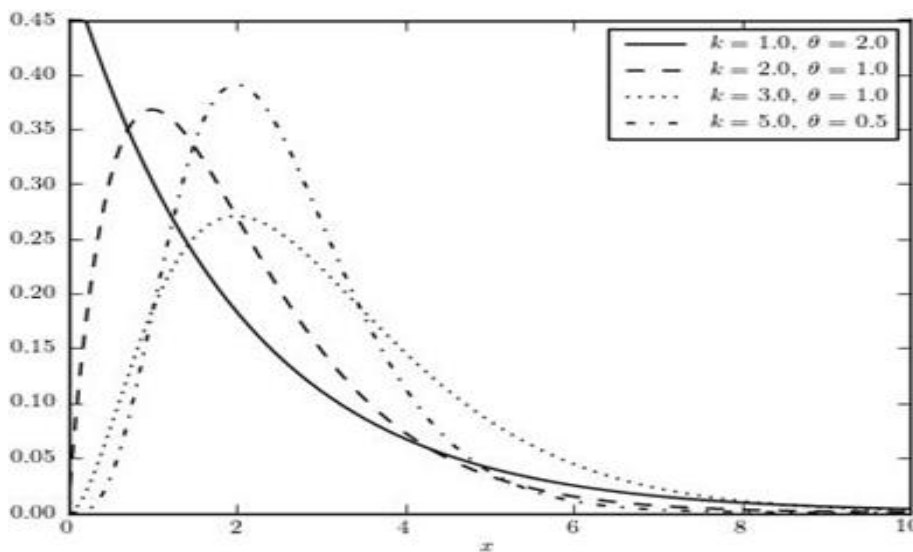
تعریف ۱-۵-۱. تابع چگالی احتمال $f(x; k, \theta)$ دارای توزیع گاما با پارامتر شکل k و پارامتر مقیاس θ خواهد بود، اگر داشته باشیم

$$f(x; k, \theta) = \begin{cases} \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

که در آن $\Gamma(k)$ ، تابع گاما است

$$\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x} dx = (k-1)\Gamma(k-1).$$

شکل زیر نمودارهای توزیع گاما با مقادیر مختلف از پارامترهای مقیاس و شکل را نشان داده است که با تغییر پارامتر شکل ($k > 1$)، منحنی تابع چگالی تغییرات زیادی در کشیدگی را دربر دارد.



شکل ۱-۳: نمودارهای توزیع گاما با مقادیر پارامترهای معلوم

قضیه ۱-۵-۲. اگر یک دنباله از متغیرهای تصادفی مستقل به صورت

$$X_1 \sim \Gamma(k_1, \theta), \dots, X_n \sim \Gamma(k_n, \theta)$$

داشته باشیم در آن صورت $Y = \sum_{i=1}^n X_i$ دارای توزیع $\Gamma(\sum_{i=1}^n k_i, \theta)$ خواهد بود [۸].

برهان. اگر $X \sim \Gamma(k, \theta)$ ، تابع مولدگشتاور X به ازای هر مقدار t به صورت زیر تعریف می شود

$$\begin{aligned}
 M_x(t) &= E(e^{tx}) = \int_0^{+\infty} e^{tx} \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x} dx \\
 &= \int_0^{+\infty} \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-(\theta-t)x} dx \\
 &= \frac{\theta^k}{(\theta-t)^k} \underbrace{\int_0^{+\infty} \frac{(\theta-t)^k}{\Gamma(k)} x^{k-1} e^{-(\theta-t)x} dx}_A.
 \end{aligned}$$

عبارت A انتگرال روی کل بازه‌ی احتمال توزیع $\Gamma(k, \theta - t)$ است و برابر یک در نظر گرفته می‌شود.

بنابراین

$$M_x(t) = \left(\frac{\theta}{\theta-t}\right)^k \quad t < \theta.$$

باتوجه به این تعریف برای متغیر تصادفی $Y = \sum_{i=1}^n X_i$ تابع مولدگشتاور به صورت زیر خواهد بود

$$\begin{aligned}
 M_Y(t) &= E(e^{tY}) = \prod_{i=1}^n E(e^{tX_i}) = \left(\frac{\theta}{\theta-t}\right)^{\sum_{i=1}^n k_i} \quad t < \theta \\
 &: Y = \sum_{i=1}^n X_i \sim \Gamma\left(\sum_{i=1}^n k_i, \theta\right).
 \end{aligned}$$

نکته ۱-۵-۳. N امین گشتاور مرکزی X با توزیع گاما برابر است با

$$E(X^N) = \frac{(k+N-1)(k+N-2) \dots k}{\theta^N}.$$

در این حالت برای میانگین و واریانس توزیع $X \sim \Gamma(k, \theta)$ خواهیم داشت

$$N = 1; \quad E(X) = \mu = \frac{k}{\theta}, \quad N = 2; \quad E(X^2) = \frac{(k+1)k}{\theta^2}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{(k+1)k}{\theta^2} - \left(\frac{k}{\theta}\right)^2 = \frac{k}{\theta^2}.$$

۶-۱ توزیع تی

تعریف ۱-۶-۱. بردار تصادفی X با درجه‌ی آزادی ν ، با پارامترهای بردار میانگین μ (با بُعد p) و ماتریس

$p \times p$ مثبت معین Σ ، دارای توزیع تی است هرگاه تابع چگالی آن به صورت زیر باشد

$$St(x; \mu, \Sigma, \nu) = \frac{|\Sigma|^{-\frac{1}{2}}}{C(\nu, p) \pi^{p/2}} \left[1 + \frac{1}{\nu} (x - \mu)^T \Sigma^{-1} (x - \mu) \right]^{-\frac{(\nu+p)}{2}}. \quad (3-1)$$

که در آن X ، مقادیر مشخص شده از یک بردار تصادفی $X, p \times 1$ می‌باشد. تابع $C(v, p)$ در این تعریف مستقل از X است و داریم

$$\Gamma\left(\frac{v+p}{2}\right) C(v, p) = v^{p/2} \Gamma\left(\frac{v}{2}\right).$$

ماتریس کواریانس برای توزیع تی به صورت $\Sigma^* = v^* \cdot \Sigma$ تعریف می‌شود که در آن

$$v^* = \frac{v}{v-2}.$$

نکته ۱-۶-۲. بردار تصادفی X دارای توزیع تی چند متغیره، می‌تواند به روش‌های مختلفی تولید شود. یک روش مرسوم، ترکیب توزیع نرمال چند متغیره $N_p(\mu; \omega^2 \Sigma)$ و ω است که دارای توزیع گامای معکوس شده است [۹]

$$h(\omega) = \frac{2 \left(\frac{v}{2}\right)^{v/2}}{\Gamma\left(\frac{v}{2}\right)} \omega^{-(v+1)} \exp\left(\frac{-v}{2\omega^2}\right).$$

v درجه‌ی آزادی توزیع گامای معکوس شده می‌باشد. در این شرایط توزیع تی به صورت زیر تولید خواهد شد

$$f(x) = \int_0^\infty \frac{|\omega^2 \Sigma|^{-\frac{1}{2}}}{(2\pi)^{p/2}} \exp\left[\frac{-(x-\mu)^T (\omega^2 \Sigma)^{-1} (x-\mu)}{2}\right] h(\omega) d\omega.$$

تعریف ۱-۶-۳. تابع چگالی احتمال مشترک از N بردار تصادفی مستقل و هم‌توزیع که هر کدام دارای توزیع تی می‌باشند، برابر است با

$$f(x_1, x_2, \dots, x_N) = f(x_1) f(x_2) \dots f(x_N) = \frac{|\Sigma|^{-\frac{N}{2}}}{C(v, p) \pi^{\frac{Np}{2}}} \left(1 + \frac{Q}{v}\right)^{-\frac{(v+Np)}{2}}.$$

که در آن $Q = \sum_{j=1}^N (x_j - \mu)^T \Sigma^{-1} (x_j - \mu)$ و $j = 1, 2, \dots, N$ مقادیر مشخص از یک بردار تصادفی X است.

نکته ۱-۶-۴. در رابطه‌ی (۳-۱)، اگر $p = 1$ ، $\mu = 0$ و $\Sigma = 1$ ، تابع چگالی احتمال توزیع تی در حالت تک متغیره به دست خواهد آمد

$$f(x) = \frac{1}{C(v, 1) \sqrt{\pi}} \left(1 + \frac{x^2}{v}\right)^{-\frac{(v+1)}{2}} \quad v > 0.$$

که ν نشان دهنده‌ی درجه‌ی آزادی توزیع است. همچنین این تابع را می‌توان به صورت

$$f(x) = \frac{1}{\sqrt{\nu} \beta \left(\frac{1}{\nu}, \frac{\nu}{\nu}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}$$

نوشت که $\beta(\cdot, \cdot)$ ، توزیع بتا است و داریم

$$\beta(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}.$$

قضیه ۱-۶-۵. تابع مشخصه‌ی X که دارای توزیع تی چند متغیره با تابع چگالی احتمال $(3-1)$ است، به

شکل زیر است

$$\Phi_X(t) = E(e^{it^T X}) = e^{it^T \mu} \frac{\|(\nu \Sigma)^{\frac{1}{\nu}} t\|^{\frac{\nu}{\nu}}}{\frac{\nu}{2^{\frac{\nu}{2}}-1} \Gamma(\frac{\nu}{2})} K_{\frac{\nu}{2}} \|(\nu \Sigma)^{\frac{1}{\nu}} t\|.$$

که در آن $K_{\frac{\nu}{2}} \|(\nu \Sigma)^{\frac{1}{\nu}} t\|$ تابع مکدونالد^۹ از مرتبه‌ی $\frac{\nu}{2}$ و با مشخصه‌ی $\|(\nu \Sigma)^{\frac{1}{\nu}} t\|$ است.

تابع مکدونالد $K_{\alpha}(t)$ با رتبه‌ی α و مشخصه‌ی t به وسیله‌ی انتگرال زیر تعریف می‌شود

$$K_{\alpha}(t) = \left(\frac{\nu}{t}\right)^{\alpha} \frac{\Gamma\left(\alpha + \frac{1}{\nu}\right)}{\sqrt{\pi}} \int_0^{\infty} (1 + u^2)^{-(\alpha+1)} \cos(tu) du \quad t > 0, \alpha > \frac{1}{\nu}.$$

□

برهان. [۱۰].

نتیجه ۱-۶-۶. برای تابع مشخصه‌ی توزیع تی در حالت تک متغیره خواهیم داشت

$$\Phi_X(t) = \frac{\nu^{\frac{\nu}{2}} |t|^{\frac{\nu}{2}}}{\frac{\nu}{2^{\frac{\nu}{2}}-1} \Gamma(\nu - 2)} K_{\frac{\nu}{2}}(\sqrt{\nu} |t|) \quad \nu > 2.$$

۷-۱ بسط تیلور^{۱۰}

در مواردی که تابع f و مشتق‌های بالاتر آن موجود باشند، قضیه‌ی تیلور برای تقریب تابع توسط چند جمله‌ای‌های خطی کاربرد دارد.

Macdonald Function ^۹

Taylor ^{۱۰}

قضیه رول^{۱۱} ۱-۷-۱. اگر تابع f در بازه $[a, b]$ پیوسته باشد و روی بازه (a, b) مشتق پذیر باشد و داشته باشیم

$$f(a) = f(b) \quad a < b.$$

در این صورت یک c در (a, b) موجود خواهد بود که $f'(c) = 0$ ، به این معناکه تابع f در بازه (a, b) یک خط مماس افقی دارد.

برهان. [۱۱]. □

قضیه تیلور ۲-۷-۱. اگر f یک تابع $(n+1)$ بار مشتق پذیر روی یک بازه شامل نقاط a و x باشد، بنابراین

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x).$$

که در آن $R_n(x)$ ، باقیمانده‌ی تقریب نام دارد و خطای تقریب تابع f را برآورد می‌کند

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} \quad a < c < x.$$

برهان. [۱۱]. □

نتیجه ۳-۷-۱. در بسط تیلور با در نظر گرفتن $a = 0$ خواهیم داشت

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + R_n(x).$$

که برای برخی نقاط c درون $(0, x)$ ، خطای تقریب به صورت زیر تعریف می‌شود

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}.$$

برای مقادیری از x که در آن $\lim_{n \rightarrow \infty} R_n(x) = 0$ ، بسط سری توانی زیر برای تابع f با عنوان بسط

مکلورن^{۱۲} از تابع f نامیده می‌شود

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n \quad ; \quad f^{(0)}(0) = f(0).$$

Roll^{۱۱}

Maclaurin^{۱۲}

۸-۱ مدل‌های مبتنی بر واریانس ناهمسان شرطی

شدت تغییرپذیری در سری‌های زمانی که به‌طور معمول با معیارهای واریانس یا انحراف معیار اندازه‌گیری می‌شوند در بسیاری از سری‌ها از جمله نرخ‌ارز، قیمت نفت، بازدهی سهام، تورم و ... ثابت نیست که به‌عنوان تغییرپذیری خوشه‌ای از آن یاد می‌شود. در این نوع تغییرپذیری نوسانات و تغییرات بزرگ منجر به تغییرات بزرگ و نوسانات و تغییرات کوچک منجر به تغییرات کوچک در سری‌های زمانی شده که این پدیده با عنوان اثر اهرم شناخته می‌شود.

۱-۸-۱ واریانس شرطی

در نوسان‌پذیری خوشه‌ای، سطح جاری تغییرپذیری رابطه‌ی مستقیم و مثبتی با مقادیر گذشته‌ی خود دارد. برای بیان جبری این فرایند می‌بایست میانگین و واریانس بلندمدت (غیرشرطی) و کوتاه‌مدت (شرطی) را تعریف کنیم.

تعریف ۱-۸-۱. چنانچه $\{Y_t\}_1^T$ نشان‌دهنده‌ی یک سری زمانی با طول T باشد امیدریاضی یا میانگین بلندمدت سری که به آن میانگین غیرشرطی نیز گفته می‌شود، به‌صورت $E(Y_t)$ نشان داده می‌شود و در مقابل برای امیدریاضی کوتاه‌مدت و یا در اصطلاح میانگین شرطی سری، خواهیم داشت

$$E_{t-1}(Y_t) = E(Y_t | Y_{t-1}, Y_{t-2}, \dots).$$

واریانس بلندمدت یا غیرشرطی یک سری زمانی به‌صورت $Var(Y_t) = E(Y - E(Y))^2$ است. نمایش شکل کوتاه‌مدت یا شرطی واریانس Y_t ، به‌شکل زیر خواهد بود

$$Var(Y_t | Y_{t-1}, Y_{t-2}, \dots) = E_{t-1}(Y_t - E_{t-1}(Y_t))^2.$$

توجه شود که منظور از کوتاه‌مدت و یا شرطی این است که اطلاعات تا دوره‌ی $t - 1$ موجود است [۱۲]، [۱۳].

۲-۸-۱ مدل واریانس ناهمسان شرطی خودبرگشت^{۱۳} (ARCH)

براساس نظریه‌ی رابرت انگل^{۱۴} (۱۹۸۲)، جزء خطا در مدل‌هایی که دچار واریانس ناهمسان شرطی هستند، در ساده‌ترین حالت به‌شکل

^{۱۳} Autoregressive Conditional Heteroscedasticity

^{۱۴} Engle

$$u_t = \varepsilon_t \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2} \quad \alpha_0 \geq 0, \quad 0 < \alpha_1 < 1.$$

مدل‌بندی می‌شوند که در آن ε_t ، فرایند اغتشاش خالص با میانگین صفر و واریانس یک است. در این گونه مدل‌ها برای میانگین و واریانس بلندمدت جزء خطا داریم

$$E(u_t) = 0, \quad \text{Var}(u_t) = \frac{\alpha_0}{1 - \alpha_1}.$$

میانگین و واریانس کوتاه‌مدت جزء خطا نیز به صورت زیر خواهد بود

$$E_{t-1}(u_t) = 0, \quad \text{Var}_{t-1}(u_t) = \alpha_0 + \alpha_1 u_{t-1}^2.$$

بنابراین چنانچه واریانس شرطی u_t را با σ_t^2 نشان دهیم مدل واریانس ناهمسان شرطی خودبرگشت به صورت یک سیستم دو معادله‌ای تعریف می‌شود [۱۴]

$$\begin{cases} y_t = \phi_1 y_{t-1} + u_t \\ \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 \end{cases}$$

این مدل از نظر میانگین $AR(1)$ و به جهت این که شامل وقفه‌ی مرتبه‌ی اول مجذور جزء خطا است، $ARCH(1)$ نام‌گذاری می‌شود، بنابراین $AR(1) + ARCH(1)$ است که با افزودن وقفه‌های زمانی بالاتر به معادلات میانگین و واریانس، می‌توان مدل $ARMA(p, q) + ARCH(m)$ را به دست آورد

$$\begin{cases} y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + u_t + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q} \\ \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_m u_{t-m}^2 \end{cases}$$

۳-۸-۱ مدل واریانس ناهمسان شرطی خودبرگشت تعمیم‌یافته^{۱۵} ($GARCH$)

بولرسلف^{۱۶} در سال ۱۹۸۶ تعمیم مدل $ARCH$ را به منظور پوشش نقطه ضعف اساسی این مدل که تعیین مقدار وقفه‌های معادله‌ی واریانس می‌باشد، ارائه داد. در حالت ساده معادله‌ی $GARCH(1,1)$ را به شکل

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2$$

تعریف می‌کنند که با جایگذاری وقفه‌های گذشته و ادامه دادن معادله‌ی بالا بدیهی است که مدل $GARCH(1,1)$ همان $ARCH(\infty)$ است، از این رو سبب رعایت اصل صرفه‌جویی مد نظر باکس و جنکینز^{۱۷} در مدل‌های $ARMA$ نیز می‌گردد.

^{۱۵} Generalized Autoregressive Conditional Heteroscedasticity

^{۱۶} Bollerslev

^{۱۷} Box & Jenkins

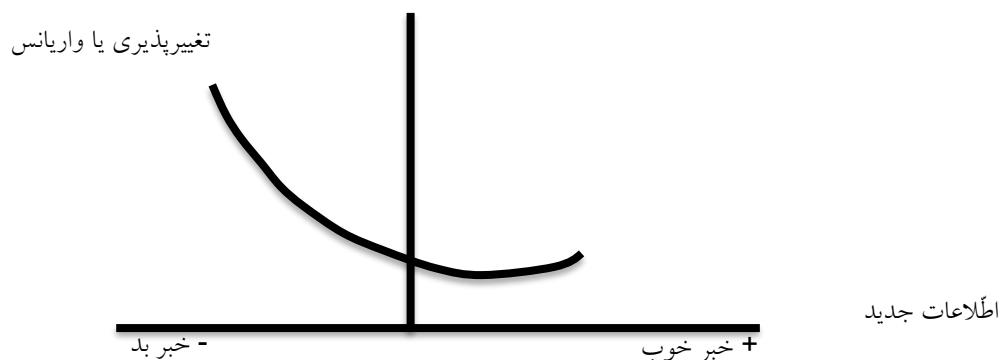
با گسترش دادن رابطه‌ی قبل به وقفه‌های زمانی بالاتر، مدل $ARMA(p, q) + GARCH(m, n)$ تعریف خواهد شد

$$\begin{cases} y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j u_{t-j} & ; \quad u_t \sim N(0, \sigma_t^2) \\ \sigma_t^2 = \alpha_0 + \sum_{l=1}^m \alpha_l u_{t-l}^2 + \sum_{k=1}^n \beta_k \sigma_{t-k}^2 & ; \quad \alpha_0 > 0 ; 0 \leq \alpha_l, \beta_k < 1 \end{cases}$$

[۱۵], [۱۶].

۴-۸-۱ مدل واریانس ناهمسان شرطی آستانه‌ای^{۱۸} ($GJR-GARCH$)

در مقایسه با مدل‌های $ARCH$ و $GARCH$ که در آن اثر شوک‌های مثبت و منفی بر معادله‌ی واریانس، یکسان و متقارن در نظر گرفته می‌شود، انتظار داریم که این تأثیر شوک‌های مثبت و منفی که به بازده‌های یک سهام وارد می‌شود، متفاوت از یکدیگر عمل کنند. شکل (۴-۱) نشان‌دهنده‌ی عملکرد نامتقارن نوسانات بازدهی سهام نسبت به اخبار خوب و بد است که بیان می‌کند نوسانات نسبت به اخبار بد کشیدگی بلندتری دارد و واکنش شدیدتری از خود نشان می‌دهند که گویای وجود اثر اهرم در نوسانات مبتنی بر واریانس ناهمسان شرطی خواهد بود [۱۷].



شکل ۴-۱: اثر اهرمی در نوسانات بازدهی سهام

مدل واریانس ناهمسان شرطی آستانه‌ای یکی از تعمیم‌هایی است که چنین عدم تقارنی در اثر اخبار و اطلاعات خوب و بد وارده بر بازارهای مالی را در نظر می‌گیرد.

^{۱۸} Glosten-Jagannathan-Runkle

در شکل عمومی این مدل به صورت زیر نمایش داده می شود

$$\sigma_t^2 = \alpha_0 + \sum_{k=1}^n \beta_k \sigma_{t-k}^2 + \sum_{h=1}^m \alpha_h u_{t-h}^2 + \sum_{h=1}^r r_h u_{t-h}^2 I_{t-h}$$

$$\begin{cases} I_{t-h} = 1 & ; & u_{t-h} < 0 \\ I_{t-h} = 0 & ; & u_{t-h} \geq 0 \end{cases}$$

که در آن $r \leq m$ می باشد و به ازای $h > r$ داریم ؛ $r_h = 0$.

در این جا $u_{t-h} \geq 0$ ، بیان گر اخبار خوب و $u_{t-h} < 0$ ، بیان گر اخبار بد است. مقدار r نیز نشان دهنده تعداد وقفه هایی است که طی آن اخبار گذشته (خوب یا بد) بر واریانس اثر می گذارد [۱۸].

۵-۸-۱ مدل واریانس ناهمسان شرطی نمایی^{۱۹} (EGARCH)

مدل واریانس ناهمسان شرطی نمایی به شکل متفاوتی از مدل آستانه ای به دنبال تعیین اثرهای نامتقارن شوک های مثبت و منفی بر واریانس شرطی است.

نلسون^{۲۰} (۱۹۹۱) شکل عمومی این مدل را به صورت زیر مطرح کرد

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{k=1}^n \beta_k \ln(\sigma_{t-k}^2) + \sum_{k=1}^n \alpha_n \left| \frac{u_{t-h}}{\sigma_{t-h}} \right| + \sum_{h=1}^r r_h \left| \frac{u_{t-h}}{\sigma_{t-h}} \right|$$

[۱۹].

مدل EGARCH مشکل مثبت بودن تمامی ضرایب معادله ی واریانس را نسبت به سایر تعمیم های مطرح شده به خاطر لگاریتمی بودن متغیر وابسته ی آن $\ln(\sigma_t^2)$ ، برطرف می کند و این مزیت را به وجود می آورد که ضرایب متغیرهای سمت راست نیز می توانند مثبت یا منفی برآورد گردند [۱۸].

^{۱۹} Exponential General Autoregressive Conditional Heteroscedastic

^{۲۰} Nelson

فصل دوّم

برآورد پارامترها و معیارهای ارزیابی مدل‌های پیشنهادی

در این بخش مفاهیم مربوط به برآورد پارامترهای یک الگوی سری‌زمانی و روش‌های این برآورد شرح داده می‌شود و به بیان معیارهایی برای ارزیابی مدل‌های برآورد شده پرداخته می‌شود.

۱-۲ برآورد پارامترهای الگوهای سری‌زمانی

بعد از شناسایی الگوی سری‌زمانی، نیازمند برآورد پارامترهای این الگو می‌باشیم. مدل‌های مختلفی برای برآورد و پیش‌بینی تلاطم وجود دارد، اما دو ویژگی اساسی سری‌های زمانی یعنی وجود دنباله‌های پهن توزیع داده‌ها و نوسانات خوشه‌ای باعث شده است که استفاده‌ی گسترده‌ای از مدل‌های خانواده‌ی $ARCH$ انجام گیرد. تاکنون برآورد پارامترهای الگوهای سری‌زمانی عمدتاً با استفاده از روش حداکثر راستنمایی بوده و توجه کمی به روش برآورد بیزی صورت گرفته است. در نمونه‌های کوچک روش حداکثر راستنمایی کارایی کمتری نسبت به روش بیزی دارد، اما با افزایش حجم نمونه، کارایی و دقت پیش‌بینی در هر دو روش همگرا می‌شود. در ادامه به بیان دقیق‌تر روش بیز و تکنیک‌های مربوط به برآورد پارامترهای مدل تحت مقایسه می‌پردازیم.

۱-۱-۲ استنتاج بیزی^{۲۱}

روش بیزی به‌عنوان یک جایگزین برای روش‌های کلاسیک آماری هم‌چون MLE ^{۲۲} سالهاست که مدنظر محققان آمار و اقتصاد قرار گرفته است. پایه و اساس برآورد پارامترهای یک مدل به روش بیزی قضیه‌ی بیز است.

۲-۱-۲ تابع زیان

برای مقادیری از پارامتر $\theta \in \Theta$ ، میزان نزدیکی مقدار برآوردگر $d \in D$ (کلاس کلیه‌ی برآوردگرها) را به مقدار واقعی پارامتر θ با تابع زیان L برپایه‌ی داده‌های تصادفی X به‌صورت زیر تعریف می‌کنیم

$$L: \Theta \times D \rightarrow [0, +\infty)$$

$$L(\theta, d) \geq 0 \quad \forall \theta, d$$

$$L(\theta, d) = 0 \quad d = \theta .$$

^{۲۱} Bayesian

^{۲۲} Maximum Likelihood Estimation

چند تابع زیان مشهور عبارتند از

$$\begin{aligned} ۱) L(\theta, d) &= (\theta - d)^2 && \text{تابع زیان درجه‌ی دوم یا مربع خطا} \\ ۲) L(\theta, d) &= |\theta - d| && \text{تابع زیان قدرمطلق} \\ ۳) L(\theta, d) &= b\{e^{c(\theta-d)} - c(\theta - d) - ۱\} && \text{تابع زیان LINEX} \end{aligned}$$

چون تابع زیان به دلیل وابستگی به $d(x)$ مقدار ثابتی ندارد و یک متغیر تصادفی است برای مقایسه‌ی دو برآوردگر، امیدریاضی تابع زیان این دو برآوردگر را باهم مقایسه می‌کنند که تابعی از θ است و به آن تابع مخاطره می‌گویند، پس

$$R(\theta, d) = E_{\theta} (L(\theta, d(x))) = \begin{cases} \int_x L(\theta, d(x)) f(x; \theta) dx & \text{برای } x \text{ پیوسته} \\ \sum_{x \in X} L(\theta, d(x)) f(x; \theta) & \text{برای } x \text{ گسسته} \end{cases}$$

در بین دو تصمیم d و d' ، تصمیم d به‌شدت تحت تصمیم d' قرار دارد، اگر که برای همه‌ی مقادیر θ داشته باشیم؛ $R(\theta, d) \leq R(\theta, d')$ و برای حداقل یک مقدار θ داشته باشیم؛ $R(\theta, d) < R(\theta, d')$ ، که با این شرایط بین دو تصمیم d و d' ، تصمیم d ترجیح داده می‌شود [۲۰].

۳-۱-۲ مخاطره و تصمیم بیزی

$\pi(\theta)$ را توزیع احتمال روی فضای پارامتری θ در نظر بگیرید که به آن توزیع پیشین می‌گویند. چون تابع مخاطره $R(\theta, d) = E_{\theta} [L(\theta, d(x))]$ تابعی از θ است و θ متغیر تصادفی است پس $R(\theta, d)$ نیز یک متغیر تصادفی است. امیدریاضی $R(\theta, d)$ نسبت به چگالی پیشین θ را مخاطره‌ی بیز می‌نامند و داریم

$$\begin{aligned} r(\pi, d) &= E[R(\theta, d)] = \int_{\theta \in \Theta} R(\theta, d) d\pi(\theta) \\ &= \begin{cases} \int_{\theta \in \Theta} R(\theta, d) \pi(\theta) d(\theta) & \text{اگر } \theta \text{ پیوسته باشد.} \\ \sum_{\theta \in \Theta} R(\theta, d) \pi(\theta) & \text{اگر } \theta \text{ گسسته باشد.} \end{cases} \end{aligned}$$

یک تابع تصمیم d را تصمیم بیزی نسبت به توزیع پیشین π می‌نامند اگر که مخاطره‌ی بیزی را در تمام توابع تصمیم، مینیمم کند بنابراین باید داشته باشیم

$$r(\pi, d) = \inf_{d' \in D} r(\pi, d') = m_{\pi}$$

در آمار بیزی x و θ هر دو به‌عنوان متغیر تصادفی با چگالی مشترک (یا تابع زیان احتمال) به‌شکل

$$f(x; \theta) = \pi(\theta) \cdot f(x|\theta)$$

نشان داده می‌شوند که در آن چگالی پیشین $\pi(\theta)$ و $f(x|\theta)$ توزیع x به شرط θ است. با در نظر گرفتن نمونه‌ی تصادفی x ، در توزیع پیشین $\pi(\theta)$ تجدیدنظر می‌کنیم و براساس نظریه‌ی بیز برای چگالی پسین θ خواهیم داشت

$$\pi(\theta|x) = \frac{f(x;\theta)}{m(x)}.$$

که در آن $m(x)$ چگالی کناری x است و مستقل از θ به صورت زیر تعریف می‌شود

$$m(x) = \int_{\theta} f(x;\theta') d\pi(\theta') = \int_{\theta} \pi(\theta') f(x;\theta') d\theta'.$$

با این شرایط تابع چگالی پسین θ برابر است با

$$\begin{aligned} \pi(\theta|x) &= \frac{\pi(\theta) \cdot f(x|\theta)}{m(x)} = C(x) \cdot \pi(\theta) \cdot f(x|\theta) \propto \pi(\theta) \cdot f(x|\theta) \\ &= \pi(\theta) \cdot L(\theta). \end{aligned}$$

در این رابطه $L(\theta)$ ، تابع درستنمایی است [۲۱]. براساس تعریف بالا برآوردگر بیز پارامتر θ به شکل زیر به دست می‌آید

$$\langle \theta \rangle = d(x) = E(\theta|x) = \frac{1}{Z} \int_{\theta} \theta \cdot \pi(\theta|x) d\theta.$$

که Z یک عبارت ثابت نرمال‌ساز برای معادله است

$$Z = \int_{\theta} \pi(\theta|x) d\theta.$$

وجود ماهیت بازگشتی معادله‌ی واریانس در الگوهای خانواده‌ی $ARCH$ و پیچیدگی‌های به وجود آمده در مدل‌های تحت پوشش استنتاج آماری، در بسیاری از موارد کاربردی به دست آوردن چگالی پسین به منظور برآورد پارامترهای معادله را به کمک روش‌های معمول میسر نمی‌سازد. در ادامه به بیان تعاریف مربوط به الگوریتم قدم‌زدن تصادفی متروپلیس-هستینگز^{۲۳} (MH) از روش‌های زنجیر مارکف مونت کارلو که در شبیه سازی عددی برآورد پارامتر بیزی کاربرد دارد، می‌پردازیم.

۲-۲ زنجیر مارکف

زنجیر مارکف معرفی شده توسط آندری مارکف^{۲۴} یک فرایند تصادفی گسسته در زمان با خاصیت تصادفی مارکف می‌باشد، که در آن احتمال شرطی رخداد آینده فقط به آخرین رخداد موجود، یعنی رخداد کنونی

^{۲۳} Metropolis-Hastings

^{۲۴} Andrey Markov

وابسته است. یک فرایند تصادفی گسسته در زمان شامل سیستمی است که در هر مرحله در حالت مشخصی قرار دارد و به صورت تصادفی در هر مرحله تغییر حالت می‌دهد.

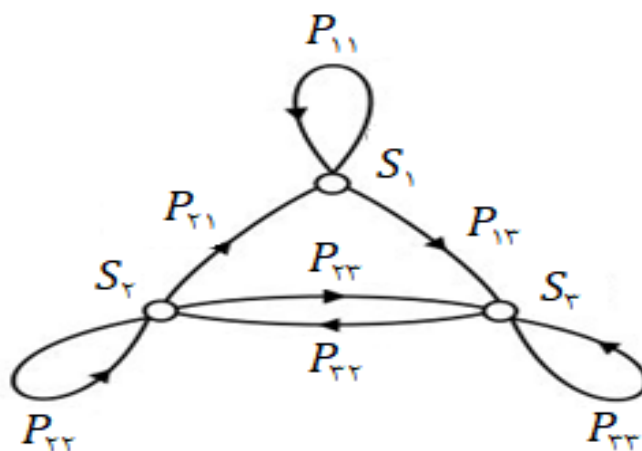
دنباله‌ای از متغیرهای تصادفی X_1, X_2, \dots را در نظر بگیرید که دارای خاصیت مارکف هستند به این معنا که احتمال حرکت به مرحله‌ی بعدی تنها به مرحله‌ی کنونی وابسته است، بنابراین خواهیم داشت

$$p(X_{n+1} = x | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = p(X_{n+1} = x | X_n = x_n).$$

زنجیر مارکف اغلب توسط دنباله‌ای از گراف‌های جهت‌دار که در آن یال‌های گراف توسط احتمال رفتن از یک حالت در زمان n به حالت‌های دیگر در زمان $n + 1$ به شکل زیر برچسب گذاری شده است، نمایش داده می‌شود

$$P_{ij} = P(X_{n+1} = j | X_n = i).$$

شکل (۱-۲) گراف یک زنجیر مارکف با فضای حالت S_1, S_2, S_3 را نشان می‌دهد، و یال‌های گراف و مسیر رفتن از یک حالت به حالت بعدی را شامل می‌شود.



شکل ۱-۲: نمونه‌ای از زنجیر مارکف با سه حالت

۱-۲-۲ ماتریس احتمال انتقال زنجیر مارکف

تغییر حالت سیستم از یک مرحله به مرحله‌ی بعدی انتقال نام دارد و احتمال‌هایی که به این تغییر حالت‌ها نسبت داده می‌شود، احتمال انتقال نام دارد. یک فرایند مارکف با یک فضای حالت شمارای S را با یک ماتریس انتقال برای توصیف احتمال‌های هر انتقال مشخص می‌کنند. براین اساس خواهیم داشت

$$(P_t)_{ij} = P(X_{t+1} = j | X_t = i) \quad \forall i, j \in S.$$

لم ۱-۲-۲. در حالت کلی برای اعضای ماتریس احتمال انتقال $P_t \cdot P_{t+1} \dots P_{t+k}$ خواهیم داشت

$$P_{ij} = P(X_{t+k+1} = j | X_t = i) \quad i, j \in S.$$

که در آن S فضای حالت شمارای زنجیر مارکف می‌باشد.

برهان. قرار می‌دهیم $M = P_t \cdot P_{t+1}$ ، با استفاده از ضرب ماتریس‌ها داریم

$$\begin{aligned} M_{ij} &= \sum_{k=1}^n (P_t)_{i.k} (P_{t+1})_{k.j} \\ &= \sum_{k=1}^n P(X_{t+1} = k | X_t = i) P(X_{t+2} = j | X_{t+1} = k) \\ &= P(X_{t+2} = j | X_t = i). \end{aligned}$$

(برای بررسی دقیق‌تر به رفرنس مراجعه شود) [۲۲].

با تعمیم $M = P_t \cdot P_{t+1} \dots P_{t+k}$ خواهیم داشت

$$P_{ij} = P(X_{t+k+1} = j | X_t = i).$$

نکته ۲-۲-۲. باتوجه به تعریف احتمال‌های انتقال بالا ماتریس احتمال انتقال k -مرحله‌ای زنجیر مارکف که

با $P_t^{(k)}$ نشان داده می‌شود به صورت زیر است

$$P_t^{(k)} = P_t \cdot P_{t+1} \dots P_{t+k-1} = \begin{pmatrix} P(X_{t+k} = 1 | X_t = 1) & P(X_{t+k} = 2 | X_t = 1) & \dots & P(X_{t+k} = n | X_t = 1) \\ P(X_{t+k} = 1 | X_t = 2) & P(X_{t+k} = 2 | X_t = 2) & \dots & P(X_{t+k} = n | X_t = 2) \\ \vdots & \vdots & \dots & \vdots \\ P(X_{t+k} = 1 | X_t = n) & P(X_{t+k} = 2 | X_t = n) & \dots & P(X_{t+k} = n | X_t = n) \end{pmatrix}.$$

یک تابع احتمال انتقال برگشت‌پذیر^{۲۵} نسبت به یک توزیع اولیه می‌باشد، اگر برای یک زنجیر مارکف

X_1, X_2, \dots ، توزیع جفت‌های (X_i, X_{i+1}) قابل تعویض باشد.

یک زنجیر مارکف برگشت‌پذیر است اگر احتمال انتقال باتوجه به توزیع اولیه خود برگشت‌پذیر باشد.

برگشت‌پذیری دلالت بر مانایی دارد اما عکس آن درست نیست [۲۳].

۲-۲-۲ روش زنجیر مارکف مونت کارلو

اگر X_1, X_2, \dots یک فرایند تصادفی و زنجیر مارکف مانا و برگشت پذیر باشد و g یک تابع حقیقی مقدار روی آن باشد، برای محاسبه‌ی تابع $\mu = E[g(X)]$ که در آن μ از روش‌های مستقیم قابل محاسبه نیست، با فرض این که بتوانیم نمونه‌ی مستقل و هم توزیع X_1, X_2, \dots را از توزیع پایه‌ی X به دست بیاوریم، تقریب مونت کارلوی تابع g بر پایه‌ی زنجیر مارکف برگشت پذیر X_1, X_2, \dots به صورت $\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n g(X_i)$ است. اگر قراردادیم؛ $Y_i = g(X_i)$ ، در این صورت $\{Y_i\}$ زنجیری مانا و برگشت پذیر با میانگین μ و واریانس $\sigma^2 = Var(g(X))$ خواهد بود. $\hat{\mu}_n$ میانگین نمونه‌ای برای Y_i است و براساس قضیه‌ی حد مرکزی داریم

$$\bar{X} = \hat{\mu}_n \approx N\left(\mu, \frac{\sigma^2}{n}\right).$$

برآورد خطای استاندارد روش مونت کارلو^{۲۶} ($MCSE$) به صورت $\frac{\hat{\sigma}_n}{\sqrt{n}}$ خواهد بود که واریانس تجربی Y_i به شکل زیر تعریف می شود

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\mu}_n)^2.$$

در نظریه‌ی مونت کارلو فرض می شود که استقلال تصادفی در زنجیر مارکف خطای استاندارد را تغییر می دهد که در آن برای واریانس Y_i خواهیم داشت

$$\sigma^2 = Var(g(X_i)) + 2 \sum_{k=1}^{\infty} Cov(g(X_i), g(X_{i+k}))$$

[۲۲], [۲۳].

۳-۲-۲ الگوریتم متروپلیس-هستینگز^{۲۷} (MH)

در سال ۱۹۷۰ هستینگز تعمیم مدل متروپلیس و یکی از روش‌های زنجیر مارکف مونت کارلو را ارائه نمود. در محاسبه‌ی امید ریاضی $E_f[h(X)] = \int h(x)f(x)dx$ ، در صورتی که تابع f در شکل بسته موجود نباشد و انتگرال امید ریاضی ناسره باشد می توان از روش‌های نمونه گیری زنجیر مارکف مونت کارلو $(MCMC)$ ^{۲۸} استفاده کرد. ایده‌ی اصلی یافتن یک زنجیر مارکف پیرامون مقادیر $X \sim f(x)$ است،

Monte Carlo Standard Error ^{۲۶}

Metropolis-Hastings ^{۲۷}

Markov Chain Monte Carlo ^{۲۸}

به شکلی که توزیع مانای زنجیر مارکف مورد نظر، $f(x)$ باشد. دنباله‌ی (x_0, x_1, x_2, \dots) به دست آمده از زنجیر مارکف پیرامون X با افزایش حجم نمونه $(n \rightarrow \infty)$ با توزیع f همگام خواهد شد $(X_n \sim f(x))$ [۲۴].

۴-۲-۲ مراحل الگوریتم متروپلیس-هستینگز

با در نظر گرفتن x_n به عنوان حالت کنونی زنجیر مارکف پیرامون X ، بر اساس یک توزیع پیشنهادی $Q(\cdot | x_n)$ مقدار x^* را از توزیع پیشنهادی برای شروع فرایند تولید زنجیر مارکف تعیین می‌کنیم. احتمال پذیرش و یا رد x^* برای رفتن به مرحله‌ی بعدی بر اساس توزیع Q و حالت کنونی x_n به صورت زیر تعریف می‌شود

$$A(x_n \rightarrow x^*) = \min\left(1, \frac{f(x^*)}{f(x_n)} \times \frac{Q(x_n | x^*)}{Q(x^* | x_n)}\right).$$

سپس در مرحله‌ی بعدی الگوریتم MH برای به دست آوردن نمونه‌ی بعدی و تعیین x_{n+1} ، یک تعداد عدد تصادفی از توزیع یکنواخت بین صفر و یک تولید می‌کنیم $(U \sim U(0, 1))$. بر اساس احتمال پذیرش الگوریتم، نمونه‌ی تصادفی در مرحله‌ی بعدی زنجیر به شکل زیر است

$$x_{n+1} = \begin{cases} x^* & U \leq A(x_n \rightarrow x^*) \\ x_n & U > A(x_n \rightarrow x^*) \end{cases}.$$

به این ترتیب الگوریتم MH شبیه‌سازی‌های x^* را با نسبت $\frac{f(x^*)}{Q(x^* | x_n)} > \frac{f(x_n)}{Q(x_n | x^*)}$ می‌پذیرد زیرا احتمال پذیرش x^* همواره برابر یک خواهد شد. همچنین با توجه به احتمال پذیرش $A(x_n \rightarrow x^*)$ اگر f یک توزیع مانا باشد و داشته باشیم

$$f(x_n) \times Q(x^* | x_n) = f(x^*) \times Q(x_n | x^*)$$

احتمال پذیرش نمونه‌ی x_{n+1} برای هر n برابر یک خواهد شد [۲۵].

نکته ۳-۲-۲. در صورتی که توزیع پیشنهادی Q متقارن انتخاب شود، $Q(x_n | x^*) = Q(x^* | x_n)$ ، یعنی احتمال تولید x^* زمانی که حالت کنونی زنجیر x_n می‌باشد با حالت عکس آن برابر باشد، احتمال پذیرش و یا رد مرحله‌ی بعدی به صورت زیر تعریف می‌شود

$$A_n(x_n \rightarrow x^*) = \min\left(1, \frac{f(x^*)}{f(x_n)}\right).$$

با این شرایط بعد از تولید نمونه‌ی تصادفی یکنواخت بین صفر و یک برای مرحله‌ی بعدی الگوریتم با استفاده از احتمال بالا خواهیم داشت

$$x_{n+1} = \begin{cases} x^* & U \leq A(x_n \rightarrow x^*) \\ x_n & U > A(x_n \rightarrow x^*) \end{cases} .$$

این الگوریتم که توزیع‌های مقارن را برای نمونه‌برداری تصادفی از توزیع f انتخاب می‌کند الگوریتم متروپلیس نام دارد [۲۶].

۳-۲ مفاهیم مربوط به ارزیابی مدل‌های سری‌زمانی

در این بخش به مفاهیم مربوط به ارزیابی مدل‌های سری‌زمانی برآزش داده شده می‌پردازیم و معیارهای اطلاعات مختلف را برای انجام آزمون نیکویی برآزش بر روی هر مدل معرفی می‌کنیم.

۱-۳-۲ عامل بیز

مقایسه‌ی مدل بیزی روشی برای انتخاب مدل بهینه بر پایه‌ی عامل بیز است که یک احتمال بیزی را به‌جای آزمون فرض آماری برای بررسی فرضیه‌ها درباره‌ی پارامترهای توزیع احتمال شامل می‌شود. با در نظر گرفتن مدل‌های M_1, M_2, \dots, M_q تحت اطلاعات X با چگالی $f(x|\theta_i)$ برای مدل M_i که در آن بردار پارامتر θ_i نامعلوم و دارای بُعد k_i است، توزیع پسین هر مدل M_i بر اساس احتمالات پیشین P_i برای هر مدل به‌شکل زیر خواهد بود

$$P(M_i|x) = \left(\sum_{j=1}^q \frac{P_j}{P_i} \cdot B_{ji} \right)^{-1} .$$

که در آن B_{ji} ، عامل بیز مدل M_j به M_i است و داریم

$$B_{ji} = \frac{m_j(x)}{m_i(x)} = \frac{\int f_j(x|\theta_j)\pi_j(\theta_j)d\theta_j}{\int f_i(x|\theta_i)\pi_i(\theta_i)d\theta_i} .$$

در این رابطه $m_j(x)$ ، چگالی حاشیه‌ای x تحت مدل M_j می‌باشد [۲۷].

برای محاسبه‌ی B_{ji} ها اغلب از توزیع‌های پیشین غیرمستقیم (پیش فرض)، $\pi_i^N(\theta_i)$ ، استفاده می‌شود که

معمولاً به‌شکل‌های توزیع پیشین یکنواخت، $\pi_i^U(\theta_i) = 1$ ، توزیع پیشین جفری،

$\pi_i^J(\theta_i) = (\det(I_i(\theta_i)))^{-\frac{1}{2}}$ ، که در آن $I_i(\theta_i)$ ماتریس اطلاعات موردانتظار فیشر مدل M_i است و

توزیع پیشین استنتاجی، $\pi_i^R(\theta_i)$ ، خواهد بود. با استفاده از این پیش فرض‌ها عامل بیز دوباره تعریف خواهد شد

$$B_{ji}^N = \frac{m_j^N(x)}{m_i^N(x)} = \frac{\int f_j(x|\theta_j)\pi_j^N(\theta_j)d\theta_j}{\int f_i(x|\theta_i)\pi_i^N(\theta_i)d\theta_i}.$$

نکته ۲-۳-۱. پیش فرض‌های π_i^N معمولاً ناسره هستند و به شکل ثابت‌های دلخواه C_i تعریف می‌شوند. از این رو B_{ji}^N به صورت $\frac{C_j}{C_i}$ دلخواه تبدیل خواهد شد. به منظور حل این مشکل با در نظر گرفتن $x(I)$ به عنوان بخشی از داده‌های مورد بررسی، $\pi_i^N(\theta_i)$ ‌های تبدیل شده به توزیع‌های پیشین سره به شکل

$$\pi_i^N(\theta_i|x(I)) = \frac{f_i(x(I)|\theta_i)\pi_i^N(\theta_i)}{m_i^N(x(I))}$$

که در آن

$$m_i^N(x(I)) = \int f_i(x(I)|\theta_i)\pi_i^N(\theta_i)d\theta_i$$

تعریف شده است. سپس عامل‌های بیز با توجه به باقی داده‌ها و با استفاده از توزیع‌های پیشین

$\pi_i^N(\theta_i|x(I))$ ، به صورت زیر به دست می‌آید

$$B_{ji}(x(I)) = B_{ji}^N \times B_{ij}^N(x(I))$$

و

$$B_{ij}^N(x(I)) = m_i^N(x(I))/m_j^N(x(I))$$

[۲۸].

۲-۳-۲ اطلاع فیشر^{۲۹}

بر اساس نظریه‌ی اطلاع فیشر می‌توان مقدار اطلاعاتی را که یک نمونه‌ی تصادفی برای شناخت از پارامتر نامعلوم جامعه (θ) در خود دارد، اندازه‌گیری کرد.

تعریف ۲-۳-۲. $f(X; \theta)$ را تابع چگالی احتمال یا تابع توزیع احتمال برای متغیر تصادفی X در نظر

بگیرید. یک تابع امتیاز را به شکل $\frac{\partial}{\partial \theta} \text{Log} f(X; \theta)$ تعریف می‌کنیم که حساسیت تابع درست‌نمایی را

نسبت به پارامتر θ تعیین می‌کند. در این صورت واریانس تابع امتیاز برابر با اطلاع فیشر خواهد بود و داریم

$$\text{Var}\left(\frac{\partial}{\partial\theta}\text{Log}f(X;\theta)\right) = E\left(\frac{\partial}{\partial\theta}\text{Log}f(X;\theta)\right)^2 - E^2\left(\frac{\partial}{\partial\theta}\text{Log}f(X;\theta)\right).$$

برای ساده کردن و انجام محاسبات مربوط به اطلاع فیشر شرایط نظم را بیان می‌کنیم. براین اساس شرایط نظم برابر است با

(۱) پارامتر مجهول θ یک فاصله‌ی باز از اعداد حقیقی است. ($\theta \in \Theta \subset \mathbb{R}$)

(۲) مشتق تابع چگالی احتمال متغیر تصادفی X وجود دارد.

(۳) جابه‌جایی بین عملگرهای مشتق و انتگرال امکان‌پذیر است.

(۴) رابطه‌ی زیر برای هر نقطه از پارامتر برقرار است. به این معنی که گشتاور دوم مشتق تابع لگاریتم درست‌نمایی موجود بوده، مخالف صفر است.

$$E\left[\left(\frac{\partial}{\partial\theta}\text{Log}f(X;\theta)\right)^2 \middle| \theta\right] > 0.$$

قضیه ۲-۳-۳. تحت شرایط نظم، اطلاع فیشر در حالت تک متغیره برابر است با

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\text{Log}f(X;\theta) \middle| \theta\right].$$

برهان. می‌توان نشان داد امیدریاضی مشتق لگاریتم تابع درست‌نمایی نسبت به پارامتر θ براساس نمونه‌ی تصادفی X صفر است، زیرا:

$$E\left[\frac{\partial}{\partial\theta}\text{Log}f(X;\theta) \middle| \theta\right] = \int \frac{\frac{\partial}{\partial\theta}f(x;\theta)}{f(x;\theta)} f(x;\theta) dx = \frac{\partial}{\partial\theta} \int f(x;\theta) dx = \frac{\partial}{\partial\theta} 1 = 0.$$

در نتیجه برای محاسبه‌ی اطلاع فیشر داریم

$$I(\theta) = E\left[\left(\frac{\partial}{\partial\theta}\text{Log}f(X;\theta)\right)^2 \middle| \theta\right] = \int \left(\frac{\partial}{\partial\theta}\text{Log}f(x;\theta)\right)^2 f(x;\theta) dx.$$

تحت شرایط نظم و با فرض وجود مشتق دوم برای تابع درست‌نمایی می‌توان شکل ساده‌تری برای محاسبه‌ی اطلاع فیشر نیز ایجاد کرد [۲۹].

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\text{Log}f(X;\theta) \middle| \theta\right].$$

تعریف ۲-۳-۴. (ماتریس اطلاع فیشر^{۳۰}) اگر توزیع آماری متغیر تصادفی X دارای N پارامتر باشد آنگاه بردار پارامتر به صورت $\theta = (\theta_1, \theta_2, \dots, \theta_N)^T$ نوشته شده و اطلاع فیشر برای این توزیع برحسب پارامتر، یک ماتریس $N \times N$ خواهد بود.

در این حالت ماتریس اطلاع فیشر دارای عناصری به صورت زیر است

$$[I(\theta)]_{i,j} = E \left[\left(\frac{\partial}{\partial \theta_i} \text{Log} f(X; \theta) \right) \left(\frac{\partial}{\partial \theta_j} \text{Log} f(X; \theta) \right) \middle| \theta \right].$$

تحت شرایط خاص^{۳۱} نظم، می‌توان عناصر ماتریس فیشر را به شکل ساده‌تری تعریف کرد تا محاسبات راحت‌تر صورت گیرد. رابطه‌ی زیر جملات یا عناصر ماتریس فیشر را براساس مشتق دوم نمایش می‌دهد

$$[I(\theta)]_{i,j} = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \text{Log} f(X; \theta) \middle| \theta \right].$$

مشخص است که این رابطه درست به مانند رابطه‌ی مربوط به متغیر تک بعدی نوشته شده است [۳۰].

۳-۳-۲ معیار اطلاعات آکائیک^{۳۱} (AIC)

یک اندازه از نیکویی برازش یک مدل برای برآورد کیفیت نسبی آن براساس مجموعه‌ای از داده‌ها برپایه‌ی نظریه‌ی اطلاعات را آکائیک (۱۹۷۱) ارائه داد. مدل عمومی $f(\cdot | \theta)$ را که با محدود کردن بردار پارامتر θ همه‌ی K مدل رقابتی از آن به دست می‌آید، در نظر بگیرید

$$M(k) | f(\cdot | \theta) ; \theta = (\theta_1, \theta_2, \dots, \theta_k, \theta_{k+1}, \dots, \theta_K).$$

بر این اساس مقادیر واقعی بردار پارامتر θ^* را تعریف می‌کنیم، داریم؛ $\theta^* \in \mathbb{R}^K$. اگر

$$\{M(k) : k = 1, 2, \dots, K\}$$

مجموعه‌ای از مدل‌های رقابتی با اندیس $k = 1, 2, \dots, K$ به عنوان تعداد پارامترهای آزاد هر مدل $M(k)$ یا رتبه‌ی مدل باشد، معیار اطلاعات آکائیک برای انتخاب یک مدل بهینه‌ی M_k میان مجموعه‌ی مدل‌ها به شکل زیر خواهد بود

$$AIC(k) = -2 \log L(\hat{\theta}_k) + 2k.$$

$L(\hat{\theta}_k)$ نشان‌دهنده‌ی بیشترین مقدار تابع درست‌نمایی است [۳۱].

^{۳۰} Fisher Information Matrix

^{۳۱} Akaike Information Criterion

نکته ۲-۳-۵. معیار اطلاعات آکائیک یک تابع افزایشی از تعداد پارامترهای برآورد شده است، در نتیجه زمانی که اندازه‌ی نمونه کوچک باشد این احتمال وجود دارد که مدل با پارامترهای بیشتر را ترجیح دهد. بر این اساس معیار اطلاعات آکائیک تصحیح شده (AIC_c) برای اندازه‌ی نمونه‌ی کوچک تعریف شده است که به مدل بستگی دارد و با فرض این که مدل یکنواخت است و پارامترهای آن خطی هستند و باقیمانده‌ها دارای توزیع نرمال می‌باشند، برابر است با

$$AIC_c = AIC + \frac{2k^2 + 2k}{n - k - 1}.$$

با افزایش حجم نمونه ($n \rightarrow \infty$) ملاحظه می‌شود که عبارت دوم به صفر میل می‌کند و AIC_c به AIC همگرا خواهد شد [۳۲].

۲-۳-۴ کمیّت اطلاعات کولبک-لیبلر^{۳۲}

معیاری برای اندازه‌گیری واگرایی یک توزیع احتمال از توزیع واقعی خود که توسط کمیّت اطلاعات کولبک-لیبلر ($K - L$) یا آنتروپی منفی اندازه‌گیری می‌شود. نیکویی برازش $f(x|\theta)$ که در آن x بردار تصادفی پیوسته و θ بردار پارامتر به شکل $\{\theta \in \mathbb{R}^K; \theta = (\theta_1, \theta_2, \dots, \theta_K)\}$ است را نسبت به بردار پارامتر واقعی θ^* از θ با چگالی احتمال $f(x|\theta^*)$ به وسیله کمیّت اطلاعات $K - L$ تعریف می‌کنند

$$\begin{aligned} I(\theta^*; \theta) &= E[-\text{Log } f(x|\theta) + \text{Log } f(x|\theta^*)] \\ &= - \int f(x|\theta^*) \text{Log } f(x|\theta) dx + \int f(x|\theta^*) \text{Log } f(x|\theta^*) dx \\ &= -H(\theta^*; \theta) + H(\theta^*; \theta^*). \end{aligned}$$

در این تعریف $H(\theta^*; \theta) = \int f(x|\theta^*) \text{Log } f(x|\theta) dx$ آنتروپی متقاطع نام دارد که نیکویی برازش از $f(x|\theta)$ به $f(x|\theta^*)$ را مشخص می‌کند و $H(\theta^*; \theta^*) \equiv H(\theta^*)$ آنتروپی منفی شانون^{۳۳} است که برای $f(x|\theta^*)$ ثابت خواهد بود.

اگر P و Q احتمال‌های تعریف شده بر روی مجموعه‌ی X باشند و P نسبت به Q به طور مطلق پیوسته باشد، کمیّت اطلاعات کولبک-لیبلر از Q به P به صورت زیر تعریف می‌شود

$$I(P; Q) = \int_x \text{Log} \left(\frac{dP}{dQ} \right) dP.$$

^{۳۲} Kullback-Leibler

^{۳۳} Shannon

$\frac{dP}{dQ}$ مشتق رادون-نیکودیم^{۳۴} P نسبت به Q و $I(P; Q)$ آنترپی نسبی P نسبت به Q است. اگر μ هر اندازه روی X تعریف شود که در آن $p = \frac{dP}{d\mu}$ و $q = \frac{dQ}{d\mu}$ و p و q نسبت به μ به طور مطلق پیوسته باشند) در این صورت کمیت $K - L$ از Q به P برابر است با

$$I(P; Q) = \int_x P \text{Log} \left(\frac{p}{q} \right) d\mu .$$

۵-۳-۲ برخی خواص کمیت کولبک-لیبلر

(۱) برای همه‌ی مقادیر پارامتر برآورد شده که در آن $f(x|\theta^*) \neq f(x|\theta)$ ، کمیت $K - L$ نامنفی است، یعنی $I(\theta^*; \theta) \geq 0$ و در محدوده‌ی احتمال x زمانی که مدل اساساً درست است، $I(\theta^*; \theta) = 0$ اگر و تنها اگر

$$f(x|\theta^*) = f(x|\theta) \quad ; \quad a. e .$$

(۲) اگر X_1, X_2, \dots, X_n متغیرهای تصادفی مستقل و هم‌توزیع باشند، براساس خاصیت جمع‌پذیری خواهیم داشت

$$I_n(\theta^*; \theta) = I_1(\theta^*; \theta) + I_2(\theta^*; \theta) + \dots + I_n(\theta^*; \theta) = nI(\theta^*; \theta) .$$

(۳) کمیت اطلاعات $K - L$ در جهت‌های تابع‌زبان احتمال (p, q) محدب است به طوری که اگر (p_1, q_1) و (p_2, q_2) دو جفت از توابع‌زبان احتمال باشند، برای $0 \leq \lambda \leq 1$ داریم

$$I(\lambda p_1 + (1 - \lambda)p_2; \lambda q_1 + (1 - \lambda)q_2) \leq \lambda I(p_1; q_1) + (1 - \lambda)I(p_2; q_2) .$$

[۳۳].

۶-۳-۲ معیار انحراف اطلاعات^{۳۵} (DIC)

معیار انحراف اطلاعات یک تعمیم بیزی بر معیار اطلاعات آکائیک است که آزمون نیکویی برازش پارامترهای به دست آمده از طریق استنتاج بیزی را انجام می‌دهد. در توزیع‌های پسین به دست آمده از شبیه‌سازی‌های $MCMC$ یک معیار انحراف به شکل زیر خواهد بود

$$\begin{aligned} D(\theta) &= 2 \text{Log}(h(y)) - 2 \text{Log} f(y|\theta) \\ &= 2 [\text{Log}(h(y)) - \text{Log} f(y|\theta)] = 2 \text{Log} \left[\frac{h(y)}{f(y|\theta)} \right] . \end{aligned}$$

^{۳۴} Radon-Nikodym

^{۳۵} Deviance Information Criterion

$f(y|\theta)$ تابع درستنمایی پارامتر θ براساس مشاهدات Y است که در آن $Y = (y_1, y_2, \dots, y_n)$ به یک بردار پارامتر p بُعدی θ وابسته است و $h(y)$ یک تابع استاندارد شده‌ی مشخص از داده‌ها است (معمولاً $h(y) = 1$ در نظر گرفته می‌شود). براساس انحراف کلاسیک $D(\theta)$ ، معیار DIC مدل که شامل دو بخش، یکی نیکویی برازش و یک عبارت اضافی برای پوشش پارامترهای مدل است، به شکل

$$DIC = \overline{D(\theta)} + P_D$$

خواهد بود که در آن $\overline{D(\theta)}$ انحراف میانگین پسین می‌باشد

$$\overline{D(\theta)} = E_{\theta|y}[D(\theta)] = E_{\theta|y}[-2 \text{Log}(f(y|\theta))] + 2 \text{Log}(h(y)).$$

P_D یک اندازه از تعداد مؤثر پارامترها در مدل است که به‌عنوان اختلاف بین انحراف میانگین پسین و انحراف ارزیابی شده در میانگین پسین $\tilde{\theta}$ تعریف می‌شود که برآوردگر پارامتر θ است و داریم

$$P_D = \overline{D(\theta)} - D(\tilde{\theta}).$$

P_D به‌طور کامل از انتخاب h مستقل است و $\bar{\theta} = E_{\theta|y}[\theta]$ می‌تواند یک برآوردگر طبیعی برای $\tilde{\theta}$ در نظر گرفته شود

$$\begin{aligned} P_D &= E_{\theta|y}[D(\theta)] - D[E_{\theta|y}(\theta)] \\ &= E_{\theta|y}[-2 \text{Log}(f(y|\theta))] + 2 \text{Log} h(y) - D[E_{\theta|y}(\theta)] \\ &= E_{\theta|y}[-2 \text{Log}(f(y|\theta))] + 2 \text{Log} f(y|\bar{\theta}). \end{aligned}$$

عبارت $-2 \text{Log}(f(y|\theta))$ نشان‌دهنده‌ی اطلاعات باقیمانده در Y به شرط θ است که به‌عنوان یک اندازه از عدم قطعیت شناخته می‌شود.

از این روابط مشخص است که

$$\overline{D(\theta)} = D(\tilde{\theta}) + P_D.$$

بنابراین با تنظیم مجدد معیار DIC خواهیم داشت

$$\begin{aligned} DIC &= \overline{D(\theta)} + P_D = D(\tilde{\theta}) + 2P_D = 2\overline{D(\theta)} - D(\tilde{\theta}) \\ &= -4E_{\theta|y}[\text{log}(f(y|\theta))] + 2 \text{log} f(y|\tilde{\theta}) \end{aligned}$$

[۳۴]، [۳۵].

فصل سوم

مدل گارچ گویا و آزمونی تجربی برای بازدهی سهام

۱-۳ مقدمه

در فصل قبل الگوریتم متروپلیس-هستینگز از روش‌های نمونه‌گیری زنجیر مارکف مونت کارلو برای نمونه برداری از یک توزیع مشخص، به منظور برآورد پارامترهای یک مدل مطرح شد. در این فصل تقریب جدیدی از یک تابع با عنوان تقریب پاده^{۳۶} ($Päde$) یا تقریب گویا تعریف خواهد شد و مدل جدیدی از گارچ نامتقارن با عنوان گارچ گویا که خطای کمتر و عملکرد بهتری نسبت به سایر مدل‌های خانواده‌ی $ARCH$ از خود نشان می‌دهد، بیان می‌کنیم.

۲-۳ تقریب گویا

اگر تابع $f(x)$ به صورت یک بسط مک‌لورن مطابق رابطه‌ی زیر در اختیار باشد، بطوری‌که تابع f و مشتق‌های آن در نقطه‌ی $x = 0$ پیوسته باشد

$$f(x) = \sum_{i=0}^{\infty} C_i x^i$$

یک تقریب پاده‌ی $\left[\frac{L}{M}\right]$ نسبت به بسط مک‌لورن $f(x)$ به شکل یک تابع گویا از دو چندجمله‌ای $P_L(x)$ و $Q_M(x)$ از درجه‌های L و M تعریف می‌کنیم

$$\left[\frac{L}{M}\right] = \frac{P_L(x)}{Q_M(x)} = \frac{a_0 + a_1 x + \dots + a_L x^L}{b_0 + b_1 x + \dots + b_M x^M} \quad ; \quad x \in D_f .$$

در این تقریب معمولاً فرض می‌شود که $b_0 = 1$ ، از این رو تعداد $L + 1$ ضریب صورت و M ضریب منخرج مجهول خواهد بود. سری توانی $f(x)$ ، $L + M + 1$ ضریب مجهول معادله را با استفاده از رابطه‌ی زیر مشخص می‌کند

$$\sum_{i=0}^{\infty} C_i x^i = \frac{a_0 + a_1 x + \dots + a_L x^L}{b_0 + b_1 x + \dots + b_M x^M} + O(x^{L+M+1})$$

یا

$$(b_0 + b_1 x + \dots + b_M x^M)(c_0 + c_1 x + \dots) = a_0 + a_1 x + \dots + a_L x^L + O(x^{L+M+1}).$$

(۱-۳)

با هم‌ارزسازی ضرایب $x^{L+1}, x^{L+2}, \dots, x^{L+M}$ برابر صفر، M معادله‌ی خطی برای ضرایب مجهول مخرج پیدا می‌کنیم

$$\begin{aligned} b_M c_{L-M+1} + b_{M-1} c_{L-M+2} + \dots + b_0 c_{L+1} &= 0 \\ b_M c_{L-M+2} + b_{M-1} c_{L-M+3} + \dots + b_0 c_{L+2} &= 0 \\ &\vdots \\ b_M c_L + b_{M-1} c_{L-M+1} + \dots + b_0 c_{L+M} &= 0 \end{aligned}$$

از آنجا که $b_0 = 1$ ، روابط بالا را می‌توان به یک دستگاه معادلات از M معادله‌ی خطی تبدیل کرد

$$\begin{bmatrix} c_{L-M+1} & c_{L-M} & \dots & c_L \\ c_{L-M+2} & c_{L-M+1} & \dots & c_{L+1} \\ \vdots & \vdots & \vdots & \vdots \\ c_L & c_{L+1} & \dots & c_{L+M-1} \end{bmatrix} \begin{bmatrix} b_M \\ b_{M-1} \\ \vdots \\ b_1 \end{bmatrix} = - \begin{bmatrix} c_{L+1} \\ c_{L+2} \\ \vdots \\ c_{L+M} \end{bmatrix}$$

با حل کردن دستگاه معادلات خطی بالا ضرایب مجهول مخرج برآورد می‌شوند. ضرایب صورت a_0, a_1, \dots, a_L نیز به‌طور مستقیم از رابطه‌ی (۳-۱) به‌وسیله‌ی سیستم معادلات زیر به‌دست می‌آیند

$$\begin{aligned} a_0 &= c_0 \\ a_1 &= c_1 + c_0 \\ a_2 &= c_2 + b_1 c_1 + b_2 c_0 \\ &\vdots \\ a_L &= c_L + \sum_{i=0}^{\min\{L, M\}} b_i c_{L-i} \end{aligned}$$

این سیستم از معادلات، با عنوان معادلات پایه برای تابع $f(x)$ شناخته می‌شود [۳۶]، [۳۷].

۳-۳ تقریب گویا برای مدل‌های ایستای میانگین

یک فرایند تصادفی مرکب $ARMA(p, q)$ را می‌توان به فرایند میانگین متحرک از مرتبه‌ی بی‌نهایت $(MA(\infty))$ به‌شکل $X_t = \Psi(B)U_t$ ، و یا فرایند خودبرگشت از مرتبه‌ی بی‌نهایت $(AR(\infty))$ به‌شکل

$$\pi(B)X_t = U_t$$

تبدیل کرد که در آن

$$\begin{aligned} \pi(B) &= \theta^{-1}(B)\phi(B) & ; & & \Psi(B) &= \phi^{-1}(B)\theta(B) \\ \phi(B) &= 1 - \phi_1 B - \dots - \phi_p B^p & , & & \theta(B) &= 1 - \theta_1 B - \dots - \theta_q B^q \end{aligned}$$

با استفاده از تقریب پایه مرتبه‌ی p و q برای فرایند مرکب تعیین می‌شود.

تعریف ۱-۳-۳. تقریب پاده از سری توانی $\sum_{i=0}^{\infty} C_i x^i$ را به صورت زیر در نظر می‌گیریم

$$\left[\frac{p}{q} \right] = \frac{P^{[p/q]}(x)}{Q^{[p/q]}(x)}$$

شرط کافی برای وجود تقریب این است که $Q^{[p/q]}(0) \neq 0$ ، در این صورت

$$P^{[p/q]}(x) = \begin{vmatrix} C_{p-q+1} & C_{p-q+2} & \cdots & C_{p+1} \\ C_{p-q+2} & C_{p-q+3} & \cdots & C_{p+2} \\ \vdots & \vdots & & \vdots \\ C_p & C_{p+1} & \vdots & C_{p+q} \\ \sum_{i=0}^{p-q} C_i x^{q+i} & \sum_{i=0}^{p-q+1} C_i x^{q+i-1} & \cdots & \sum_{i=0}^p C_i x^i \end{vmatrix}$$

و

$$Q^{[p/q]}(x) = \begin{vmatrix} C_{p-q+1} & C_{p-q+2} & \cdots & C_p & C_{p+1} \\ C_{p-q+2} & C_{p-q+3} & \cdots & C_{p+1} & C_{p+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ C_p & C_{p+1} & \cdots & C_{p+q+1} & C_{p+q} \\ x^q & x^{q-1} & \cdots & x & 1 \end{vmatrix}$$

با استفاده از مقادیر $\{C(p/q) ; p, q = 0, 1, 2, \dots\}$ از طریق جدول C به شکل زیر رتبه‌ی مدل $ARMA$

مشخص می‌شود

	p					
	0	1	2	3	...	
q						
0	$C(0/0)$	$C(1/0)$	$C(2/0)$	$C(3/0)$		
1	$C(0/1)$	$C(1/1)$	$C(2/1)$	$C(3/1)$		
2	$C(0/2)$	$C(1/2)$	$C(2/2)$	$C(3/2)$		
\vdots	\vdots	\vdots	\vdots	\vdots		

جدول ۱-۳: ورودی‌های جدول C برای تعیین رتبه‌ی مدل $ARMA$

که در آن

$$C(p/q) = Q^{[p/q]}(\circ) = \begin{vmatrix} C_{p-q+1} & C_{p-q+2} & \cdots & C_p \\ C_{p-q+2} & C_{p-q+3} & \cdots & C_{p+1} \\ \vdots & \vdots & \vdots & \vdots \\ C_p & C_{p+1} & \cdots & C_{p+q+1} \end{vmatrix}. \quad (2-3)$$

در تعیین ورودی‌های جدول C چهار اصل زیر برقرار است

- (۱) سطر اول به صورت $p = 0, 1, 2, \dots$; $C(p/0) = 1$ تعریف می‌شود.
- (۲) سطر دوم به صورت $p = 0, 1, 2, \dots$; $C(p/1) = C_p$ تعریف می‌شود.
- (۳) مقادیر ستون اول به صورت $q = 0, 1, 2, \dots$; $C(0/q) = (-1)^{q(q-1)/2} C_0^q$ خواهد بود.
- (۴) باقی ورودی‌ها از رابطه‌ی زیر محاسبه می‌شود

$$C(p/q + 1) = \frac{C(p + 1/q)C(p - 1/q) - C(p/q)^2}{C(p/q - 1)} \quad (3-3)$$

[۳۸].

مثال ۳-۳-۲. جدول C برای مدل $ARMA(1,1)$

مدل $ARMA(1,1)$ را در نظر بگیرید.

$$Z_t = \phi_1 Z_{t-1} + X_t - \theta_1 X_{t-1}$$

$$(1 - \phi_1 B)Z_t = (1 - \theta_1 B)X_t.$$

این مدل را می‌توان به شکل یک فرایند خودبرگشت از مرتبه‌ی بی‌نهایت تبدیل کرد

$$X_t = \frac{(1 - \phi_1 B)}{(1 - \theta_1 B)} Z_t = \pi(B)Z_t$$

که در آن

$$\pi(B) = 1 + \pi_1 B + \pi_2 B^2 + \cdots + \pi_k B^k + \cdots.$$

مشخص است که

$$\pi_k = \theta_1^{k-1} (\theta_1 - \phi_1) \quad k = 1, 2, \dots.$$

هم‌چنین در مورد مدل‌های $ARMA$ ، $C_k = \pi_k$ خواهد بود.

با استفاده از دترمینان معادله‌ی (۲-۳) و رابطه‌ی (۳-۳) ورودی‌های جدول C برای مدل $ARMA(1,1)$ به

دست می‌آید،

به عنوان نمونه

$$C(1/2) = \begin{vmatrix} 1 & C_1 \\ C_1 & C_2 \end{vmatrix} = C_2 - C_1^2 = \phi_1(\theta_1 - \phi_1)$$

$$C(2/2) = \begin{vmatrix} C_1 & C_2 \\ C_2 & C_3 \end{vmatrix} = C_1 C_3 - C_2^2 = 0$$

$$C(3/2) = C_2 C_4 - C_3^2 = 0 \dots$$

هم چنین بر اساس معادله ی (۳-۳) داریم

$$C(1/3) = \frac{C(0/2)C(2/2) - C(1/2)^2}{C(1/1)} = 1 - \phi_1^2(\theta_1 - \phi_1).$$

باقی ورودی های جدول C به شکل مشابه تعریف خواهند شد [۳۹]. با جایگذاری مقادیر به دست آمده، جدول C متناظر با مدل $ARMA(1,1)$ تعریف می شود.

q	p	۰	۱	۳	۴	...
۰		۱	۱	۱	۱	...
۱		۱	$(\theta_1 - \phi_1)$	$\theta_1(\theta_1 - \phi_1)$	$\theta_1^2(\theta_1 - \phi_1)$...
۲		-۱	$\phi_1(\theta_1 - \phi_1)$	*	*	
۳		-۱	$-\phi_1^2(\theta_1 - \phi_1)$	*	*	
⋮						

جدول ۳-۲: جدول C متناظر با مدل $ARMA(1,1)$

در جدول فوق ملاحظه می شود که از وقفه های دوم به بعد بلوک نامحدود صفر نمایان است، بنابراین رتبه ی مدل $ARMA$ می تواند از طریق این جدول مشخص شود.

۴-۳ مدل گارچ گویا ($RGARCH$)

بر اساس اعمال تقریب پایه روی مدل گارچ از خانواده ی مدل های $ARCH$ ، تابع گویایی که خواص نامتقارن فرایند نوسانات را دریافت می کند به صورت تقسیم دو چند جمله ای از بازده ها و نوسانات گذشته تعریف می شود

$$\sigma_t^2 = \frac{P(r_{t-1}, r_{t-2}, \dots; \sigma_{t-1}, \sigma_{t-2}, \dots)}{Q(r_{t-1}, r_{t-2}, \dots; \sigma_{t-1}, \sigma_{t-2}, \dots)}.$$

یک بازدهی دارایی در زمان t به شکل $r_t = \sigma_t \varepsilon_t$ خواهد بود که ε_t اجزای خطا از یک توزیع مشخص با میانگین صفر و واریانس یک در نظر گرفته می شود. می توان برای دریافت بهتر کشیدگی در بازده های غیرشرطی از توزیع های تی و نرمال برای ε_t ها استفاده کرد. ساده ترین حالت تقریب گویا برای مدل های نامتقارن گارچ شامل اولین وقفه زمانی از نوسانات و بازده ها با عنوان $RGARCH(1,1)$ است در این صورت داریم

$$\sigma_t^2 = \frac{\omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2}{1 + \delta r_{t-1}} \quad (4-3)$$

با در نظر گرفتن ضریب مخرج $\delta = 0$ ، مدل $RGARCH$ به مدل $GARCH(1,1)$ تبدیل خواهد شد. در معادله ی فرایند نوسانی گویا، δ نامتقارنی را در مقابل r_{t-1} نشان می دهد همچنین اگر δ مثبت باشد نوسانات واکنش شدیدتری نسبت به اخبار بد (بازده های منفی) نشان می دهند که گویای وجود اثر اهرم در فرایند نوسانی است [۴۰].

در حالتی که مخرج معادله ی (۳-۴) منفی باشد ($1 < -\delta r_{t-1}$)، فرایند نوسانی به خوبی تعریف نمی شود که ممکن است برای بازده های بزرگ و یا دورافتاده اتفاق بیفتد. برای پوشش این مشکل مدل جایگزینی از مدل گارچ گویا با عنوان $RGARCH - EXP$ ارائه شده است

$$\sigma_t^2 = \frac{\omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2}{\exp(\delta r_{t-1})} \quad (5-3)$$

که می تواند یک تعمیم بهینه از مدل گارچ گویا در جهت بهبود کارایی آن معرفی شود. برای یک مقدار δr_{t-1} کوچک، معادله ی (۳-۵) به معادله ی (۳-۴) تبدیل می شود.

۳-۴-۱ مدل گارچ و تبدیل های نامتقارن آن

برای مقایسه ی مدل جدید $RGARCH$ با سایر مدل های خانواده ی $ARCH$ ، از مدل های $GARCH$ ، $EGARCH$ و GJR که ویژگی های نامتقارن دارند استفاده می کنیم. در همگی این مدل ها نوسانات به صورت تابعی از بازده ها و نوسانات گذشته خواهد بود. می توان برای پیش بینی دقیق تر نوسانات آینده ی فرایند از شکل های پیچیده تری برای هر کدام استفاده کرد اما برآورد پارامترهای آن ناممکن خواهد بود. با در نظر گرفتن پایین ترین وقفه زمانی، فرایند نوسانات برای هر کدام تعریف می شود.

مدل GARCH

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2.$$

مدل EGARCH

$$\begin{aligned} \ln(\sigma_t^2) &= \omega + \beta \ln(\sigma_{t-1}^2) + \theta Z_{t-1} + r(|Z_{t-1}| - E(|Z_{t-1}|)) \\ &= \omega + \beta \ln(\sigma_{t-1}^2) + \theta Z_{t-1} + r(|Z_{t-1}| - \sqrt{\frac{2}{\pi}}). \end{aligned}$$

که در آن خواهیم داشت

$$Z_{t-1} = \frac{r_{t-1}}{\sigma_{t-1}}.$$

مدل (TARCH)GJR

$$\sigma_t^2 = \begin{cases} \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 & , r_t \geq 0 \\ \omega + (\alpha + \rho) r_{t-1}^2 + \beta \sigma_{t-1}^2 & , r_{t-1} < 0 \end{cases}.$$

در این مدل‌ها پارامترهای θ و ρ نامتقارنی را در فرایند نوسانات نشان می‌دهند [۴۰]، [۴۱].

۲-۴-۳ استنتاج بیزی

به منظور برآورد پارامترهای مدل گارچ گویا از روش استنتاج بیزی و نظریه‌ی بیز که در فصل قبل به آن اشاره شد، استفاده می‌کنیم. بر این اساس، توزیع پسین $P(\theta|r)$ با بردار مشاهدات $r = (r_1, r_2, \dots, r_n)$ برای بردار پارامتر $\theta = (\alpha, \beta, \omega, \delta)$ به شکل زیر تعریف می‌شود

$$P(\theta|r) \propto L(r|\theta)\pi(\theta).$$

که در آن $L(r|\theta)$ تابع درستنمایی و $\pi(\theta)$ چگالی پیشین پارامتر θ است. جمله‌ی خطای ε_t در فرایندهای نوسانی از یک توزیع مشخص با میانگین صفر و واریانس یک در نظر گرفته می‌شود که از توزیع‌های نرمال و تی می‌توان استفاده کرد.

در این حالت تابع درستنمایی بازده‌های دارای r بر اساس پارامترهای مدل تعریف خواهد شد

توزیع نرمال برای خطاها

$$L(r|\theta) = \prod_{i=1}^n f(r_i|\theta) = (\sqrt{\pi})^{-n} \prod_{i=1}^n \frac{1}{\sigma_i} \exp\left(\frac{-r_i^2}{\sigma_i^2}\right) \quad [۴۲].$$

توزیع تی برای خطاها

$$L(r|\theta) = \prod_{i=1}^n f(r_i|\theta) = \prod_{i=1}^n \frac{\Gamma\left(\frac{\mu+1}{2}\right)}{\Gamma\left(\frac{\mu}{2}\right) \sqrt{\pi\sigma_i^2(\mu-2)}} \left(1 + \frac{r_i^2}{(\mu-2)\sigma_i^2}\right)^{-\left(\frac{\mu+1}{2}\right)} \quad [۴۳].$$

با استفاده از چگالی پسین $P(\theta|r)$ برآوردگر بیز بردار پارامتر θ به صورت یک مقدار موردانتظار به دست خواهد آمد

$$\langle \theta \rangle = E(\theta|r) = \frac{\int \theta P(\theta|r) d\theta}{\int P(\theta|r) d\theta}.$$

که $Z = \int P(\theta|r) d\theta$ یک ثابت استاندارد ساز در مخرج کسر است [۴۴].

با ثابت در نظر گرفتن توزیع پیشین بردار θ ، $\pi(\theta)$ ، توزیع پسین مدل متناسب با تابع درستنمایی $L(r|\theta)$ می باشد. در برخی موارد توزیع پسین $P(\theta|r)$ در شکل بسته موجود نمی باشد از این رو با استفاده از روش زنجیر مارکف مونت کارلو بردار θ برآورد می شود. ایده‌ی اصلی این روش به این صورت است که ابتدا یک مقدار اولیه از پارامتر θ از توزیع $P(\theta|r)$ نمونه‌گیری می شود سپس با به کارگیری یک زنجیر مارکف حول P ، نمونه‌ی $(\theta^{(1)}, \dots, \theta^{(k)}) = \theta^{(i)}$ به دست می آید که k متناظر با تعداد نمونه‌های استخراج شده است.

بعد از جمع آوری نمونه‌ی تصادفی، برآورد پارامتر موردنظر به صورت زیر نمایش داده می شود

$$\langle \theta \rangle = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k \theta^{(i)}.$$

بر اساس انتخاب توزیع خطا با میانگین صفر و واریانس یک برای هر مدل، خطای آماری برای k داده‌ی مستقل برابر $\frac{1}{\sqrt{k}}$ خواهد بود، اما داده‌های به دست آمده از شبیه‌سازی $MCMC$ همبسته هستند زیرا انتخاب هر نمونه بر اساس نمونه‌ی قبلی انجام می گیرد در نتیجه خطای آماری متناظر با $\sqrt{\frac{\nu T}{k}}$ می باشد که T نشان‌دهنده‌ی خودهمبستگی زمانی بین داده‌های نمونه برداری شده است [۴۵].

۳-۴-۳ شبیه‌سازی الگوریتم متروپلیس-هستینگز

یک الگوریتم بسیار کارآمد برای نمونه‌برداری از یک توزیع داده شده، الگوریتم متروپلیس-هستینگز می‌باشد که تعمیمی بر روش متروپلیس است.

با شروع از نمونه‌ی اولیه‌ی θ از توزیع هدف

$$P(\theta|r) = L(r|\theta) = \prod_{i=1}^n (\nu\pi\sigma_i^2)^{-\frac{1}{\nu}} \exp\left(-\frac{r_i^2}{\nu\sigma_i^2}\right)$$

یک چگالی پیشنهادی $q(\theta' | \theta)$ از توزیع تی چند متغیره از بُعد تعداد پارامترهای مدل ارائه می‌شود

$$q(\theta) = \frac{\Gamma(\frac{\nu+p}{\nu})/\Gamma(\frac{\nu}{\nu})}{\det\Sigma^{1/\nu}(\nu\pi)^{p/\nu}} \times \left[1 + \frac{(\theta - M)^T \Sigma^{-1} (\theta - M)}{\nu}\right]^{-\frac{\nu+p}{\nu}}.$$

بر اساس بردار پارامتر θ در مدل گارچ گویا، $p = 4$ خواهد بود. از آنجاکه با افزایش پارامتر درجه‌ی آزادی، توزیع تی به توزیع هدف نرمال همگرا می‌شود، $\nu = 10$ قرار می‌دهیم. بردارهای پارامتر M و θ در توزیع تی به صورت زیر نشان داده می‌شوند

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \omega \\ \delta \end{bmatrix}, \quad M = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix}$$

که $M_i = E(\theta_i)$ بردار میانگین توزیع تی است. Σ یک ماتریس معین مثبت نشان‌دهنده‌ی ماتریس

کواریانس توزیع به شکل زیر است

$$\frac{\nu\Sigma}{\nu-2} = E[(\theta - M)(\theta - M)^T].$$

برای برآورد پارامترهای M و Σ از توزیع q ، ابتدا الگوریتم متروپلیس را با در نظر گرفتن q به عنوان توزیع

هدف، به کار می‌گیریم. مقادیر اولیه‌ی دلخواه پیشنهادی برای شروع الگوریتم عبارت‌انداز

$$\Sigma = \begin{bmatrix} \circ/1 & \circ & \circ & \circ \\ \circ & \circ/2 & \circ & \circ \\ \circ & \circ & \circ/3 & \circ \\ \circ & \circ & \circ & \circ/4 \end{bmatrix}, \quad M = \begin{bmatrix} \circ/1 \\ \circ/4 \\ \circ/2 \\ \circ/3 \end{bmatrix}$$

با به کارگیری الگوریتم در یک مقدار نمونه‌ی برآورد شده‌ی بزرگ، پارامترهای M و Σ پیشنهادی به مقادیر واقعی خود همگرا می‌شوند. توزیع پیشنهادی برای اجرای شبیه‌سازی، توزیع نرمال چند متغیره با میانگین θ و ماتریس کواریانس Σ در نظر گرفته شده است. بعد از حذف ۵۰۰۰ نمونه‌ی ابتدایی به دست آمده به عنوان فرآیند سوخته، ۵۵۰۰۰ نمونه‌ی پارامتر از توزیع q استخراج شد که مقادیر M و Σ برآورد شده به صورت زیر است

$$\Sigma = \begin{bmatrix} 0.06575159 & -0.01094036 & 0.01735915 & 0.03527002 \\ -0.01094036 & 0.16921456 & 0.01661529 & 0.01430978 \\ 0.01735915 & 0.01661529 & 0.26095376 & 0.09571424 \\ 0.03527002 & 0.01430978 & 0.09571424 & 0.63836144 \end{bmatrix}$$

$$M = \begin{bmatrix} 0.3997452 \\ 0.4611378 \\ 0.3473960 \\ 0.1959700 \end{bmatrix}$$

اکنون با در نظر گرفتن توزیع $q(\theta)$ با پارامترهای معلوم به عنوان توزیع پیشنهادی الگوریتم MH با تابع هدف $P(\theta|r)$ ، پارامترهای مدل گارچ گویا برآورد خواهد شد [۴۱]، [۴۴]، [۴۵].

۵-۳ نتایج تجربی

به منظور ارزیابی مدل $RGARCH$ و مقایسه‌ی آن با سایر مدل‌های گارچ، از مجموعه داده‌های سری زمانی زیر استفاده شده است

S_1 (قیمت روزانه‌ی بسته شده‌ی سهام شرکت IBM^{37} ، ۲۹ ژوئن ۱۹۵۹ - ۳۰ سپتامبر ۱۹۶۰)

S_2 (شاخص سهام سه ماهه‌ی $S\&P^{38}$ ۵۰۰، ۱۹۰۰ - ۱۹۹۶)

S_3 (نرخ بهره‌ی ماهانه‌ی اوراق قرضه‌ی بانک مرکزی دولت استرالیا، ۶ ژوئن ۱۹۶۹ - سپتامبر ۱۹۹۴)

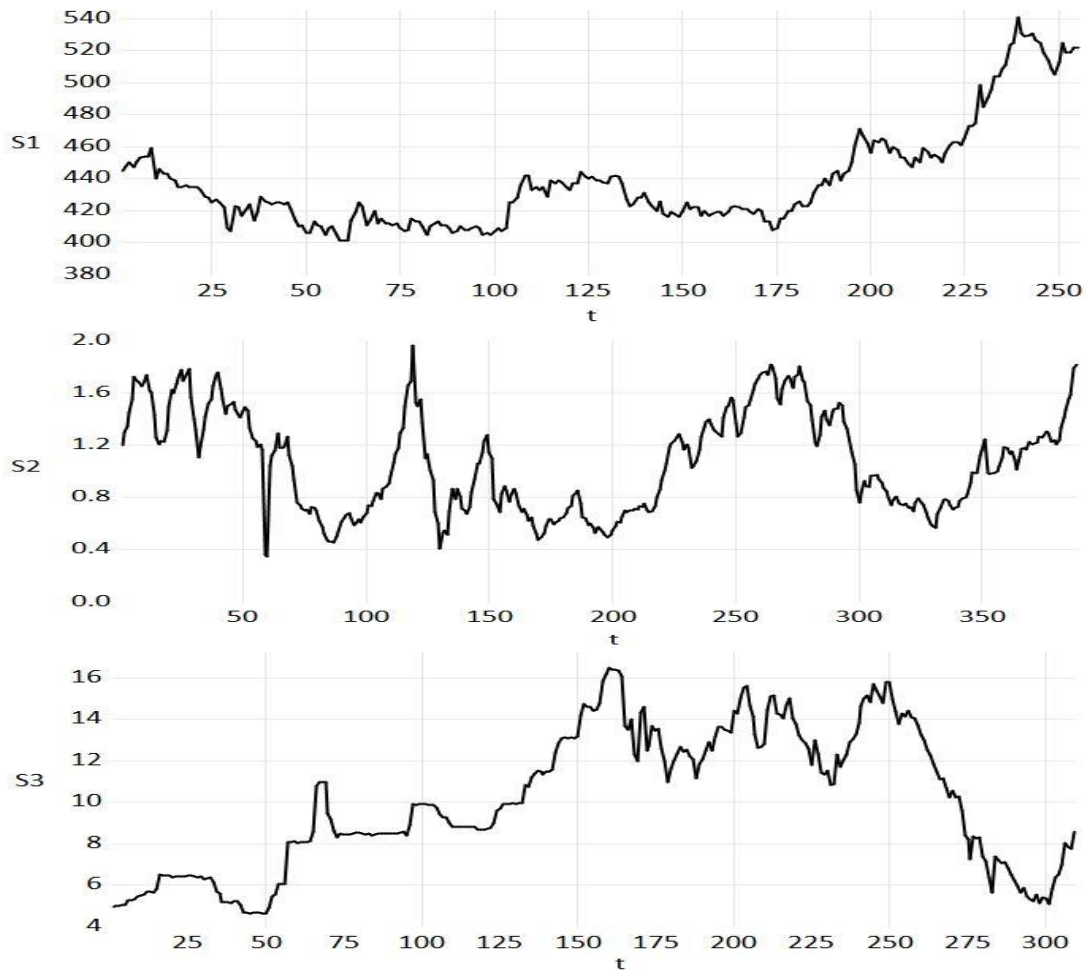
[۴۹].

برای تعیین الگوی خطی بر روی هر یک از داده‌های سری زمانی و پیش‌بینی مقادیر آینده می‌بایست تغییرات روند هر کدام مورد بررسی قرار گیرد و نامانایی برطرف شود. سری‌های تحت بررسی روندهای نامانایی را در میانگین خود دارند.

International Business Machines Corporation ³⁷

Standard and Poor's index ³⁸

شکل (۱-۳) روند تغییرات سری‌های زمانی S_1 ، S_2 و S_3 را در طول زمان نشان می‌دهد که بیان‌گر وجود نامانایی در میانگین و واریانس برای هر کدام است.

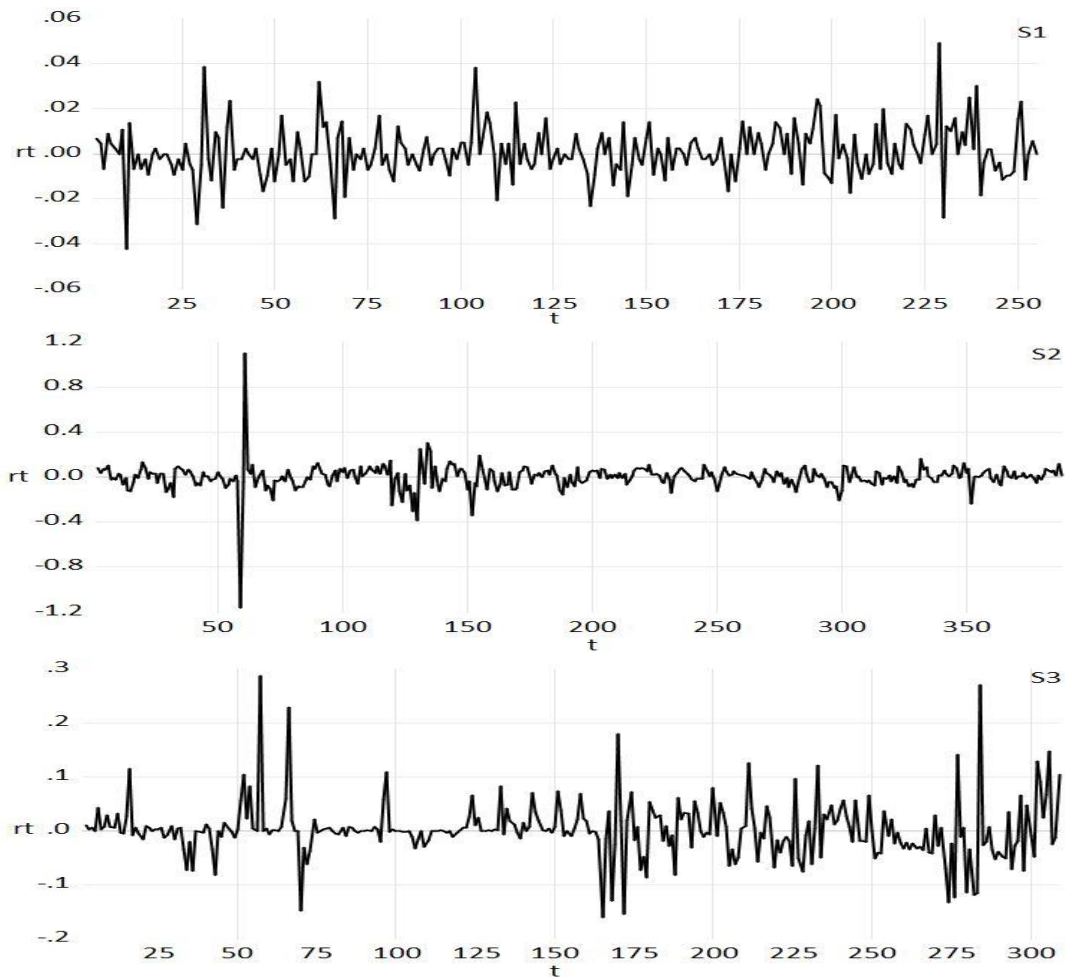


شکل ۱-۳: نمودارهای تغییرات روند سری‌های زمانی S_1 ، S_2 و S_3

بر اساس شکل فوق ملاحظه می‌شود که داده‌های سری‌های زمانی تغییرات نامانایی را در میانگین و واریانس خود نشان می‌دهند. به منظور حذف تغییرات میانگین در طول زمان برای هر سری، بازده‌های مانا برای مدل‌بندی با

استفاده از تبدیل لگاریتم تفاضلی $r_t = 100 \times (Ln(P_t) - Ln(P_{t-1}))$ به دست آمده است که P_t قیمت بسته شده در زمان t است.

شکل (۲-۳) نمودار بازده‌های مانای تبدیل لگاریتم برای هر یک از سری‌ها را نشان می‌دهد که دارای روند مانا در میانگین است.



شکل ۳-۲: نمودارهای بازده‌های S_1, S_2, S_3 و S_3 بعد از اعمال تبدیل لگاریتم تفاضلی

برای تعیین معادله‌ی واریانس ناهمسان شرطی و برآورد پارامترهای مدل از استنتاج بیزی استفاده می‌کنیم، که برای انجام آن از الگوریتم MH با چگالی پیشنهادی توزیع تی چند بُعدی کمک گرفته شده است. با شروع از یک مقدار اولیه‌ی دلخواه همانند بخش قبل پارامترهای به‌دست آمده در هر مرحله به‌روزرسانی خواهد شد و بعد از حذف ۲۰۰۰ نمونه‌ی اول فرآیند سوخته برای کاهش خطای تقریب و ایجاد برآورد دقیق‌تر از توزیع $P(\theta|r)$ ، ۱۰۰۰۰ نمونه‌ی مونت‌کارلو برای برآورد نهایی پارامترهای مدل مورد بررسی قرار می‌گیرد.

برای مقایسه‌ی کارایی مدل گارچ گویا، نتایج به‌دست آمده از آن با مدل‌های $EGARCH$ ، GJR و $GARCH$ که خواص نامتقارن فرایند نوسانات را دربر می‌گیرند، مورد تحلیل قرار گرفته است. شکل (۳-۳) مقادیر خودهمبستگی و خودهمبستگی جزئی سری‌های زمانی S_1 ، S_2 و S_3 را بعد از اعمال تبدیل لگاریتم تفاضلی نمایش می‌دهد که براساس آن معادله‌ی میانگین برای هرکدام تنظیم خواهد شد.

S1							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	-0.073	-0.073	1.3616	0.243	
		2	-0.046	-0.051	1.9055	0.386	
		3	0.033	0.026	2.1804	0.536	
		4	0.019	0.021	2.2725	0.686	
		5	-0.062	-0.056	3.2658	0.659	
		6	0.126	0.119	7.4113	0.284	
		7	0.027	0.039	7.6043	0.369	
		8	0.073	0.094	8.9951	0.343	
		9	-0.125	-0.119	13.171	0.155	
		10	0.044	0.027	13.678	0.188	
		11	0.078	0.082	15.317	0.168	
		12	-0.036	-0.031	15.671	0.207	
		13	-0.006	-0.002	15.682	0.267	
		14	-0.028	-0.077	15.895	0.320	
		15	-0.024	0.000	16.046	0.379	
		16	-0.020	-0.025	16.155	0.442	
		17	0.055	0.050	16.981	0.456	
		18	-0.017	-0.029	17.059	0.519	
		19	0.034	0.034	17.381	0.564	
		20	-0.052	-0.018	18.129	0.579	

S2							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	0.157	0.157	9.6277	0.002	
		2	-0.180	-0.210	22.345	0.000	
		3	0.041	0.116	22.989	0.000	
		4	-0.028	-0.106	23.305	0.000	
		5	-0.086	-0.030	26.209	0.000	
		6	0.012	0.006	26.267	0.000	
		7	-0.079	-0.113	28.739	0.000	
		8	-0.087	-0.036	31.732	0.000	
		9	-0.060	-0.096	33.174	0.000	
		10	0.027	0.046	33.473	0.000	
		11	-0.017	-0.069	33.588	0.000	
		12	0.062	0.093	35.128	0.000	
		13	0.037	-0.033	35.677	0.001	
		14	-0.029	-0.010	36.026	0.001	
		15	-0.009	-0.013	36.060	0.002	
		16	0.034	0.007	36.539	0.002	
		17	0.037	0.049	37.094	0.003	
		18	0.026	0.002	37.373	0.005	
		19	-0.082	-0.069	40.093	0.003	
		20	-0.100	-0.073	44.229	0.001	

S3							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	0.079	0.079	1.9444	0.163	
		2	-0.006	-0.013	1.9568	0.376	
		3	0.117	0.119	6.2373	0.101	
		4	-0.011	-0.031	6.2763	0.179	
		5	-0.064	-0.058	7.5507	0.183	
		6	0.047	0.043	8.2455	0.221	
		7	0.176	0.176	18.022	0.012	
		8	-0.041	-0.058	18.559	0.017	
		9	0.189	0.198	29.929	0.000	
		10	-0.022	-0.110	30.080	0.001	
		11	-0.005	0.047	30.089	0.002	
		12	0.041	0.004	30.632	0.002	
		13	-0.019	-0.016	30.750	0.004	
		14	-0.053	-0.068	31.648	0.004	
		15	-0.010	0.001	31.682	0.007	
		16	-0.013	-0.090	31.737	0.011	
		17	-0.087	-0.015	34.242	0.008	
		18	-0.032	-0.093	34.583	0.011	
		19	0.001	0.039	34.584	0.016	
		20	-0.031	-0.035	34.897	0.021	

شکل ۳-۳: نمودارهای خودهمبستگی و خودهمبستگی جزئی بازده‌های مانای سری‌های زمانی S_1 ، S_2 و S_3

مقادیر ستون AC نشان‌دهنده‌ی خودهمبستگی سری زمانی در وقفه‌های مختلف است و براساس رابطه‌ی زیر به دست می‌آید

$$\tau_k = \frac{\sum_{t=k+1}^T ((Y_t - \bar{Y})(Y_{t-k} - \overline{Y_{t-k}})) / (T - K)}{\sum_{t=1}^T (Y_t - \bar{Y})^2 / T}$$

که در آن میانگین نمونه‌ای Y به عنوان سری زمانی مورد بررسی است و $\overline{Y_{t-k}} = \sum Y_{t-k} / T - k$. خطوط نقطه‌چین در دو طرف نمودار خودهمبستگی دو حد خطای استاندارد به صورت $\pm 2/(\sqrt{T})$ است که T تعداد کل مشاهدات می‌باشد. مقادیر τ_k درون این بازه در یک فاصله‌ی اطمینان ۵٪ موافق صفر هستند و در تعیین معادله میانگین نقشی ایفا نمی‌کنند.

ستون PAC مربوط به مقادیر خودهمبستگی جزئی در وقفه‌ی k ام است و داریم

$$\phi_k = \begin{cases} \tau_1, & k = 1 \\ \frac{\tau_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \tau_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \tau_{k-j}}, & k > 1 \end{cases}$$

τ_k خودهمبستگی برآوردشده در وقفه‌ی k ام می‌باشد و داریم

$$\phi_{k,j} = \phi_{k-1,j} - \phi_k \phi_{k-1,k-j}$$

همانند ستون خودهمبستگی خطوط نقطه‌چین در دو طرف نمودار خودهمبستگی جزئی دو حد خطای استاندارد به صورت $\pm 2/(\sqrt{T})$ است که در سطح معنی داری ۵٪ موافق صفر است.

دو ستون آخر شکل (۳-۳) مربوط به آماره‌ی Q آزمون لجانگ-باکس^{۳۹} و مقادیر p -value برای هر کدام است. آماره‌ی Q در وقفه‌ی k ام، یک آزمون آماری برای فرضیه‌ی صفری است که بیان می‌کند هیچ

همبستگی تا مرتبه‌ی k ام وجود ندارد و به شکل زیر تعریف می‌شود

$$Q_{LB} = T(T + 2) \sum_{j=1}^k \frac{\tau_j^2}{T - j}$$

[۴۶].

با به کارگیری الگوریتم MH و نمونه برداری از توزیع $P(\theta|r)$ مقادیر پارامتر معادله‌ی واریانس شرطی برای هر مدل به صورت جدول (۳-۳) شکل می‌گیرد. با افزایش تعداد نمونه‌های دریافتی از تابع هدف، پارامترهای

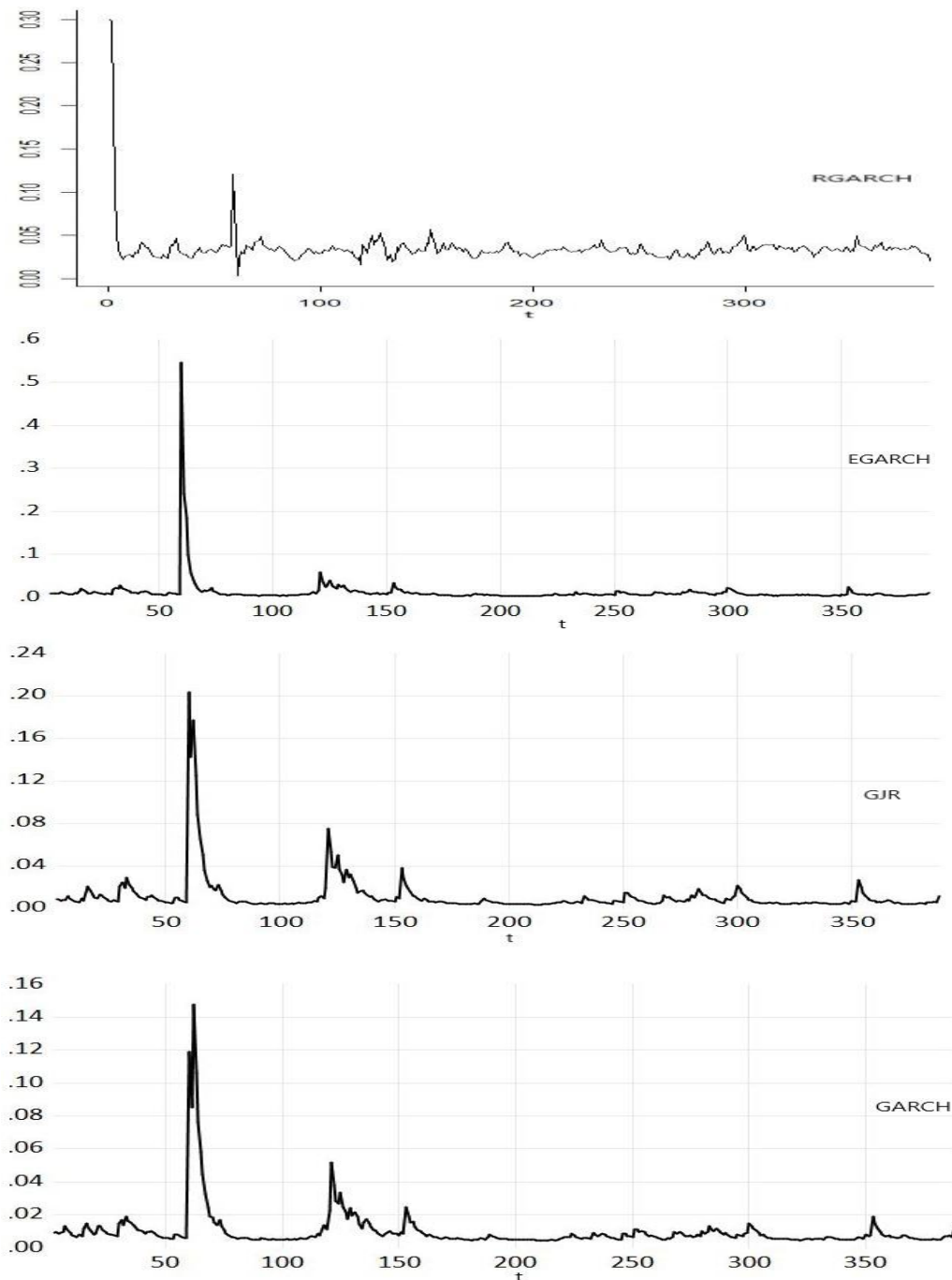
به دست آمده به مقادیر واقعی خود همگرا خواهند شد، در نتیجه پارامترهای مدل برای هر سری براساس توزیع نرمال برای جملات خطای مدل به دست آمده است.

<i>RGARCH</i>	S_1	S_2	S_3
α	۰/۲۱۳۶۲۵۸	۰/۰۰۲۱۴۰۷۶	۰/۲۳۷۹۴۰۳
β	۰/۴۶۱۸۱۹۴	۰/۲۷۹۸۷۵۵	۰/۲۹۹۹۰۲۶
ω	۰/۷۷۲۸۲۵	۰/۰۲۶۸۷۷۰۷	۰/۵۲۰۹۰۳۸
δ	۰/۲۴۱۴۱۰۱۱	-۰/۰۵۰۵۹۰۰۹	۰/۷۲۱۸۶۲۶
<i>EGARCH</i>	S_1	S_2	S_3
β	۰/۷۱۹۰۷۸	۰/۸۶۹۳۶۲	۰/۹۴۰۰۸۱
ω	۰/۷۷۴۰۶۳	-۰/۹۰۶۱۱۸	-۰/۲۳۴۵۰۶
γ	۰/۱۶۴۶۴۶	۰/۳۹۲۵۷۳	۰/۲۴۷۷۳۱
θ	۰/۱۱۳۲۱۴	-۰/۰۷۸۵۱۹	۰/۰۵۸۳۰۴
<i>GJR</i>	S_1	S_2	S_3
α	۰/۰۸۸۵۱۱	۰/۱۵۸۷۰۳	۰/۱۸۹۶۴۰
β	۰/۷۷۵۲۲۵	۰/۶۹۴۸۷۲	۰/۸۴۶۲۱۹
ω	۴/۴۴۱۶۶۳	۰/۰۰۱۰۸۶	۰/۰۰۷۷۹۲
ρ	-۰/۰۹۵۵۰۱	۰/۱۵۲۴۰۰	-۰/۰۶۸۲۶۳
<i>GARCH</i>	S_1	S_2	S_3
α	۰/۰۷۷۶۵۴	۰/۱۷۷۹۳۱	۰/۱۰۹۷۵۳
β	۰/۶۶۶۲۲۵	۰/۷۰۲۸۹۹	۰/۸۷۶۴۴۰
ω	۶/۵۴۸۶۷۱	۰/۰۰۱۲۸۶	۰/۰۰۸۳۷۹

جدول ۳-۳: نتایج برآورد پارامتر مدل‌های خانواده‌ی گارچ و مدل گارچ گویا با توزیع خطای نرمال

در جدول فوق مقادیر پارامترهای δ ، ρ و θ که عامل ایجاد نامتقارنی در فرایندهای نوسانی خانواده‌ی *ARCH* و مدل جدید گویا هستند، به دست آمده است. تمامی این مقادیر با عدم تقارن پدیده‌ی اثر اهرم و وجود آن روی هر فرایند سازگار است.

شکل (۳-۴) نوسانات برآورد شده توسط مدل‌های تحت مقایسه را به همراه مدل جدید با در نظر گرفتن توزیع نرمال برای فرایند خطای سری زمانی S_2 نشان می‌دهد. در تمامی این نوسانات تغییرات مشابهی در طول زمان مشاهده می‌شود و در یک دوره‌ی زمانی ابتدایی مشخص یک خوشه‌ی نوسانی بزرگ قابل مشاهده است.



شکل ۳-۴: نوسانات برآوردشده‌ی سری زمانی S_t از مدل‌های خانواده‌ی گارچ و مدل گارچ گویا

۳-۶ معیار انحراف اطلاعات (DIC) برای آزمون نیکویی برازش هر مدل

برای مقایسه‌ی مدل‌های مختلف تحت بررسی، از معیار بیزی انحراف اطلاعات استفاده می‌کنیم. همان‌طور

که در فصل قبل عنوان شد، براساس تابع درستنمایی مشاهدات هر مدل، $L(r|\theta)$ ، یک معیار انحراف

به‌صورت زیر تعریف می‌شود

$$D(\theta) = -\nu \text{Log}(L(r|\theta)) + \nu \text{Log}(f(r)).$$

که در آن θ ، مجموعه‌ی پارامترهای مدل و $f(r)$ ، یک تابع مشخص از مشاهدات به‌تنهایی است. در این‌جا این عبارت ثابت و به‌شکل $f(r) = 1$ در نظر گرفته می‌شود. براساس این تعریف، معیار انحراف اطلاعات برای هر مدل مشخص با پارامترهای معلوم به‌صورت زیر خواهد بود

$$DIC = 2\overline{D(\theta)} - D(\tilde{\theta}) = -4E_{\theta|r}[\text{Log}(L(r|\theta))] + 2\text{Log}(L(r|\tilde{\theta})) \quad (3-6)$$

بعد از تعیین پارامترهای مدل از طریق شبیه‌سازی مقدار تابع درست‌نمایی هر مدل براساس مشاهدات سری زمانی مربوط به‌دست می‌آید که معیار انحراف اطلاعات برای هر کدام از مدل‌ها از طریق رابطه‌ی (۳-۶) محاسبه خواهد شد. جدول زیر مقادیر DIC برای هر کدام از مدل‌ها را دربر دارد و براساس این مقادیر رتبه بندی هر مدل برای مقایسه‌ی کارایی آن‌ها در ثبت پاسخ‌های نامتقارن فرایند نوسانی مشخص شده است. از بین مدل‌های مقایسه‌ای، مدلی که مقدار معیار DIC کمتری نسبت به باقی تعمیم‌ها داشته باشد کارایی بهتری را در ثبت نوسانات فرایند از خود نشان می‌دهد که در میان آن‌ها، مدل گارچ گویا در اکثر مواقع رتبه‌ی اول را در بین سایر مدل‌ها دارد.

	S_1	S_2	S_3
<i>RGARCH</i>	۷۹۴/۰۳۲۴	-۳۶۳/۳۹۲۱	۱۲۸۹/۵۷۱
<i>EGARCH</i>	۸۲۲/۲۷۹۵	-۱۳۰/۲۳۹۶	۴۴۴۳/۲۸۱
<i>GJR</i>	۱۲۳۷/۳۸	-۷۹/۹۵۳۰۴	۱۲۹۱/۷۲۹
<i>GARCH</i>	۱۲۳۷/۷۸۳	-۱۵۶/۶۹۲۵	۱۲۸۸/۷۲۵

جدول ۳-۴: مقادیر DIC برای مدل‌های خانواده‌ی گارچ براساس توزیع خطای نرمال

	S_1	S_2	S_3
<i>RGARCH</i>	۱	۱	۲
<i>EGARCH</i>	۲	۳	۴
<i>GJR</i>	۳	۴	۳
<i>GARCH</i>	۴	۲	۱

جدول ۳-۵: رتبه‌بندی مدل‌ها براساس معیار انحراف اطلاعات

۷-۳ منحنی تأثیر اخبار^{۴۰} (NI)

با وجود این‌که مدل‌های خانواده‌ی $ARCH$ خواص نامتقارن فرایند نوسانات را ثبت می‌کنند و اثر اهرم برای

هر مدل قابل مشاهده است، هر مدل پاسخ نوسانی متقارن متفاوتی نسبت به بازده‌ها در زمان t دارد. منحنی‌های تأثیر اخبار با ثابت نگه‌داشتن اطلاعات از زمان $t - 2$ ، تأثیر اخبار روی واریانس ناهمسان شرطی و تعیین ارتباط بین بازده‌ها و نوسانات را مشخص می‌کنند.

تعریف ۱-۷-۳. فرض کنید y_t ، نشان‌دهنده‌ی نرخ بازدهی یک سهام مشخص در زمان t باشد و F_{t-1} حاوی تمام اطلاعات تا زمان $t - 1$ شامل مقادیر واقعی پارامترها باشد. بر این اساس مقدار موردانتظار شرطی y_t نسبت به F_{t-1} و واریانس شرطی y_t نسبت به F_{t-1} به صورت زیر تعریف می‌شود

$$\sigma_t^2 = \text{Var}(y_t | F_{t-1}) \quad , \quad m_t = E(y_t | F_{t-1}).$$

با استفاده از این تعریف یک بازدهی غیر قابل انتظار در زمان t به صورت $r_t = y_t - m_t$ خواهد بود. r_t به عنوان یک اندازه‌ی جمعی از اخبار در زمان t شناخته می‌شود [۴۷].

نکته ۲-۷-۳. ارتباط بین بازده‌ها و نوسانات در مواجهه با اثر اهرم روی فرایند نوسانی برای هر مدل از خانواده‌ی ARCH براساس منحنی‌های تأثیر اخبار هرکدام به صورت زیر تعریف می‌شود که در آن‌ها، $\sigma_t^2 = \text{Var}(r_t)$ ، بیان‌گر واریانس بازده‌ها است.

مدل RGARCH

$$\sigma_t^2 = \frac{\omega + \beta\sigma^2 + \alpha r_{t-1}^2}{1 + \delta r_{t-1}}.$$

مدل EGARCH

$$\sigma_t^2 = \begin{cases} \sigma^2 \beta \exp\left(\omega - r\sqrt{2/\pi}\right) \exp\left(\frac{\theta + \gamma}{\sigma} r_{t-1}\right) & , r_{t-1} \geq 0 \\ \sigma^2 \beta \exp\left(\omega - r\sqrt{2/\pi}\right) \exp\left(\frac{\theta + \gamma}{\sigma} r_{t-1}\right) & , r_{t-1} < 0 \end{cases}.$$

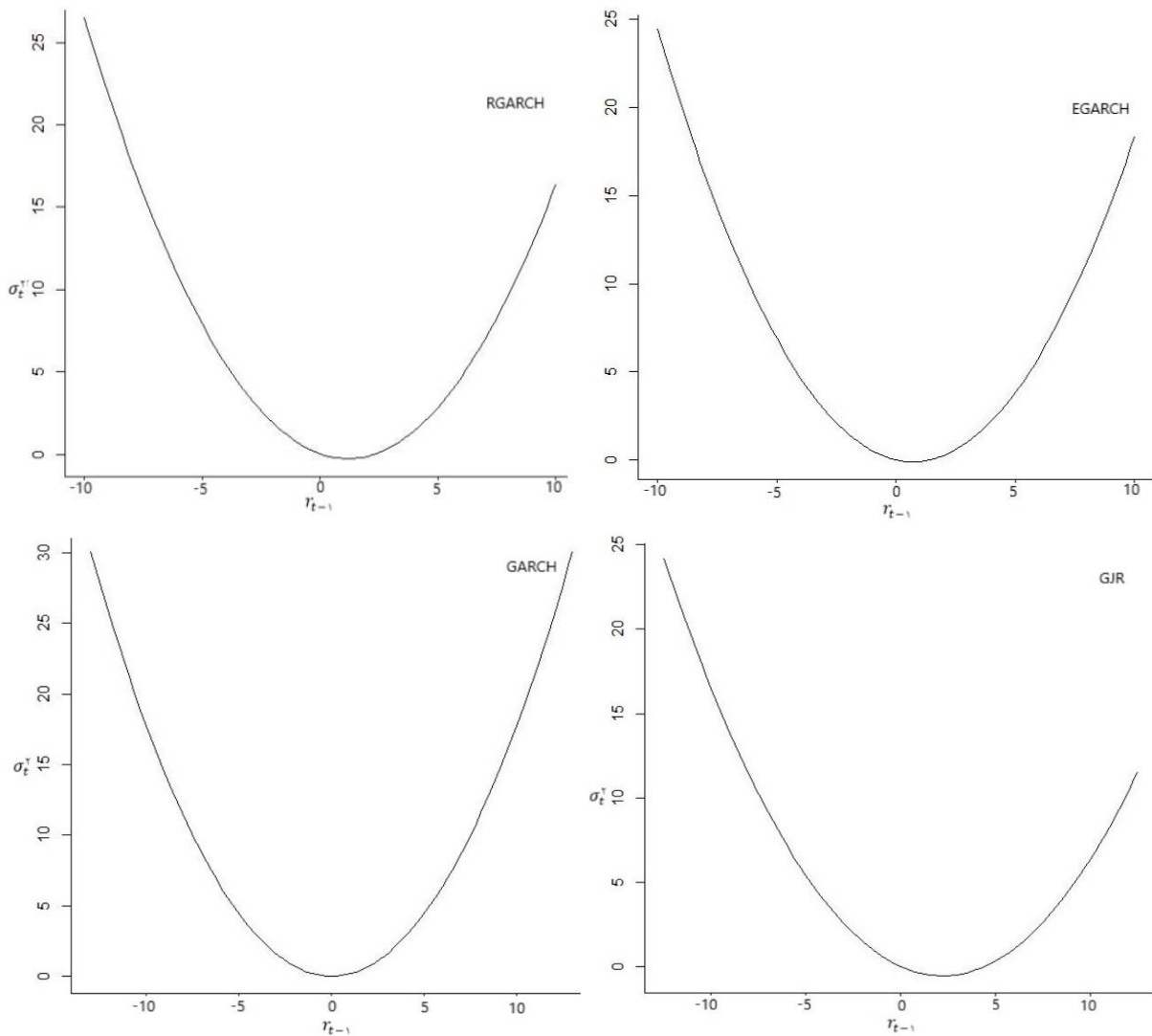
مدل GJR

$$\sigma_t^2 = \begin{cases} \omega + \beta\sigma^2 + \alpha r_{t-1}^2 & r_{t-1} \geq 0 \\ \omega + \beta\sigma^2 + (\alpha + \rho) r_{t-1}^2 & r_{t-1} < 0 \end{cases}.$$

مدل GARCH

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma^2.$$

شکل (۳-۵) نمایش منحنی‌های تاثیر اخبار مدل‌های تحت مقایسه‌ی خانواده‌ی $ARCH$ را همراه با مدل گارچ گویا با جملات خطای نرمال برای هر کدام نشان می‌دهد.



شکل ۳-۵: نمایش منحنی‌های NI مدل‌های خانواده‌ی گارچ و مدل گارچ گویا برای سری زمانی S_p

نتایج به دست آمده مربوط به سری زمانی S_p بر اساس مدل‌های مختلف واریانس می‌باشد. منحنی‌های تاثیر اخبار مدل‌های نامتقارن گارچ پاسخ‌های نوسانی بلندتری را نسبت به بازده‌های منفی به جا می‌گذارند به طوری که اگر r_{t-1} یک بازدهی مثبت در زمان $t - 1$ باشد خواهیم داشت

$$\sigma_t^2(-r_{t-1}) > \sigma_t^2(r_{t-1}) \quad [۴۸].$$

در بین منحنی‌های NI مدلی که کشیدگی بلندتری را در قسمت بازده‌های منفی از خود نشان می‌دهد عملکرد بهتری در ثبت نوسانات فرایند و بیان پدیده‌ی اثر اهرم دارد. از این منحنی‌ها این چنین استنباط می‌شود که

در بین مدل‌های تحت مقایسه، مدل گارچ و مدل گارچ گویا پاسخ‌های مشابهی را در فرایند نوسانات شرطی دارند و هر دو مدل اثرات نامتقارن اهرم را در بازده‌های منفی به‌خوبی نشان می‌دهند و رتبه‌های اول و دوم را در بین سایر مدل‌های نامتقارن خانواده گارچ دارند.

۸-۳ نتیجه‌گیری

در این پایان‌نامه، مدل جدیدی از خانواده‌ی گارچ نامتقارن براساس اعمال تقریب گویا بر روی مدل گارچ ارائه شد و از طریق تکنیک برآورد بیزی، پارامترهای معادله‌ی واریانس ناهمسان شرطی به‌دست آمد که برای آن از روش شبیه‌سازی الگوریتم متروپلیس-هستینگز از روش‌های زنجیر مارکف مونت کارلو کمک گرفته شده است. با در نظر گرفتن توزیع خطای نرمال برای جملات خطای هر مدل، مدل جدید براساس معیار انحراف اطلاعات و منحنی‌های تأثیر اخبار با سایر تعمیم‌های خانواده‌ی گارچ مقایسه شد و کارایی بهتری را در ثبت نوسانات فرایند و پیش‌بینی تلاطم آینده دارا بود.

۹-۳ پیشنهاد برای ادامه‌ی کار

- (۱) به‌کارگیری تقریب گویا برای مدل‌های ایستای میانگین برای بهبود کارایی مدل‌ها و کاهش خطای تقریب.
- (۲) بررسی مدل جدیدی از گارچ گویا با عنوان مدل گارچ گویای نمایی به‌منظور برطرف کردن مشکل منفی بودن ضرایب در معادله‌ی واریانس.
- (۳) برآورد پارامترهای مدل با روش‌های دیگر هم‌چون MLE .

دستورات نرم افزار R

(*i*) دستور اجرای الگوریتم متروپلیس برای تابع هدف توزیع تی و به دست آوردن پارامترهای مربوط به آن

```
S <- matrix(c(0.1,0,0,0,0,0.2,0,0,0,0,0.3,0,0,0,0,0.4), 4,4)
Mu <- c(0.1,0.4,0.2,0.3)
Df <- 10
N <- 1000
Teta <- c(0.1,1.4,0.2,0.3)
Target <- function(x){dmvt(Teta, Mu, S, Df, Log = FALSE)}
rwmetro <- function(Target, N, Teta, S, burn in = 0)
{
  require(MASS)
  sample <- Teta
  for(i in 2: (burn in + N))
  {
    prop <- mvrnorm(n = 1, Teta, S)
    proposal <- function(y){dmvn(y, Teta, S, Log = FALSE)}
    if(runif(1) < min(1, Target(prop) * proposal(Teta)/Target(Teta) *
    proposal(prop)))
    Teta <- prop
  }
  sample <- rbind(sample, Teta)
}
Samples[(burn in + 1): (N + burn in)]
Ringsample4D <- rwmetro(TargetN, Teta, S)
a <- Ringsample4D[, 1]
b <- Ringsample4D[, 2]
c <- Ringsample4D[, 3]
d <- Ringsample4D[, 4]
```

```

Mu1 < -(sum(a[,1]))/1000
Mu2 < -(sum(a[,2]))/1000
Mu3 < -(sum(a[,3]))/1000
Mu4 < -(sum(a[,4]))/1000
Mu < -c(Mu1, Mu2, Mu3, Mu4)
S < -(Df - 2) /Df * matrix(c(var(a), cov(a, b), cov(a, c), cov(a, d),
cov(b, a), var(b), cov(b, c), cov(b, d), cov(c, a), cov(c, b), var(c), cov(c, d), cov(d, a),
cov(d, b), cov(d, c), var(d)),4,4)

```

(ii) دستور مربوط به برآورد پارامترهای مدل گارچ گویا از توزیع پسین با استفاده از الگوریتم متروپلیس-هستینگز براساس داده‌های سری زمانی S_1

```

Read(S1 ibm)
Z < -c(S1 ibm)
for(i in 2: 255){
r < -100 * (Log(Z[i]) - Log(Z[i - 1]))
}
y < -c(r^2)
sigma < -c(1: 254)
L < -c(1: 254)
mu < -c(0.03997452,0.46113780,0.34739600,0.1959700)
S < -matrix(c(0.06575159, -0.01094036,0.01735915,0.03527002,
-0.01094036,0.16921456,0.01661529,0.01430978,0.01735915,0.01661529,
0.26095376,0.09571424,0.03527002,0.01430978,0.09571424,0.6383614),4,4)
Df < -10
u < -1e-300
proposal < -function(p){dmvt(p, mu, S, Df, Log = FALSE)}
Teta < -c(0.02924496,0.5547858,0.9304921,0.6659138)
M < -3.023627e-107
current = Teta
for(i in 1: 55000){

```

```

proposed = c((rmvt(n = 1, mu, S, Df)))
a < -proposed[1]
b < -proposed[2]
c < -proposed[3]
d < -proposed[4]
for(i in 2:254){
A[i] = (c + a * y[i - 1] + b * A[i - 1])/1 + d * r[i - 1]
if(A[i] <= 0){
A[i] = 0.05
}}
for(i in 1:254){
L[i] < -dnorm(r[i], mean = 0, sd = sqrt(A[i]), Log = FALSE)
if(prod(L) >= u){
rho = (prod(L) * proposal(current))/(M * proposal(proposed))
if(runif(1) < min(1, rho))
print(proposed)
M < -prod(L)
current = proposed
}}

```

(iii) دستور مربوط به رسم منحنی‌های تأثیر اخبار برای هر مدل براساس اطلاعات سری‌زمانی S_2

```

Read(S2 s&p 500)
Z < -c(S2 s&p 500)
P < -c(1:387)
for(i in 2:387){
P[i] < -100 * (Log(Z[i]) - Log(Z[i - 1]))
r < -c(1:386)
for(i in 1:386){
r[i] < -Z[i + 1]
}
}

```


$Y < -c(r^2)$

RGARCH

$X < -c(0.00214076, 0.2798755, 0.02687707, -0.05059009)$

$Alfa < -X[1]$

$Beta < -X[2]$

$Omega < -X[3]$

$Delta < -X[4]$

$V < -var(Z)$

$Sigma < -function(r)\{$

$Sigma = (Omega + Beta * V + Alfa * Y)/1 + Delta * r$

$\}$

$Curve(Sigma, -15, 15)$

GARCH

$X < -c(0.177931, 0.702899, 0.001286)$

$Alfa < -X[1]$

$Beta < -X[2]$

$Omega < -X[3]$

$Sigma < -function(r)\{$

$Sigma = Omega + Alfa * Y + Beta * V$

$\}$

$Curve(Sigma, -15, 15)$

GJR – GARCH

$X < -c(0.158703, 0.694872, 0.001086, 0.152400)$

$Alfa < -X[1]$

$Beta < -X[2]$

$Omega < -X[3]$

$Ro < -X[4]$

$Sigma < -function(r)\{$

$if(r >= 0)\{$

$Sigma = Omega + Beta * V + Alfa * Y$

```

}else{
Sigma = Omega + Beta * V + (Alfa + Ro) * Y
}}
Curve(Sigma, -15,15)

```

EGARCH

```

X <- c(0.869362, -0.906118, 0.392573, -0.078519)
Beta <- -X[1]
Omega <- -X[2]
Ro <- -X[3]
Teta <- -X[4]
A <- -V^Beta * exp(Omega - Ro * sqrt(2/pi))
Sigma <- function(r){
if(r >= 0){
Sigma <- -A * exp(((Teta + Ro)/sqrt(V)) * r)
}else{
Sigma <- -A * exp(((Teta - Ro)/sqrt(V)) * r)
}}
Curve(Sigma, -15,15)

```

منابع و مأخذ

- [1] G. Žitković, Introduction to Stochastic Processes- Lecture Notes, Department of Mathematics, The University of Texas at Austin, Last Updated: December 24, 2010.
- [2] O. Calin, An Introduction to Stochastic Calculus with Applications to Finance, Department of Mathematics, Eastern Michigan University, Ypsilanti, MI 48197 USA, ocalin@emich.edu, October 2012.
- [3] Paul G. Hoel, Sidney C. Port, Charles J. Stone, Introduction to Stochastic Processes, Waveland Press, 1986, ISBN: 1478608994, 9781478608998.
- [4] Andres M. Alonso, Carolina G. Martos, Time Series Analysis, Autoregressive MA and ARMA Processes, Universidad Carlos III de Madrid, June- July, 2012.
- [5] Z. Ivanovski, A. Milenkovski, Z. Narasanov, Time Series Forecasting Using a Moving Average Model for Extrapolation of Number of Tourist, January 2018.
- [6] S. Gordon, The Normal Distribution, Mathematics Learning Centre, University of Sydney, NSW 2006.
- [7] M. I. Ribeiro, Gaussian Probability Density Functions: Properties and Error Characterization, Institute for Systems and Robotics, Instituto Superior Tcnico, February 2004, mir@isr.ist.utl.pt.
- [8] Ayienda K. Carolynne, Gamma and Related Distributions, School of Mathematics, University of Nairobi, November, 2013.
- [9] B. M. Golam Kibria, A. Joarder, A Short Review of Multivariate T-Distribution, Journal of Statistical Research 2006, Vol. 40, No. 1, pp. 59-72 Bangladesh, ISSN 0256 - 422 X.
- [10] M. Roth, On the Multivariate t Distribution, Department of Electrical Engineering, 17th April 2013, <http://www.control.isy.liu.se>.
- [11] G. Sobczyk, Generalized Taylor's Theorem, Departamento de Actuaría, Física y Matemáticas, 72820 Cholula, Mexico, February 2, 2008.
- [12] A.A. Christie, The stochastic behavior of common stock variances: Value, leverage and interest rate effects, J. Finan. Econ. 10 (4) (1982) 407-432.

- [13] F. Black, Studies of stock market volatility changes, in: 1976 Proceedings of the American Statistical Association, Business and Economic Statistics, Section, 1976, pp. 177–181.
- [14] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*: J. Econometr. Soc. (1982) 987–1007.
- [15] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, *J. Econometrics* 31 (1986) 307–327.
- [16] Dennis S. Mapa, Adolfo M. de Guzman, The Generalized AutoRegressive Conditional Heteroskedasticity Parkinson Range (GARCH-PARK-R) Model for Forecasting Financial Volatility, School of Statistics, University of the Philippines, October 4-5, 2004.
- [17] D. Nelson, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59 (1991) 347–370.
- [18] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung, *Time Series Analysis.: Forecasting and Control*, Wiley Series in Probability and Statistics, Wiley, 2015, ISBN: 1118674928, 9781118674925.
- [19] T. Bollerslev, A conditionally heteroskedastic time series model for speculative prices and rates of return, *Rev. Econ. Stat.* (1987) 542–547.
- [20] Y. Wang, D. Zhang, Risk and Loss Functions, Columbia University, Friday, February 6, 2004.
- [21] George E. P. Box, George C. Tiao, *Bayesian Inference in Statistical Analysis*, Volume 40 of Wiley Classics Library, John Wiley & Sons, 2011, ISBN: 111803144X, 9781118031445.
- [22] W. K. Hastings, Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrika*, Volume 57, Issue 1, April 1970, Pages 97109, <https://doi.org/10.1093/biomet/57.1.97>.
- [23] Dani Gamerman, Hedibert F. Lopes, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Second Edition, Chapman & Hall/CRC Texts in Statistical Science, CRC Press, 2006, ISBN: 1584885874, 9781584885870.
- [24] M. Asai, Comparison of mcmc methods for estimating GARCH models, *J. Japan Statist. Soc.* 36 (2) (2006) 199–212.

- [25] I. Yildirim, Bayesian Inference: Metropolis-Hastings Sampling, Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, August 2012.
- [26] S. Chib, E. Greenberg, Understanding the Metropolis-Hastings Algorithm, the American statistical association, November 1995, Vol. 49. No. 4.
- [27] Leonhard Held, Introducing Bayes Factors, Division of Biostatistics, University of Zurich, 25 November 2011.
- [28] Aitkin, M. (1991). Posterior Bayes factors (with discussion). *Journal of the Royal Statistical Society, Series B*, 53: 111-142.
- [29] Robert L. Wolpert, Fisher Information & Efficiency, Department of Statistical Science, Duke University, Durham, NC, USA.
- [30] Jay I. Myung, Daniel J. Navarro, Information Matrix, Department of Psychology, Ohio State University, 1827 Neil Avenue, Columbus OH 43210, USA, {myung.1, navarro.20}@osu.edu.
- [31] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: 2nd International Symposium on Information Theory, 1973, pp. 267–281.
- [32] Cavanaugh, J. E. (1997), "Unifying the derivations of the Akaike and corrected Akaike information criteria", *Statistics & Probability Letters*, 31: 201–208, doi:10.1016/s0167-7152(96)00128-9.
- [33] Anna-Lena Popkes, Kullback-Leibler Divergence, February 2, 2019.
- [34] Eric F. Lock, Deviance Information Criterion: Bayes Decision Theory and Data Analysis, UMN Division of Biostatistics, SPH elock@umn.edu, 04/29/2019.
- [35] Spiegelhalter DJ, Best NG, Carlin BP and Van der Linde A, "Bayesian Measures of Model Complexity and Fit (with Discussion)", *Journal of the Royal Statistical Society, Series B*, 2002 64(4):583-616.
- [36] J. Nuyts, I. Platten, Phenomenology of the term structure of interest rates with padé Approximants, *Physica A* 299 (2001) 528–546.
- [37] P.C. Phillips, Best uniform and modified padé approximants to probability densities in econometrics, *Adv. Econom.* 51 (1982) 1–24.

- [38] K. Kumar, Padé Approximation And its Application in Time Series Analysis, Applied Mathematics and Computation, volume 48, Issues2-3, April 1992, Pages 139-151.
- [39] Z. Kalateh Bojdi, S. Ahmadi-Asl, A. Aminataei, A New Extended Padé Approximation and Its Application, Advances in Numerical Analysis, Volume 2013, Article ID 263467, 8 pages, <http://dx.doi.org/10.1155/2013>.
- [40] T. Takaishi, Rational GARCH model: An empirical test for stock returns, Hiroshima University of Economics, Hiroshima 731-0192, Physica A 473 (2017) 451–460.
- [41] T.T. Chen, T. Takaishi, Empirical study of the GARCH model with rational errors, in: Journal of Physics: Conference Series, Vol. 454, IOP Publishing, 2013, p. 012040.
- [42] T. Takaishi, Bayesian estimation of GARCH model with an adaptive proposal density, New Advances in Intelligent Decision Technologies, Studies in Computational Intelligence 199 (2009) 635–643.
- [43] T. Takaishi, Bayesian inference on QGARCH model using the adaptive construction scheme, in: Proceedings of 8th IEEE/ACIS International Conference on Computer and Information Science, 2009, pp. 525–529. <http://dx.doi.org/10.1109/ICIS.2009.173>.
- [44] T. Takaishi, Bayesian inference with an adaptive proposal density for GARCH models, in: Journal of Physics: Conference Series, Vol. 221, IOP Publishing, 2010, p. 012011.
- [45] T. Takaishi, T.T. Chen, Bayesian inference of the GARCH model with rational errors, Internat. Proc. Econ. Development Res. 29 (2012) 303–307.
- [46] P. Burns, Robustness of the Ljung-Box Test and its Rank Equivalent, This Draft: 6 October 2002, <http://www.burns-stat.com>.
- [47] R. Engle, V. Ng, Measuring and testing the impact of news on volatility, J. Finance 48 (1993) 1749–1778.
- [48] M. Takahashi, Y. Omori, T. Watanabe, News Impact Curve for Stochastic Volatility Models, , Volume 120, Issue 1, July 2013, pages130-134
- [49] <https://www.datamarket.com>.

واژه‌نامه فارسی به انگلیسی

Probability	احتمال
Experiment	آزمایش
Decrease	افزایش
Entropy	آنتروپی
Transfer	انتقال
Deviance	انحراف
Estimator	برآوردگر
Scale Parameter.....	پارامتر مقیاس
Location Parameter.....	پارامتر مکان
Acceptance.....	پذیرش
Posterior	پسین
Complexity.....	پیچیدگی
prior	پیشین
Continuous	پیوسته
Likelihood Function	تابع درستنمایی
Risk Function.....	تابع زیان
Gamma Function	تابع گاما
Character Function.....	تابع مشخصه
Loss Function	تابع مخاطره
Moment-Generating Function	تابع مولد گشتاور
Decision	تصمیم
Generalize.....	تعمیم

Univariate	تک متغیره
Distribution	توزیع
Impact	تأثیر
Density	چگالی
Property	خاصیت
Partial Autocorrelation	خودهمبستگی جزئی
Cluster	خوشه
Autocorrelation	خودهمبستگی
Degree of Freedom	درجه آزادی
Period	دوره
Simulation	شبیه‌سازی
Conditional	شرطی
Time Series	سری زمانی
Stock	سهام
Factor	عامل
Backshift Operator	عملگر پسرو
Random Process	فرایند تصادفی
Burn-in Process	فرایند سوخته
Stationary Process	فرایند مانا
Hypothesis	فرضیه
State Space	فضای حالت
Performance	کارایی
Kurtosis	کشیدگی
Quantity	کمیت
Discrete	گسسته

Rational.....	گویا
Strictly Stationary	مانای اکید
Weakly Stationary	مانای ضعیف
Variable.....	متغیر
Independent Random Variable.....	متغیر تصادفی مستقل
Convex	محدب
Finite.....	محدود
Limitation	محدودیت
Composite.....	مرکب
Derivative	مشتق
Independent	مستقل
Expected Value	مقدار موردانتظار
Mean	میانگین
Asymmetry	نامتقارن
Heteroskedasticity	ناهمسان
Volatility	نوسان
Convergence	همگرایی

Abstract:

A new model of asymmetric GARCH, rational GARCH and an empirical test for stock return

Amin Mansouri Kohestani

In this thesis, a new model of the GARCH family models is defined as the rational GARCH model with normal error sentences. Based on the Bayesian theorem and the Bayesian parameter estimation, the parameters of the new model are obtained with the other models under comparison for which the Metropolis-Hastings algorithm is assisted by the Markov chain Monte Carlo Methods. By using the deviance information criterion and the news impact curves, the efficiency of each model is examined based on which each model's ranking is used to record the random process fluctuation and predict future turbulence.

Key words: Rational GARCH model, Asymmetric GARCH models, Leverage effect, Rational approximation, Deviance information criterion, News impact curve



University
of
Guilan

Sector of Educational and Research Studies

Declaration of the Thesis/Dissertation Originality

Hereby, I declare that this thesis, entitled..... **A new model of asymmetric GARCH, rational GARCH and an empirical test for stock return.....**

Which has been defended for an M.A. /Ph.D. degree in **Mathematical Statistics.....** in.....2019/03/10.....

is the result of my own research except as cited in the references . The Thesis has not been submitted in candidature of any other degree inside or outside of Iran.

I also declare that my making use of the thesis, in any type and aspect, will be under the permission of my supervisor and educational department; Otherwise, I will be punished legally as far as annulling my academic degree.

Full name: Amin Mansouri Kohestani

Signature: 

Date: 1398.04.05



University of Guilan

In The Name of God

Proceeding for Master of Science Dissertation

Defense session for granting M.Sc. dissertation on award to Mr Amin Mansouri Kohestani majoring in Statistics with minor in Mathematical Statistics title of dissertation A New Model of Asymmetric GARCH, Rational GARCH And An Empirical Test for Stock Returns in Guilan province with 6 hours credit, held on Mar. 10, 2019 at 17:30, in Faculty of Mathematical Sciences, University of Guilan. The juries announced the following decision:

- The dissertation is accepted with score ... of 20 ranged as Excellent, Very good, Good, Acceptable.
The dissertation is acceptable with minor corrections with score 17.5 of 20 ranged as Excellent, Very good, Good, Acceptable.
The dissertation with present status did not accepted and suggested that.....

Table with 5 columns: Juries' members, Ranks, Specialization, Affiliation, Signature. Rows include Major Advisors (Dr. Hossein Samimi Haghgozar), Minor Advisor, and Examiners (Dr. Behrouz Fathi, Dr. Farshid Mehrdoust).

Table with 4 columns: Representative of College Graduate Office, Rank, Affiliation, Signature. Row includes Dr. Maryam Seifaddini, Assistant Professor, Computer Science.

After completing of all documents four copies of the proceeding have to be submitted and handed to the Director of College Graduate Office. From four copies, one will be kept in the student department, one in Graduate Office, one in the file of student and the last hands to the student.



University Of Guilan

Department of Statistics
Mathematical Statistics

**A new model of asymmetric GARCH, rational GARCH
and an empirical test for stock return**

By

Amin Mansouri Koohestani

Supervisor

Dr. Hossein Samimi Haghgozar

September 2019

