ریاضی و جامعه
شاپا (چاپی): ۶۴۹۳-۲۳۴۵، شاپا(الکترونیکی): ۶۵۰۷-۲۳۴۵
جلد۲ ، شماره ۴ (۱۳۹۶)، صص. ۱۰-۱
© ۱۳۹۶ دانشگاه اصفهان

http://math-sci.iranjournals.ir
http://www.ui.ac.ir

برآورد کمترین توان‌های دوم پارامترهای مدل آرچ در حضور داده‌های گم‌شده

محمدرضا صالحی راد، رضا حبیبی* و مریم سادات غیبی

چکیده. بسیاری از سری‌های زمانی مالی و اقتصادی در دوره‌هایی با نوسانات زیادی همراه هستند و متعاقب آن در دوره‌هایی نوسان‌پذیری (واریانس) کمی دارند یعنی نوسان‌پذیری خوشه‌ای است. در این موقعیت فرض وجود واریانس ثابت معقول نیست. یکی از ابزارهای مدل‌بندی تغییرات، استفاده از مدل‌های واریانس ناهمسان شرطی آرچ و گارچ است. مدل آرچ در زمینه‌های اقتصادسنجی و مالی از جمله در بررسی تغییرپذیری مربوط به شاخص بورس، قیمت طلا و نرخ ارز کاربردهای فراوانی دارد.یکی از روش‌های رایج برآورد پارامترهای مدل آرچ روش شبه ماکسیمم درست‌نمایی است. از آنجایی‌که این برآوردگر فرم بسته‌ای ندارد و برآوردگرهای دیگر نظیر برآوردگر کمترین توان‌های دوم کارایی بالاتری دارد برای برآورد پارامترهای مدل آرچ از برآوردگر کم‌ترین توان‌های دوم استفاده می‌شود. وجود داده‌های گم‌شده در سری‌های زمانی امری معمول است. از این‌رو برآوردگر کمترین توان‌های دوم دو مرحله‌ای را در حضور داده‌های گم‌شده به دست می‌آوریم. این برآوردگر کارایی مجانبی یکسانی با برآوردگر شبه ماکسیمم درست‌نمایی دارد. سازگاری قوی و توزیع مجانبی نرمال این برآوردگر را اثبات و سپس این نتایج را با استفاده از شبیه‌سازی تأیید می‌کنیم. نتایج را بر روی داده‌های شاخص کل بورس تهران به کار برده می‌شود.

۱. مقدمه

تحلیل‌های سری‌های زمانی اغلب بر پایه ایستایی سری زمانی و در حالت خاص، فرض همسانی واریانس‌ها بنا شده‌اند. این فرض در بسیاری از داده‌های سری‌های زمانی از جمله داده‌های اقتصادی ممکن است برقرار نباشد. برای رفع ناهمسانی، روش معمول بر اساس تبدیل داده‌ها به نحوی است که همسانی واریانس‌ها حاصل شود. البته بهتر آن است که از مدل‌هایی استفاده شود که شرط ناهمسانی را در برازش مدل در نظر بگیرد. یک خانواده معروف از این مدل‌ها، خانواده مدل‌های اتورگرسیو واریانس ناهمسان شرطی (آرچ)[1] است. روش مرسوم برای برآورد پارامترهای مدل آرچ، استفاده از تابع شبه درست‌نمایی می‌باشد که لگاریتم تابع شبه درست‌نمایی را ماکسیمم می‌کند. ویس [۵] سازگاری و توزیع مجانبی نرمال برآوردگر شبه ماکسیمم درست‌نمایی[2] را اثبات کرد. با این حال، این برآوردگر را نمی‌توان به صورت یک فرم بسته بیان کرد، یعنی نمی‌توان آن را بر حسب یک ضابطه‌ی صریح نمایش داد. به همین دلیل، باید از روش‌های دیگری، برای مثال روش‌های عددی کمک بگیریم که محاسبه‌ی آن آسان نیست [۴]. بوز و موخرجی [۲] برای برطرف کردن این مشکلات، پیشنهاد استفاده از برآوردگر حداقل مربعات به روش حل معادلات خطی[3]را دادند. به‌دست آوردن این برآوردگر نیاز به حل دو مجموعه از معادلات خطی دارد لذا به دست آوردن آن آسان است. این برآوردگر از نظر دقت و صحت با برآوردگر شبه ماکسیمم درست‌نمایی قابل مقایسه می‌باشد. این برآوردگر خطی دارای واریانس کمتری در مقایسه با برآوردگر شبه ماکسیمم درست‌نمایی است. همچنین روش برآوردگر

عبارات و کلمات کلیدی. مدل‌های آرچ و گارچ، مشاهدات گم‌شده، ناهمسانی شرطی، برآوردگر کم‌ترین توان‌های دوم، قضیه حد مرکزی مارتینگل.
*نویسنده مسئول
تاریخ دریافت: ۱۳۹۳/۰۷/۰۹ تاریخ پذیرش: ۱۳۹۵/۱۰/۱۶ .

[1] Autoregressive Conditional Heteroscedastic (ARCH) [2] Quasi Maximum Likelihood Estimator (QMLE) [3] Linear Equations(LE)

خطی حتی برای نمونه‌هایی با اندازه کوچک (مثلاً اندازه نمونه ۳۰) نتیجه بهتری نسبت به روش شبه ماکسیمم درست‌نمایی دارد. $P_n(z) = 0$،

۲. برآوردگر کم‌ترین توان‌های دوم در حضور داده‌های گم‌شده

فرض می‌کنیم که $(\Omega, \mathcal{F}, P)$ فضای احتمال روی متغیر X_t است و X_t را سری زمانی $ARCH(p)$ تعریف می‌کنیم به طوری‌که X_t در معادله‌ی

$$X_t = \sigma_t(\boldsymbol{\alpha})\epsilon_t \tag{۱.۲}$$

صدق می‌کند. در این معادله ϵ_t دنباله‌ای از متغیرهای تصادفی مستقل با $E(\epsilon_t) = 0$ و $E(\epsilon_t^2) = 1$ است و σ_t یک فرایند نامنفی با معادله تصادفی است که به صورت زیر تعریف می‌شود

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 \tag{۲.۲}$$

که $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_p]'$ پارامترهای نامعلوم با شرط $\alpha_0 > 0$ و برای $i = 1, 2, \ldots, p$ ، $\alpha_i \geq 0$ می‌باشد . حال فرض می‌کنیم در سری زمانی مورد بررسی داده گم شده وجود دارد. برای این منظور، سری زمانی با مشاهدات گم‌شده را منطبق بر کار پارزن (۱۹۶۳) به صورت سری زمانی تعدیل‌شده‌ای از داده‌های اصلی در نظر می‌گیریم. یعنی

$$X_t^* = a_t X_t \tag{۳.۲}$$

به طوری‌که X_t برای هر t تعریف شده و مقدار دنباله دودویی a_t برابر با یک است اگر X_t مشاهده شده باشد و در غیر این صورت مقدار آن صفر است. در واقع زمانی که X_t گم‌شده نیست X_t^* مقدار واقعی مشاهدات X_t را نشان می‌دهد و زمانی که X_t گم‌شده است مقدار صفر جایگذاری می‌شود. دو فرض زیر را در مورد a_t در نظر می‌گیریم

الف) فرآیند $\{a_t\}$ مانای اکید و آمیخته ضعیف است،

ب) فرآیندهای $\{a_t\}$ و $\{X_t\}$ مستقل هستند.

در ادامه برخی نمادهای مورد نیاز مقاله را معرفی می‌کنیم. فرض کنیم

$$A_t = \prod_{i=0}^{p} a_{t-i}.$$

و قرار دهید

$$Y_t = X_t^2, \quad \boldsymbol{Z}_t = [1, Y_{t-1}, \ldots, Y_{t-p}]', \quad \eta_t = \epsilon_t^2 - 1$$

و لذا می‌توان دید که

$$\begin{aligned} \sigma_t^2(\boldsymbol{\alpha}) &= \boldsymbol{Z}_t'\boldsymbol{\alpha} \\ Y_t &= \boldsymbol{Z}_t'\boldsymbol{\alpha} + \sigma_t^2(\boldsymbol{\alpha})\eta_t \end{aligned} \tag{۴.۲}$$

که در آن

$$E(\sigma_t^2(\boldsymbol{\alpha})\eta_t \mid \mathcal{F}_{t-1}^X) = \sigma_t^2(\boldsymbol{\alpha})E(\eta_t \mid \mathcal{F}_{t-1}^X) = 0$$

و $\mathcal{F}_{t-1}^X$ سیگما بر تولیدشده توسط $(X_j)_{j\leq t}$ است. رابطه فوق نشان می‌دهد Y_t از مدل اتورگرسیو با مرتبه‌ی p که عامل ابداع یک دنباله تفاضل مارتینگل است پیروی می‌کند. نکته حائز اهمیت این است که عامل ابداع یک ناهمسان شرطی است. با اندکی عملیات جبری می‌توان دید که

$$Y_t^* = A_t Y_t = A_t \boldsymbol{Z}_t'\boldsymbol{\alpha} + A_t\sigma_t^2(\boldsymbol{\alpha})\eta_t = \boldsymbol{Z}_t^{*\prime}\boldsymbol{\alpha} + \sigma^{*2}(\boldsymbol{\alpha})\eta_t \qquad p+1 \leq t \leq n$$

که در آن

$$\boldsymbol{Z}_t^{*'} = A_t \boldsymbol{Z}_t'$$
$$\sigma^{*^2}(\boldsymbol{\alpha}) = A_t \sigma_t^2(\boldsymbol{\alpha})$$

حال با صرف نظر کردن از تصادفی بودن $\sigma^{*^2}(\boldsymbol{\alpha})$ و نیز حضور $\boldsymbol{\alpha}$ در آن، برآوردگر کم‌ترین توان‌های دوم مقدماتی $\boldsymbol{\alpha}$ می‌تواند به صورت زیر به دست آید.

$$\hat{\boldsymbol{\alpha}}_{Pr} = \Big[\sum_{t=p+1}^{n} \boldsymbol{Z}_t^* \boldsymbol{Z}_t^{*'}\Big]^{-1} \sum_{t=p+1}^{n} \boldsymbol{Z}_t^* Y_t^* \tag{۵.۲}$$

در ادامه منطبق بر مقاله بوز و موخرجی از $\hat{\boldsymbol{\alpha}}_{Pr}$ برای ساختن برآوردگر بهبود یافته $\hat{\boldsymbol{\alpha}}$ استفاده می‌کنیم. با تقسیم طرفین رابطه‌ی

$$Y_t^* = \boldsymbol{Z}_t^{*'} \boldsymbol{\alpha} + \sigma^{*^2}(\boldsymbol{\alpha}) \eta_t$$

بر $\sigma_t^2(\boldsymbol{\alpha})$ ، معادله‌ی

$$\frac{Y_t^*}{\sigma_t^2(\boldsymbol{\alpha})} = \frac{\boldsymbol{Z}_t^{*'}}{\sigma_t^2(\boldsymbol{\alpha})} \boldsymbol{\alpha} + A_t \eta_t$$

را به دست می‌آوریم. در این عبارت، خطاهای $A_t \eta_t$ ناهمسان هستند. با جایگذاری $\hat{\boldsymbol{\alpha}}_{Pr}$ به جای $\boldsymbol{\alpha}$ در $\sigma_t^2(\boldsymbol{\alpha})$ ، $\boldsymbol{\alpha}$ را با روش کم‌ترین توان‌های دوم معمولی با نماد $\hat{\boldsymbol{\alpha}}$ به صورت زیر برآورد می‌کنیم

$$\hat{\boldsymbol{\alpha}} = \Big[\sum_{t=p+1}^{n} \frac{\boldsymbol{Z}_t^* \boldsymbol{Z}_t^{*'}}{\sigma_t^4(\hat{\boldsymbol{\alpha}}_{Pr})}\Big]^{-1} \sum_{t=p+1}^{n} \frac{\boldsymbol{Z}_t^* Y_t^*}{\sigma_t^4(\hat{\boldsymbol{\alpha}}_{Pr})}. \tag{۶.۲}$$

۳. سازگاری و توزیع نرمال مجانبی برآوردگرها

لم ۱.۳. *فرض کنید X_t فرآیند $ARCH(p)$ تعریف شده در بخش ۲ باشد و a_t در شرایط الف و ب بخش ۲ صدق می‌کند. همچنین فرض می‌کنیم $A_0 \neq 0$ ، بنابراین زمانی‌که $n \to \infty$ داریم*
۱) هرگاه $EX_0^4 < \infty$ باشد، آن‌گاه $\hat{\boldsymbol{\alpha}}_{Pr} \xrightarrow{a.s} \boldsymbol{\alpha}$.
۲) هرگاه $EX_0^8 < \infty$ باشد، آن‌گاه

$$n^{1/2}(\hat{\boldsymbol{\alpha}}_{Pr} - \boldsymbol{\alpha}) \xrightarrow{d} N\Big[0, (EA_0)^{-1} Var(\epsilon_0^2)\{E\boldsymbol{Z}_0\boldsymbol{Z}_0'\}^{-1} E\{(\boldsymbol{\alpha}'\boldsymbol{Z}_0)^2 \boldsymbol{Z}_0\boldsymbol{Z}_0'\}\{E\boldsymbol{Z}_0\boldsymbol{Z}_0'\}^{-1}\Big] \tag{۱.۳}$$

قضیه زیر سازگاری و توزیع نرمال مجانبی $\hat{\boldsymbol{\alpha}}$ را نشان می‌دهد.

قضیه ۲.۳. *فرض کنید X_t فرآیند $ARCH(p)$ است و a_t نیز در مجموعه‌ی $\{0, 1\}$ تعریف‌شده در الف و ب صدق می‌کند همچنین $A_0 \neq 0$. وقتی $n \to \infty$ ، داریم*
۱) هرگاه $EX_0^6 < \infty$ ، آن‌گاه $\hat{\boldsymbol{\alpha}} \xrightarrow{a.s.} \boldsymbol{\alpha}$.
۲) هرگاه $EX_0^8 < \infty$ ، آن‌گاه

$$n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \xrightarrow{d} N\Big(0, (EA_0)^{-1} Var(\epsilon_0^2)\Big\{E\{(\boldsymbol{\alpha}'\boldsymbol{Z}_0)^{-2} \boldsymbol{Z}_0\boldsymbol{Z}_0'\}\Big\}^{-1}\Big)$$

در صورتی که داده‌ها فاقد داده‌ی گم‌شده باشند. توزیع مجانبی برآوردگر مقدماتی و برآوردگر اصلی به صورت زیر به دست می‌آیند:
فرض می‌کنیم شرط گشتاوری زیر به ازای $1 \leq j,k,l,m \leq p$

$$E\{Y_jY_kY_lY_m\} < \infty \tag{۲.۳}$$

برقرار است. در این صورت

$$n^{1/2}(\hat{\boldsymbol{\alpha}}_{Pr} - \boldsymbol{\alpha}) \xrightarrow{d} N\Big(0, \sum\Big) \tag{۳.۳}$$

که در آن

$$\sum = Var(\epsilon_1^2)\{E(\boldsymbol{Z}_0\boldsymbol{Z}_0')\}^{-1}E\{(\boldsymbol{\alpha}'\boldsymbol{Z}_0)^2\boldsymbol{Z}_0\boldsymbol{Z}_0'\}\{E(\boldsymbol{Z}_0\boldsymbol{Z}_0')\}^{-1}.$$

قضیه ۳.۳. *فرض می‌کنیم*

$$n^{1/2}(\hat{\boldsymbol{\alpha}}_{Pr} - \boldsymbol{\alpha}) = O_p(1), \tag{۴.۳}$$

و دنباله‌ی متغیرهای تصادفی $\{X_t; 1-p \leq t \leq 0\}$ *به ازای هر* $1 \leq j,k,l \leq p$ *در ویژگی گشتاوری*

$$E\Big\{\frac{Y_{-j}Y_{-k}Y_{-l}}{(\boldsymbol{\alpha}'\boldsymbol{Z}_0)^3}\Big\} < \infty \tag{۵.۳}$$

صدق می‌کند. لذا

$$n^{1/2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \xrightarrow{d} N\Big(0, Var(\epsilon_1^2)\Big\{E\{\boldsymbol{Z}_0\boldsymbol{Z}_0'(\boldsymbol{\alpha}'\boldsymbol{Z}_0)^{-2}\}\Big\}^{-1}\Big). \tag{۶.۳}$$

۴. شبیه سازی

برآوردگر کم‌ترین توان‌های دوم دومرحله‌ای پارامترهای مدل آرچ در مجموع بهتر از برآوردگر شبه ماکسیمم درست‌نمایی عمل می‌کند. این امر چه بدون داده‌های گم‌شده و چه در حضور داده‌های گم‌شده حتی برای نمونه‌ای به اندازه ۳۰ برقرار است. مدل آرچ از مرتبه‌ی $p = 1$ با پارامترهای $\boldsymbol{\alpha} = [0/01, 0/8]'$ و خطاهای نرمال شبیه‌سازی کرده‌ایم. برای i امین آزمایه $(1 \leq i \leq 10)$ ، α را به طور تصادفی از توزیع χ^2 ، (چون بایستی مثبت باشد) و برای اینکه شرط $\alpha_1 < 1$ برقرار باشد α_1 را از توزیع $U(0,1)$ تولید کرده‌ایم. عمل شبیه‌سازی را به تعداد $k = 1000$ بار تکرار و در هر بار تکرار $\{X_t : 0 \leq t \leq n = 30\}$ را مشاهده می‌کنیم و بر پایه‌ی آن‌ها برآوردگرهای α_0 و α_1 را با استفاده از هر دو روش معادلات خطی و QML محاسبه می‌کنیم.
S_0 و S_1 را به ترتیب قرار می‌دهیم

$$S_0 = \sum_{k=1}^{1000} \frac{\Big(\beta_0 - \hat{\beta}_{0k}\Big)^2}{1000}$$

و

$$S_1 = \sum_{k=1}^{1000} \frac{\Big(\beta_1 - \hat{\beta}_{1k}\Big)^2}{1000}.$$

برآوردگرهای $\hat{\beta}_{0k}$ و $\hat{\beta}_{1k}$ با استفاده از روش معادلات خطی در k امین تکرار به دست آمده‌اند. S_{0Q} و S_{1Q} نیز مشابه بالا تعریف شده با این تفاوت که برآوردگرهای $\hat{\alpha}_{0k}$ و $\hat{\alpha}_{1k}$ به روش QML به دست آمده‌اند.
عمل شبیه‌سازی را به تعداد $k = 1000$ بار تکرار و در هر بار تکرار $\{X_t : 0 \leq t \leq n = 30\}$ را مشاهده می‌کنیم و بر پایه‌ی آن‌ها برآوردگرهای α_0 و α_1 را با استفاده از هر دو روش معادلات خطی و QML محاسبه می‌کنیم.

$S_۰$ و $S_۱$ را به ترتیب قرار می‌دهیم

$$S_۰ = \sum_{k=۱}^{۱۰۰۰} \frac{\left(\alpha_۰ - \hat{\alpha}_{۰k}\right)^۲}{۱۰۰۰}$$

و

$$S_۱ = \sum_{k=۱}^{۱۰۰۰} \frac{\left(\alpha_۱ - \hat{\alpha}_{۱k}\right)^۲}{۱۰۰۰}.$$

برآوردگرهای $\hat{\alpha}_{۰k}$ و $\hat{\alpha}_{۱k}$ با استفاده از روش معادلات خطی در k امین تکرار به دست آمده‌اند. $S_{۰Q}$ و $S_{۱Q}$ نیز مشابه بالا تعریف شده با این تفاوت که برآوردگرهای $\hat{\alpha}_{۰k}$ و $\hat{\alpha}_{۱k}$ به روش شبه روش ماکسیمم درستنمایی به دست آمده‌اند.

جدول ۱: مقادیر $S_۰$، $S_{۰Q}$، $S_۱$ و $S_{۱Q}$ بدون داده‌های گم‌شده

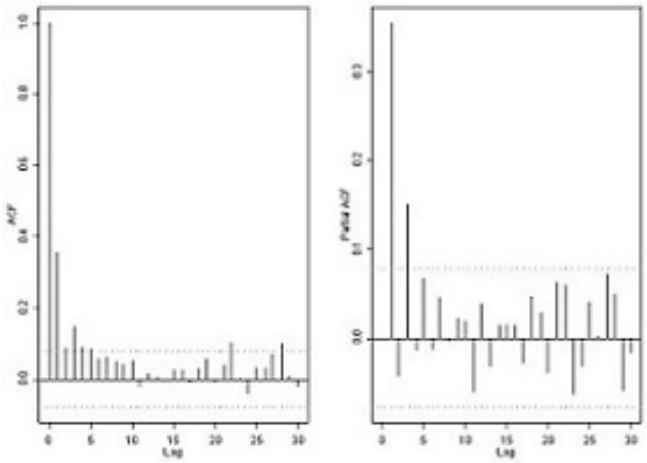
$Trial$	$\alpha_۰$	$\alpha_۱$	$S_۰$	$S_{۰Q}$	$S_۱$	$S_{۱Q}$
۱	۰/۱۳۹۱۰۶	۰/۵۳۲۴۹۶۲	۰/۰۱۵۶۹۳۱۱	۰/۱۶۴۰۵۳۲	۰/۱۷۰۲۵۰۶	۰/۵۱۶۴۰۶۸
۲	۰/۰۶۴۳۴۴۹۳	۰/۸۲۶۹۲۴۳	۰/۰۰۲۵۹۳۴۹۴	۰/۰۰۲۹۳۷۴۰۶	۰/۱۷۱۵۵۹۷	۰/۶۷۶۰۶۲۹
۳	۰/۰۱۷۵۴۶۵۷	۰/۹۸۷۴۲۶	۰/۰۰۰۰۷۱۴۴	۰/۰۰۰۰۷۶۶۲	۰/۳۲۶۶۶۵۱	۰/۶۸۳۷۳۴۱
۴	۰/۷۲۰۷۰۳۲	۰/۶۸۰۷۱۷۶	۰/۴۹۸۷۴۲۹	۰/۵۰۲۷۹۸۶	۰/۱۶۳۹۷۱۵	۰/۱۹۲۴۷۱۵
۵	۰/۲۱۱۱۶۰۷	۰/۱۹۱۰۷۸	۰/۰۳۸۱۷۳۷۶	۰/۳۹۹۸۸۸	۰/۲۸۲۲۷۳۴	۰/۵۴۱۵۹۲۴
۶	۰/۸۷۰۶۱۲۴	۰/۴۹۶۳۲۴۷	۰/۷۳۲۸۹۰۹	۰/۷۳۹۷۴۳۴	۰/۱۶۴۹۵۸۱	۰/۲۵۲۹۰۲۹
۷	۰/۰۲۳۰۴۴۴۱	۰/۷۲۷۳۵۳۹	۰/۰۰۰۱۵۰۰۲۴	۰/۰۰۰۱۷۹۴۹۷	۰/۲۱۳۹۵۲۶	۰/۷۴۱۵۹۸
۸	۰/۱۸۷۲۰۰۸	۰/۲۱۱۱۹۹۹	۰/۰۲۹۸۵۱۸۹	۰/۰۳۱۱۸۴۴۶	۰/۲۱۳۲۹۵	۰/۴۶۳۰۶۵۵
۹	۰/۱۳۷۷۴۹	۰/۱۹۵۸۲۷۷	۰/۰۱۵۵۳۱۳۷	۰/۰۱۶۲۲۷۴۵	۰/۲۲۵۱۱۱	۰/۴۶۵۳۸۵۵
۱۰	۰/۶۶۲۳۶۹۵	۰/۷۹۳۲۹۷۲	۰/۴۱۹۴۹۵۲	۰/۴۲۴۸۷۴۷	۰/۱۷۴۶۳۸۱	۰/۲۴۷۵۸۰۳

جدول ۲: مقادیر $S_۰$، $S_{۰Q}$، $S_۱$ و $S_{۱Q}$ در حضور داده‌های گم‌شده

$Trial$	$\alpha_۰$	$\alpha_۱$	$S_۰$	$S_{۰Q}$	$S_۱$	$S_{۱Q}$
۱	۰/۷۴۹۶۰۶	۰/۲۶۵۱۹۹	۰/۵۳۵۳۱۶	۰/۵۴۰۰۷۱۱	۰/۱۵۹۲۵۸	۰/۲۷۲۸۵۶
۲	۰/۰۰۱۱۱۸	۰/۷۷۲۵۴۴۱	۰/۰۰۰۴۹۸۷	۰/۰۰۰۲۹۵۳	۰/۳۰۱۲۰۸۸	۰/۵۷۲۱۸۹
۳	۰/۷۲۱۷۵۷۹	۰/۲۵۸۵۱۵۲	۰/۴۹۸۹۲۱۷	۰/۵۰۳۰۵۴۹	۰/۱۶۳۲۱۸۸	۰/۳۰۵۰۸۲۴
۴	۰/۰۲۰۸۹۱۱۴	۰/۹۰۲۰۰۷	۰/۰۰۰۱۱۱۵۴	۰/۰۰۰۱۲۶۵۶	۰/۳۵۰۳۲۸	۰/۵۰۶۴۵۷
۵	۰/۳۸۴۸۵۷	۰/۲۱۹۱۹۱	۰/۱۳۴۸۳۷	۰/۱۳۶۸۸۹	۰/۱۴۸۰۰۰۹	۰/۲۴۵۶۳۹۷
۶	۰/۰۲۴۱۱۷۷	۰/۶۶۹۴۵۴	۰/۰۰۰۱۶۸۶	۰/۰۰۰۲۰۸۴	۰/۲۲۰۲۰۲	۰/۶۴۲۸۸۷
۷	۰/۱۲۳۹۶۵۳	۰/۵۶۹۰۲۴	۰/۰۱۱۵۷۵۸	۰/۱۲۰۹۶۱۲	۰/۱۶۰۵۴۷۹	۰/۳۸۶۹۲۱
۸	۰/۲۱۲۱۵۸	۰/۶۹۱۴۶۷	۰/۰۳۷۹۱۶۹	۰/۳۹۵۱۵۲	۰/۲۳۸۰۰۶۵	۰/۳۷۴۱۸۶۵
۹	۰/۳۹۶۷۵۱۵	۰/۴۶۴۳۹۶	۰/۱۴۶۳۲۶	۰/۱۴۸۲۰۷	۰/۱۳۴۰۴۵	۰/۲۷۲۴۳۱
۱۰	۰/۸۱۸۴۱۸	۰/۳۵۲۷۷۱	۰/۶۴۵۳۷۰	۰/۶۴۰۰۸۸	۰/۱۳۹۷۷۶	۰/۲۸۱۵۰۸

۵. مثال کاربردی

در این بخش به بررسی داده‌های مربوط به شاخص کل بورس تهران از آذر ۱۳۸۹ تا شهریور ۱۳۹۲ پرداخته‌ایم. در این داده‌ها از فاصله‌های موجود بین پنج روز کاری هفته صرف نظر کرده و روزهای تعطیل به مناسبت‌های مختلف را به عنوان داده گم‌شده در نظر گرفته‌ایم. ابتدا بدون در نظر گرفتن روزهای تعطیل، مدل‌بندی کرده و مدل آرچ مناسب را اختیار کرده و برآورد پارامترها را به دست آورده‌ایم. سپس با جای‌گذاری صفر به جای داده‌های گم‌شده در باقی‌مانده‌هایی که از برازش مدل اولیه داشته‌ایم مدل آرچ مناسب را اختیار کرده و برآورد پارامترهای متناظر را به دست آورده‌ایم. لازم به ذکر است که مدل‌بندی و محاسبات را بر روی تفاضل لگاریتم شاخص کل به عبارتی $R_t = lnP_t - lnP_{t-۱}$ انجام داده‌ایم.

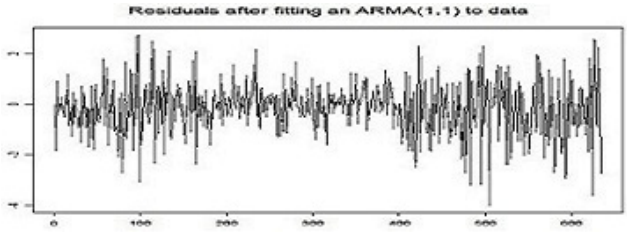


شکل ۱: نمودار خودهمبستگی و خودهمبستگی جزیی تفاضل لگاریتم شاخص کل بورس تهران

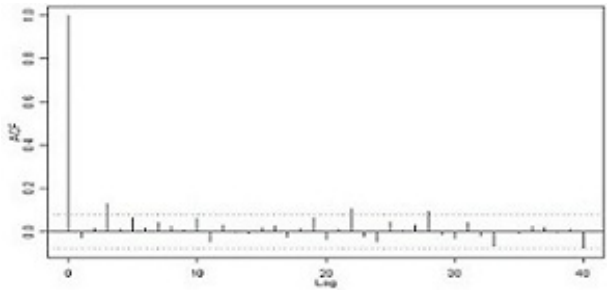
بعد از رسم نمودارهای خودهمبستگی و خودهمبستگی جزیی مربوط به تفاضل لگاریتم شاخص کل و برازاندن چند مدل مختلف و محاسبه معیار آکاییک مدل $ARMA(۱,۱)$ با کم‌ترین میزان آکاییک را به عنوان مدل برای داده‌ها در نظر گرفتیم معادله این به مدل به صورت زیر می‌باشد:

$$R_t = ۰/۰۸۸۴R_{t-۱} + X_t - ۰/۳۳۲۲X_{t-۱}.$$

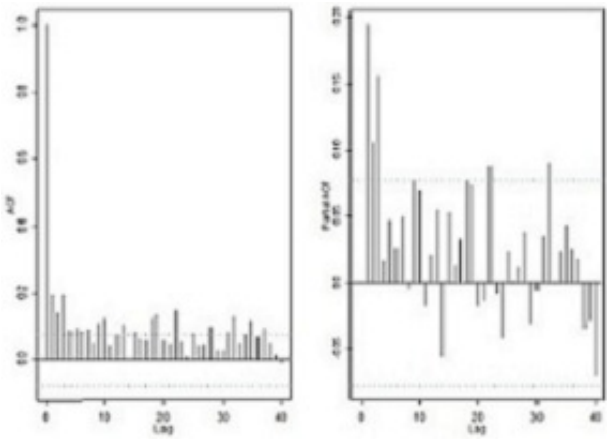
بعد از برازاندن مدل $ARMA$ نمودار باقی‌مانده‌ها و نمودار خودهمبستگی باقی‌مانده‌های این مدل را رسم می‌کنیم. نمودارهای خودهمبستگی و خودهمبستگی جزیی مربعات باقی‌مانده‌ها وجود اثر آرچ در باقی‌مانده‌ها را تأیید می‌کنند.



شکل ۲: نمودار باقی‌مانده‌های مدل ARMA



شکل ۳: نمودار خودهمبستگی باقی‌مانده‌های مدل ARMA



شکل ۴: نمودار خودهمبستگی و خودهمبستگی جزیی مربعات باقی‌مانده‌های مدل ARMA

با این حال قبل از برازش مدل آرچ، به آزمون اثرات آرچ می‌پردازیم. که مقدار $p.value$ وجود اثر آرچ را در داده‌ها نشان می‌دهد. این آزمون در ماژول $S+finmetrics$ نرم افزار $SPLUS$ انجام گرفته است.

Null Hypothesis: no ARCH effects
Test Stat : 76.8634
p.value : 0.0000

ـ با توجه به نمودارهای خودهمبستگی و خودهمبستگی جزیی سه مدل $ARCH(۱)$ ، $ARCH(۳)$ و $GARCH(۲,۳)$ را به داده‌های موجود برازش دادیم که با توجه به این‌که مقدار آکاییک مدل $ARCH(۳)$ از مقدار آکاییک بقیه مدل‌ها کم‌تر بود مدل آرچ از مرتبه $p=۳$ را برای داده‌ها در نظر گرفتیم.

	$Value$	$Std.Error$	$tvalue$	$Pr(>\|t\|)$
A	0.4190	0.05049	8.298	$6.661e-016$
$ARCH(1)$	0.2002	0.05661	3.536	$4.353e-004$
$ARCH(2)$	0.2102	0.05491	3.828	$1.422e-004$
$ARCH(3)$	0.2117	0.06626	3.195	$1.470e-003$

می‌توانیم معادله مدل آرچ به دست آمده را به صورت زیر بنویسم

$$X_t = ۰/۴۱۹۰ + ۰/۲۰۰۲X_{t-۱} + ۰/۲۱۰۲X_{t-۲} + ۰/۲۱۱۷X_{t-۳} + \epsilon_t.$$

با استفاده از روش LE که توسط بوس و موخرجی ارایه شده است برآورد پارامترهای مدل را به دست می‌آوریم که به صورت زیر می‌باشد:

$$X_t = ۰/۴۷۵۰ + ۰/۱۷۸۹X_{t-۱} + ۰/۱۷۹۶X_{t-۲} + ۰/۱۹۲۴X_{t-۳} + \epsilon_t.$$

حال روزهای تعطیل به غیر از تعطیلات آخر هفته‌ها را به عنوان داده‌های گم‌شده در نظر می‌گیریم. با ضرب باقی‌مانده‌های مدل ARMA در فرآیند داده‌های a_t به دست می‌آید که به جای داده‌های گم‌شده صفر قرار داده شده است. نمودار مربوط به این داده‌ها و هم‌چنین نمودار خودهمبستگی این داده‌ها به صورت شکل ۵ و ۶ می‌باشد. با آزمون ضرایب لاگرانژ بررسی می‌کنیم که آیا اثرات آرچ در داده‌های جدید وجود دارد یا خیر.

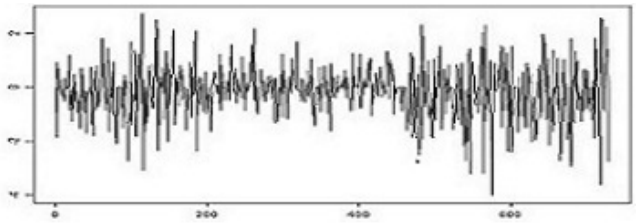
Null Hypothesis: no ARCH effects
Test Stat: 71.6771
p.value: 0.0000

مقدار p.value وجود اثرات آرچ را تأیید می‌کند. برای تشخیص مرتبه‌ی مدل آرچ نمودارهای خود همبستگی و خودهمبستگی جزیی دو مدل ARCH(۳) و GARCH(۳،۲) مدل مناسب‌تری به نظر می‌آیند. پس از برازش هر دو مدل و محاسبه‌ی معیار آکاییک مدل ARCH(۳) مدل مناسب‌تری به نظر می‌آید.

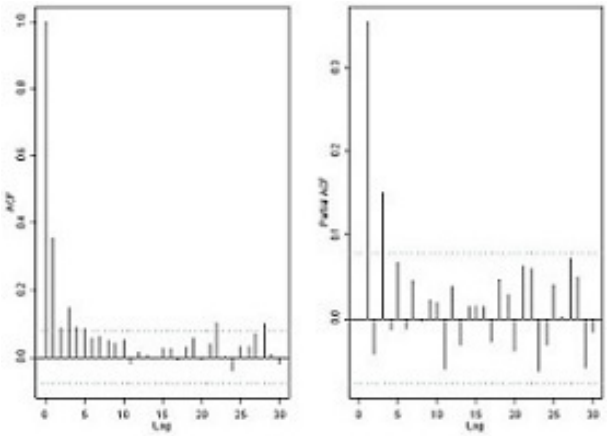
	$Value$	$Std.Error$	$tvalue$	$Pr(>\|t\|)$
A	0.4853	0.02934	16.538	0.0000000
$ARCH(1)$	0.1791	0.04702	3.810	0.0001510
$ARCH(2)$	0.1066	0.03879	2.749	0.0061213
$ARCH(3)$	0.1501	0.04814	3.117	0.0018973

با توجه به ضرایب به دست آمده مدل به صورت زیر می‌باشد

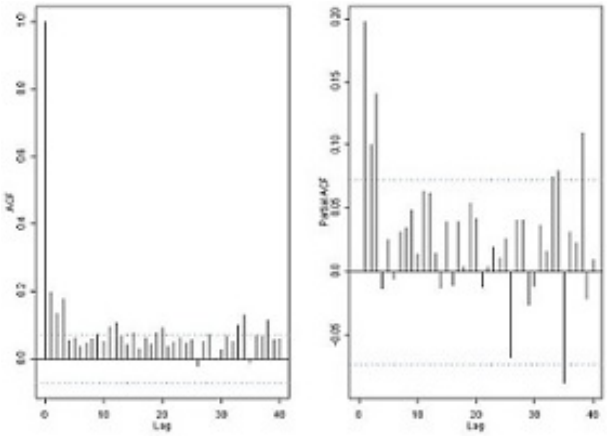
$$X_t = ۰/۴۸۵۳ + ۰/۱۷۹۱X_{t-۱} + ۰/۱۰۶۶X_{t-۲} + ۰/۱۵۰۱X_{t-۳} + \epsilon_t.$$



شکل ۵: نمودار سری‌های زمانی پس از جانهی داده‌های گم‌شده



شکل ۶: نمودار خودهمبستگی و خودهمبستگی جزیی سری‌های زمانی پس از جانهی داده‌های گم‌شده



شکل ۷: نمودار خودهمبستگی و خودهمبستگی جزیی مربع سری زمانی پس از جانهی داده‌های گم‌شده

حال با روشی که توسط بوز و موخرجی ارایه شده است برآوردهای پارامترها را به دست می‌آوریم و مدل را به صورت زیر می‌نویسم

$$X_t = ۰/۴۷۷۹ + ۰/۱۹۰۰X_{t-۱} + ۰/۱۰۸۳X_{t-۲} + ۰/۱۶۲۸X_{t-۳} + \epsilon_t.$$

۶. نتیجه گیری

در این مقاله، نحوه تعامل با داده های گمشده در مدل آرچ بررسی می گردد. بدین صورت که در سری زمانی آرچ به جای داده های گمشده صفر قرار می دهیم. سپس با تغییر مدل آرچ به فرم یک مدل رگرسیونی برآوردگرهای پارامتر مدل آرچ موسوم به براورگرد خطی در حضور داده گمشده ارائه می گردد. سازگاری قوی و توزیع مجانبی نرمال این برآوردگر اثبات شده است و با شبیه سازی نتایج تایید شده و در نهایت نتایج بر روی داده ای شاخص کل بورس تهران بکار برده شده است.

مراجع

[1] R. Fletcher and S. Leyffer, Nonlinear programming without a penalty function, *Math. Program.*, **91** no. 2 (2002) 239–269.

[2] A. Bose and K. Mukherjee, Estimating the arch parameters by solving linear equations, *J. Time Ser. Anal.*, **24** no. 2 (2003) 127–136.

[3] E. Parzen, On spectral analysis with missing observations and amplitude modulation, *Sankhyā Ser. A*, **25** (1963) 383–392.

[4] N. Shephard, *Statistical aspect of ARCH and stochastic volatility*, In D. R. Cox, D. V. Hinkley, O. E. Barndorff-Nielsen (Eds.), Time series models in econometrics, finance and other fields, London: Chapman and Hall, 1996.

[5] A. A. Weiss, Asymptotic theory for ARCH models: Estimation and testing, *Econometric Theory*, **2** no. 1 (1986) 107–131.

محمدرضا صالحی راد

تهران، دانشگاه علامه طباطبائی، دانشکده اقتصاد، گروه آمار

moresara۲۰@yahoo.com

محمدرضا صالحی راد فارغ التحصیل دکتری آمار از دانشگاه فردوسی مشهد در سال ۱۳۸۳ است. رساله دکتری ایشان با عنوان "کنترل و تکمیل سرویس درسیستم‌های صف بندی" می‌باشد. ایشان حدود ۳۶ مقاله چاپ شده در مجلات ISI، علمی پژوهشی، علمی ترویجی و کنفرانس‌های داخی و خارجی دارند. سمت فعلی ایشان معاون آموزشی و پژوهشی پردیس تحصیلات تکمیلی دانشگاه علامه طباطبائی می‌باشد.

رضا حبیبی

تهران، موسسه عالی آموزش بانکداری ایران، بانک مرکزی

r_habibi@ibi.ac.ir

رضا حبیبی فارغ التحصیل دکتری آمار از دانشگاه شیراز در سال ۱۳۸۴ است. رساله دکتری ایشان با عنوان "آشکار سازی نقطه تغییر در خانواده کلی از توزیع‌ها" تحت نظارت آقایان دکتر صدوقی و دکتر نعمت اللهی دفاع گردید. ایشان رتبه یک کنکور ارشد آمار در سال ۱۳۷۷ هستند. در سال ۱۳۹۱ ایشان به سمت هیات علمی موسسه عالی آموزش بانکداری درآمدند. به تازگی (در سال ۱۳۹۴) ایشان مدرک کارشناسی ارشد دیگری در مهندسی مالی از دانشگاه امیرکبیر اخذ نموده‌اند. زمینه‌های تحقیقاتی ایشان سری‌های زمانی مالی، مهندسی مالی، مدیریت ریسک و داده کاوی است. چندین مقاله علمی از ایشان در نشریات معتبر داخلی و خارجی به چاپ رسیده است.

مریم السادات غیبی

تهران، دانشگاه علامه طباطبائی، دانشکده اقتصاد، کارشناس ارشد گروه آمار

مریم سادات غیبی فارغ التحصیل کارشناسی ارشد آمار اقتصادی- اجتماعی از دانشگاه علامه طباطبایی در سال ۱۳۹۲ است. پایان نامه کارشناسی ارشد ایشان با عنوان "برآورد کم ترین توانهای دوم مدل آرچ با داده های گم شده" تحت نظارت آقایان دکتر حبیبی و دکتر صالحی راد دفاع گردید. ایشان جز نفرات برتر المپیاد آمار در سال ۱۳۸۹ هستند.