مجله علوم آماری، پاییز و زمستان ۱۳۹۵
جلد ۱۰، شماره ۲، ص ۱۸۵-۲۰۲
DOI: 10.18869/acadpub.jss.10.2.185

برآورد معیارهای ریسک ارزش در معرض خطر و میانگین ارزش در معرض خطر با استفاده از مدل رگرسیونی چندکی ترکیبی

علی آقامحمدی[۱]، مهدی سجودی[۲]
[۱] گروه آمار، دانشگاه زنجان
[۲] گروه ریاضی، دانشگاه تحصیلات تکمیلی علوم پایه زنجان

تاریخ دریافت: ۱۳۹۴/۲/۶ تاریخ آخرین بازنگری: ۱۳۹۴/۹/۲۹

چکیده : ارزش در معرض خطر و میانگین ارزش در معرض خطر که ریسک بازار را با یک عدد بیان می‌کنند، دو نوع از مهمترین معیارهای اندازه‌گیری ریسک مبتنی بر روشهای آماری در بازارهای مالی هستند. اخیراً رهیافت مدل‌های رگرسیونی خطی نظیر روش کمترین توان‌های دوم و چندکی برای برآورد این معیارها ارائه شده است. در این مقاله روش برآورد این دو معیار ریسک با استفاده از مدل رگرسیونی چندکی ترکیبی بررسی می‌شود. برای مقایسه کارایی مدل ارائه شده با مدل‌های متداول در این زمینه، مطالعهٔ شبیه‌سازی انجام شده و در پایان نیز نحوهٔ کاربست مدل‌ها در قالب مثال کاربردی برای داده‌هایی از بازار سهام ایران نشان داده خواهد شد.

واژه‌های کلیدی : ارزش در معرض خطر، میانگین ارزش در معرض خطر، رگرسیون چندکی ترکیبی، استنباط آماری

آدرس الکترونیک مسئول مقاله: علی آقامحمدی، Aghamohammadi.ali@znu.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): ۶۲P۰۵، ۶۲J۰۵

۱ مقدمه

جهانی سازی تجارت و تغییر ساختار بازارهای مالی بین‌المللی، بنگاه‌ها و موسسات مالی را با انواع ریسک‌ها مواجه ساخته است. ریسک در عُرف عبارت است از خطری که به دلیل عدم اطمینان در مورد وقوع حادثه‌ای در آینده پیش می‌آید و هر چه این عدم اطمینان بیشتر باشد، اصطلاحاً گفته می‌شود ریسک بالاتر است. اما در فرهنگ سرمایه‌گذاری، ریسک همان زیان بالقوه سرمایه‌گذاری است که قابل محاسبه است (راعی و سعیدی، ۱۳۸۸). امروزه هزینه‌ها و درآمدهای موسسات و بنگاه‌های مالی با ریسک‌های پیچیده‌ای مواجه‌اند که از یک سو از تعاملات کسب‌وکار جهانی و تصمیم‌گیری‌های مالی نشات می‌گیرند و از سوی دیگر از عدم اطمینان در مورد قیمت کالا، نرخ‌های ارز، نرخ‌های بهره و ارزش سهام. این ریسک‌ها، تصمیم گیری در کسب‌وکار را پیچیده کرده و موسسات و بنگاه‌های مالی را با وقایعی مواجه می‌کنند که می‌تواند ارزش بنگاه‌ها را تحت تاثیر قرار دهد. به همین دلیل ریسک یکی از مهمترین دغدغه‌های سرمایه‌گذاران است و به‌عنوان معیاری مهم در تصمیم‌گیری‌های سرمایه‌گذاری محسوب می‌شود. در واقع تن‌دادن به یک سرمایه‌گذاری بدون توجه به ریسک آن، قرار گرفتن در شرایط زیان‌بار است. بنابراین مفهوم ریسک نقشی کلیدی در بازارهای مالی ایفا می‌کند که شناسایی انواع ریسک‌ها، اندازه‌گیری و مدیریت آن از اهمیت بالایی برخوردار است. تاکنون برای اندازه‌گیری ریسک، معیارهای مختلفی از طریق شاخص‌های پراکندگی آماری ارائه شده است.

همزمان با ظهور ابزارهای مشتقه در دهه‌های اخیر میلادی، مدیریت ریسک نیز با چالش‌های جدیدی مواجه گردید. در واقع روش‌های سنتی مدیریت ریسک، دیگر پاسخ‌گوی کنترل ریسک‌های ناشی از این نوع ابزارهای نوپا نبودند. گروه ده[۱] که به‌عنوان یک مرکز بین‌المللی برای وضع قوانین بانک‌داری شناخته می‌شود، در سال ۱۹۹۳ میلادی ضمن ارائه گزارشی در مورد ریسک بانک‌ها، تعریف معیاری مناسب و جامع برای اندازه‌گیری ریسک را ضروری دانست. از این رو تعریف و

[۱] G-10

تبیین معیارهای جدیدی از اندازه‌های ریسک بر اساس روشهای آماری، نظیر ارزش در معرض خطر[۲] (VaR) توسط بانک‌های آمریکایی شروع شد و به سرعت نیز در بازارهای مالی عمومیت پیدا کرد. زیرا این معیار، تمامی انواع ریسک را در یک عدد خلاصه کرده و مقدار سرمایه‌ای که ممکن است مورد زیان قرار گیرد را تعیین کرده و از این رو مدیران و حسابداران را از انجام انواع محاسبات مربوط به ریسک باز می‌دارد. در واقع از طریق این معیار می‌توان ریسک را به‌طور کمی محاسبه و برای مقابله با آن، تمهیدات لازم را اندیشید. معیار VaR، برابر حداکثر زیانی است که در یک دوره زمانی معین (یک روز، یک هفته و غیره) در سطح اطمینان مشخصی، ممکن است برای مجموعه‌ای از دارایی‌ها رخ دهد. به‌عنوان مثال، وقتی گفته می‌شود که ارزش در معرض خطر یک دارایی در سطح اطمینان ۹۹ درصد روزانه یک میلیون دلار است، به این معنا است که انتظار داریم در هر ۱۰۰ روزی که مبادلات تجاری در آنها صورت می‌گیرد، تنها در یک روز زیانی بالاتر از یک میلیون دلار داشته باشیم (شرکت ماتریس تحلیل‌گران سیستم‌های پیچیده، ۱۳۸۸). معمولاً سرمایه‌ای که لازم است تا یک بنگاه مالی برای مقابله با خطرات احتمالی در دسترس نگه دارد، دست کم سه برابر میزان برآورد شده برای VaR تعیین می‌شود. از خطرات احتمالی که بانک‌ها و بنگاه‌های مالی را تهدید می‌کنند، می‌توان به نکول مشتریان از بازپرداخت وام، بحران‌های مالی که ممکن است رخ دهند و یا هر اتفاق دیگری که سرمایه‌های بنگاه اقتصادی و یا زیرمجموعه‌های وابسته به آن را تهدید می‌کنند، اشاره کرد.

برخی تحقیقات نشان داده‌اند در معاملاتی که ضرر بالقوه آن‌ها زیاد است، ممکن است معیار VaR ریسک را کمتر از آنچه که هست برآورد کند و سرمایه‌گذاری را به سمت زیان دهی سوق دهد. بنابراین در این گونه موارد، VaR معیاری مناسب برای سنجش ریسک محسوب نمی‌شود. معیار مناسب دیگری که برای سنجش ریسک استفاده می‌شود، میانگین ارزش در معرض خطر[۳] (AVaR) است. این معیار برای سنجش حداکثر زیان در شرایط نامطلوب بازار به‌کار می‌رود. بر خلاف VaR که ریسک را در سطح اطمینان مشخصی بیان می‌کند، معیار AVaR ریسک را برای تمام

[۲] Value at Risk
[۳] Average Value at Risk

سطوح بالاتر از سطح اطمینان مشخص محاسبه می‌کند (راعی و سعیدی، ۱۳۸۸).

برآورد این دو معیار ریسک، همواره یکی از موضوعات مهم در مدیریت ریسک محسوب می‌شود. به همین دلیل در سال‌های اخیر روش‌های مناسب و کارا برای برآورد این معیارها بسیار مورد توجه قرار گرفته است. از جمله می‌توان به روش‌هایی مانند مونت‌کارلو، بوت‌استرپ، آمار بیزی و روش‌های آماری دیگر اشاره کرد. اما در بسیاری از موارد، بازده آتی دارایی‌ها ممکن است وابسته به متغیرهای توصیفی دیگری باشد که استفاده از این متغیرها می‌تواند در تحلیل رفتار بازده موثر واقع شود. در این راستا چان و همکاران (۲۰۱۲) از روش رگرسیون خطی میانگین[۴] و چندکی[۵] برای برآورد معیارهای VaR و AVaR استفاده کردند. در این مقاله هدف بهبود برآورد این معیارها با مدل رگرسیونی چندکی ترکیبی[۶] است که اولین بار توسط زو و یوان (۲۰۰۸) ارائه و توسط تانگ و همکاران (۲۰۱۲) و تانگ و همکاران (۲۰۱۵) بسط داده شد. در بخش ۲ ابتدا روش‌های محاسبه VaR و AVaR ارایه شده و سپس برآورد آنها با روش رگرسیونی میانگین و چندکی توصیف خواهد شد. در بخش ۳ مدل رگرسیونی چندکی ترکیبی برای برآورد VaR و AVaR بررسی خواهند شد. بخش ۴ شامل ارزیابی روش‌ها با استفاده از شبیه‌سازی و تحلیل داده‌هایی از بازار سهام ایران است. بخش ۵ نیز به بحث و نتیجه‌گیری می‌پردازد.

۲ معیارهای VaR و AVaR

فرض کنید متغیر تصادفی $Y_t = p_0 - P_t$، که در آن p_0 ارزش دارایی یا مجموعه‌ای از دارایی‌ها در مبدأ زمان و متغیر تصادفی P_t ارزش آن در زمان آتی t باشد. توجه کنید p_0 مقداری معلوم ولی P_t متغیر تصادفی است، بنابراین Y_t نیز متغیری تصادفی است که کاهش ارزش دارایی را در دوره آتی مشخص می‌کند. از نظر ریاضی ارزش در معرض خطر Y_t، به‌صورت

$$VaR_\alpha(Y_t) = \inf\{y;\ F_{Y_t}(y) \geq \alpha\}, \qquad (1)$$

۴ Mean Regression

۵ Quantile Regression

۶ Composite Quantile Regression

تعریف می‌شود، که در آن $\alpha \in (۰, ۱)$ اندازه اطمینان و معمولاً برابر ۹۵ یا ۹۹ درصد در نظر گرفته می‌شود و $F_{Y_t}(y)$ نیز تابع توزیع تجمعی متغیر Y_t در نقطه y را نشان می‌دهد (چان و همکاران، ۲۰۱۲). رابطهٔ (۱) به این معنی است که احتمال اینکه کاهش ارزش دارایی در دوره آتی بیش از ارزش در معرض خطر باشد، حداکثر برابر $۱ - \alpha$ است. یا به عبارت دیگر احتمال اینکه زیان دارایی در دوره آتی کمتر از ارزش در معرض خطر باشد، حداقل برابر α است. توجه شود اگر متغیر Y_t پیوسته باشد، آنگاه $\mathrm{VaR}_\alpha(Y_t) = F_{Y_t}^{-۱}(\alpha)$ و همچنین $\Pr\{P_t - p_۰ < -\mathrm{VaR}_\alpha(\mathrm{Y_t})\} \leq 1 - \alpha$، بنابراین احتمال اینکه سود دارایی در دوره آتی کمتر از "ارزش در معرض خطر" باشد، حداکثر برابر $۱ - \alpha$ است (مهدوی و ماجدی، ۱۳۸۹). ارزش در معرض خطر را هم می‌توان برای زیان و هم سود یا همان بازده آتی دارایی تعریف کرد. در اینجا $-\mathrm{VaR}_\alpha(\mathrm{Y_t})$ با $\mathrm{VaR}_{1-\alpha}(\mathrm{Y_t})$ نشان داده می‌شود.

معیار ریسک AVaR برای Y_t در سطح اطمینان α به‌صورت

$$AVaR_\alpha(Y_t) = \inf_{y \in R} \left\{ y + (۱ - \alpha)^{-۱} E(Y_t - y)_+ \right\}, \qquad (۲)$$

تعریف می‌شود، که در آن $a_+ = \max\{۰, a\}$ (راکفلر و اُراسیو، ۲۰۰۲؛ چان و همکاران، ۲۰۱۲) و یا به‌طور معادل (آسربای، ۲۰۰۲):

$$AVaR_\alpha(Y_t) = \frac{۱}{(۱ - \alpha)} \int_\alpha^۱ VaR_\tau(Y_t) d\tau. \qquad (۳)$$

یکی دیگر از روش‌های کارا برای محاسبه VaR و AVaR، استفاده از مدل‌های رگرسیونی خطی است. زیرا نوسانات قیمت سهام یا مجموعه‌ای از دارایی‌ها در بازار، علاوه بر نوسانات قیمت سهامی از نوع خود، وابسته به قیمت سهام صنایع و شاخص‌های دیگر نیز است. در این مدل‌ها فرض می‌شود، نوسانات همزمان قیمت‌ها با یکدیگر ناشی از یک یا چند عامل عمومی است. بنابراین با استفاده از مدل رگرسیونی می‌توان قیمت سهام را به‌صورت

$$Y = \beta_۰ + \boldsymbol{x}'\boldsymbol{\beta} + \epsilon, \qquad (۴)$$

مدل‌بندی کرد، که در آن Y قیمت سهام، $\boldsymbol{x}$ عوامل تاثیر گذار بر آن و ϵ مولفه خطا است. فرض کنید تابع $\rho(.)$ یک اندازه ریسک دلخواه باشد (آرتزنر و همکاران، ۱۹۹۹) با استفاده از مدل رگرسیونی رابطه (۴) و با توجه به استقلال خطا و متغیرهای توضیحی، داریم

$$\rho(Y|\boldsymbol{x}) \;=\; \rho(\beta_0 + \boldsymbol{x}'\boldsymbol{\beta} + \epsilon) = \beta_0 + \boldsymbol{x}'\boldsymbol{\beta} + \rho(\epsilon),$$

که در آن $\rho(\epsilon)$، اندازه ریسک را برای متغیر ϵ نشان می‌دهد (چان و همکاران، ۲۰۱۲). در برآورد $\rho(Y|\boldsymbol{x})$، هدف برآورد ضرایب نامعلوم β_0، $\boldsymbol{\beta}$ و نیز برآوردی مناسب از $\rho(\epsilon)$ است. حال با جایگذاری دو اندازه ریسک VaR و AVaR به جای تابع ریسک دلخواه $\rho(.)$، می‌توان ارزش در معرض خطر و میانگین ارزش در معرض خطر Y را به صورت

$$\begin{aligned} VaR_\alpha(Y|\boldsymbol{x}) &= \beta_0 + \boldsymbol{x}'\boldsymbol{\beta} + VaR_\alpha(\epsilon), \\ AVaR_\alpha(Y|\boldsymbol{x}) &= \beta_0 + \boldsymbol{x}'\boldsymbol{\beta} + AVaR_\alpha(\epsilon), \end{aligned} \qquad (۵)$$

در نظر گرفت (چان و همکاران، ۲۰۱۲؛ گالیانو و همکاران، ۲۰۱۱).

۳ برآورد VaR و AVaR با استفاده از مدل‌های رگرسیونی خطی

واضح است در عمل محاسبه VaR و AVaR در روابط (۱)، (۲) و (۵) به دلیل نامعلوم بودن توزیع Y_t و ϵ، امکان پذیر نیست. لذا باید این دو معیار با استفاده از مشاهدات یا داده‌های تاریخی[۷] برآورد شوند. در این بخش برآورد این معیارها با استفاده از مدل‌های رگرسیونی خطی بر اساس مشاهدات تاریخی $(Y_1, \boldsymbol{x}_1), \ldots, (Y_n, \boldsymbol{x}_n)$ بررسی می‌شوند.

۱.۳ مدل رگرسیونی میانگین و چندکی

چون هدف برآورد $\mathrm{VaR}_\alpha(Y|\boldsymbol{x})$ و $\mathrm{AVaR}_\alpha(Y|\boldsymbol{x})$ با استفاده از رابطه (۵) است، در روش مدل رگرسیونی میانگین چان و همکاران (۲۰۱۲) پیشنهاد کردند، ابتدا

۷ Historical data

پارامترهای $\boldsymbol{\beta}$ و β_0 را با روش کمترین توان‌های دوم خطا برآورد کرده و سپس با محاسبه مانده‌های مدل، یعنی $e_i = Y_i - \hat{\beta}_0 - \boldsymbol{x}_i'\hat{\beta}$ برای $i = 1,\dots,n$، $\mathrm{VaR}_\alpha(\epsilon)$ نیز به‌صورت

$$V\hat{a}R_\alpha(\epsilon) = \hat{F}_{e,n}^{-1}(\alpha) = e_{([n\alpha]+1)},$$

برآورد شود، که در آن $\hat{F}_{e,n}(.)$ تابع توزیع تجربی $e_{([n\alpha])}$، آماره ترتیبی $([n\alpha])$ام برای e_iها است و $[a]$ نیز برابر کوچکترین عدد صحیح مساوی یا بزرگتر از a می‌باشد. برآورد معیار $\mathrm{AVaR}_\alpha(\epsilon)$ نیز در این مدل با توجه به رابطه (۲) به‌صورت

$$\begin{aligned} A\hat{V}aR_\alpha(\epsilon) &= \inf_{t\in R}\{t + \frac{1}{(1-\alpha)n}\sum_{i=1}^{n}(e_i - t)_+\} \\ &= V\hat{a}R_\alpha(\epsilon) + \frac{1}{(1-\alpha)n}\sum_{i=1}^{n}\Big(e_i - V\hat{a}R_\alpha(\epsilon)\Big)_+ \\ &= e_{([n\alpha])} + \frac{1}{(1-\alpha)n}\sum_{i=[n\alpha]+1}^{n}\left(e_{(i)} - e_{([n\alpha])}\right). \end{aligned}$$

به دست می‌آید (چان و همکاران، ۲۰۱۲). بنابراین با برآورد ضرایب $\boldsymbol{\beta}$ و β_0 به روش کمترین توان‌های دوم و نیز برآوردهای ارائه شده برای $\mathrm{VaR}_\alpha(\epsilon)$ و $\mathrm{AVaR}_\alpha(\epsilon)$، برآوردهای $\mathrm{VaR}_\alpha(Y|\boldsymbol{x})$ و $\mathrm{AVaR}_\alpha(Y|\boldsymbol{x})$ با استفاده از رابطهٔ (۵) قابل محاسبه هستند.

با توجه به رابطهٔ (۱)، در واقع VaR همان چندک α-ام متغیر تصادفی Y_t است. بنابراین استفاده ازمدل‌های رگرسیونی چندکی که تاثیر متغیرهای توضیحی را در چندک‌های مختلف متغیر پاسخ بررسی می‌کنند، ایده مناسبی برای برآورد دو معیار VaR و AVaR است (چان و همکاران، ۲۰۱۲؛ گالیانو و همکاران، ۲۰۱۱). اگر فرض شود در مدل رابطهٔ (۴) چندک αام متغیر ϵ برابر صفر است، آنگاه چندک αام متغیر Y برابر $\beta_0 + \boldsymbol{x}'\boldsymbol{\beta}$ خواهد بود. با این فرض برای ϵ، برآورد $\boldsymbol{\beta}$ و β_0 بر اساس n مشاهده در مدل رگرسیونی چندکی به‌صورت

$$(\hat{\beta}_0, \hat{\boldsymbol{\beta}}) = \arg\min_{\boldsymbol{\beta},\beta_0}\sum_{i=1}^{n}\rho_\alpha(y_i - \beta_0 - \boldsymbol{x}_i'\boldsymbol{\beta})$$

به‌دست می‌آید، که در آن تابع زیان $\rho_\alpha(t)$ به‌صورت

$$\rho_\alpha(t) = \frac{|t| + (۲\alpha - ۱)t}{۲}$$

تعریف می‌شود (کوئنکر و باست، ۱۹۷۸؛ کوئنکر و ماکادو، ۱۹۹۹). برای حل مسئله بهینه‌سازی فوق می‌توان از بستهٔ quantreg در نرم‌افزار R استفاده کرد. توجه شود در رگرسیون چندکی، برآورد ضرایب مدل از مینیمم کردن مجموع قدر مطلق انحرافات موزون به‌دست می‌آیند، لذا این روش نسبت به داده‌های دور افتاده استوار است. چون در این مدل چندک αام متغیر ϵ برابر صفر است، بنابراین $\mathrm{VaR}_\alpha(\epsilon) = 0$ و با استفاده از رابطهٔ (۵) داریم (چان و همکاران، ۲۰۱۲)

$$V\hat{a}R_\alpha(Y|\boldsymbol{x}) = \hat{\beta}_\circ + \boldsymbol{x}'\hat{\boldsymbol{\beta}}.$$

برای برآورد $\mathrm{AVaR}_\alpha(\mathrm{Y})$ در مدل رگرسیونی چندکی از تقریب انتگرال به روش مونت کارلو استفاده می‌شود. با توجه به رابطهٔ (۳) که بیان انتگرالی AVaR است، می‌توان این معیار را به‌صورت

$$AV\hat{a}R_\alpha(Y) = \frac{۱}{(۱-\alpha)}\int_\alpha^۱ V\hat{a}R_\tau(Y)d\tau \simeq \frac{۱}{r}\sum_{j=۱}^{r} V\hat{a}R_{\alpha_j}(Y),$$

برآورد کرد، که در آن $j = ۱, \dots, r$ $\quad \alpha_j = \alpha + (j - \frac{۱}{۲})\Delta$ و $\Delta = \frac{۱-\alpha}{r}$ (چان و همکاران، ۲۰۱۲). معمولاً در محاسبات مقدار r برابر جز صحیح عدد ثابت $(۱-\alpha)n$ در نظر گرفته می‌شود.

۲.۳ مدل رگرسیونی چندکی ترکیبی

مدل رگرسیونی چندکی که توسط کوئنکر و باست (۱۹۷۸) ارائه شد، مدلی تک سطحی است، به این معنی که برآوردها براساس فقط یک مقدار مشخص چندک مانند α به دست می‌آیند. رگرسیون چندکی بر اساس یک سطح چندک، ممکن است برآوردهای کارایی ایجاد نکند. به همین دلیل روش کمترین توان‌های دوم اغلب در مقایسه با این روش کارا است (زو و یوان، ۲۰۰۸). بنابراین زو و یوان

(۲۰۰۸)، مدل رگرسیونی چندکی ترکیبی که در آن برآورد پارامترها بر اساس ترکیبی از سطوح متفاوت چندک‌ها محاسبه می‌شوند، ارائه کردند. همچنین آنها نشان دادند این روش در مقایسه با روش کمترین توان‌های دوم و چندکی کارا است. در روش رگرسیونی چندکی ترکیبی، برآورد پارامترهای $\beta_۰$ و $\boldsymbol{\beta}$ در مدل (۴) از طریق تابع زیان ترکیبی، به‌صورت

$$(\hat{\boldsymbol{\beta}}, \hat{\beta}_۰, \hat{b}_{\tau_۱}, \ldots, \hat{b}_{\tau_K}) = \arg\min_{\boldsymbol{\beta}, \beta_۰, b_{\tau_۱}, \ldots, b_{\tau_K}} \sum_{i=۱}^{n} \sum_{k=۱}^{K} \rho_{\tau_k}(y_i - b_{\tau_k} - \beta_۰ - \boldsymbol{x}_i'\boldsymbol{\beta}), \quad (۶)$$

محاسبه می‌شوند، که در آن $۱ > \tau_K > \ldots > \tau_۱ > ۰$، همان K سطح چندک را نشان می‌دهند. طبق مدل (۴) برای سطح τ_k، چندک τ_kام Y_i به شرط $\boldsymbol{x}_i$، برابر $\beta_۰ + \boldsymbol{x}_i'\boldsymbol{\beta} + b_{\tau_k}$ و b_{τ_k} نیز چندک τ_kام متغیر ϵ_i است. در تحلیل داده‌ها و نیز مطالعه شبیه‌سازی زو و یوان (۲۰۰۸) و تانگ و همکاران (۲۰۱۵)، مقدار K را برابر ۲۰ قرار داده و τ_kها را به‌صورت $\tau_k = \frac{k}{K+۱}$، $k = ۱, \ldots, ۲۰$ در نظر گرفتند.

برای برآورد معیار VaR در سطح α در روش رگرسیونی چندکی ترکیبی، τ_kها برای $k = ۱, \ldots, ۱۹$، مانند زو و یوان (۲۰۰۸) و تانگ و همکاران (۲۰۱۵) در نظر گرفته و $\tau_{۲۰}$ برابر همان α قرار داده می‌شود. با توجه به رابطهٔ (۵)، در روش رگرسیونی چندکی ترکیبی داریم

$$V\hat{a}R_\alpha(Y|\boldsymbol{x}) = \hat{\beta}_۰ + \boldsymbol{x}'\hat{\boldsymbol{\beta}} + \hat{b}_{\tau_{۲۰}},$$

که در آن $\hat{\boldsymbol{\beta}}$، $\hat{\beta}_۰$ و $\hat{b}_{\tau_{۲۰}}$ از رابطه (۶) به دست می‌آیند. توجه شود در این حالت برآوردها بر اساس ۲۰ سطح متفاوت چندک‌ها محاسبه می‌شوند و همچنین $\hat{b}_{\tau_{۲۰}}$ برآورد $\mathrm{VaR}_\alpha(\epsilon)$ است.

برای برآورد AVaR در سطح α مانند روش رگرسیونی چندکی (چان و همکاران، ۲۰۱۲) از تقریب انتگرال استفاده می‌شود. اگر قرار دهیم $\tau_{۲۰} = \alpha_j = \alpha + (j - \frac{۱}{۲})\Delta, \ j = ۱, \ldots, r$، آنگاه با توجه به رابطهٔ (۵) برآورد AVaR در سطح α به‌صورت

$$AV\hat{a}R_\alpha(Y|\boldsymbol{x}) = \frac{۱}{(۱-\alpha)} \int_\alpha^۱ V\hat{a}R_\tau(Y|\boldsymbol{x}) d\tau \simeq \frac{۱}{r} \sum_{j=۱}^{r} V\hat{a}R_{\alpha_j}(Y|\boldsymbol{x}),$$

به‌دست می‌آید، که در آن

$$V\hat{a}R_{\alpha_j}(Y|\boldsymbol{x}) = \hat{\beta}_۰ + \boldsymbol{x}'\hat{\boldsymbol{\beta}} + \hat{b}_{\alpha_j}, \quad j = ۱, \ldots, r,$$

و $\hat{\beta}_۰$، $\hat{\boldsymbol{\beta}}$ و $\hat{b}_{\alpha_j}$ با فرض $\tau_{۲۰} = \alpha_j$ از رابطهٔ (۶) محاسبه می‌شوند. برای حل مسئلهٔ بهینه سازی رابطه (۶) بسته نرم‌افزاری خاصی موجود نیست. اما با توجه به تعریف تابع $\rho(.)$ که یک تابع محدب است، می‌توان آن را به‌عنوان یک مسئله بهینه‌سازی محدب در نظر گرفته و با تعریف متغیرهای کمکی آن را به یک مسئلهٔ بهینه‌سازی خطی به‌صورت

$$\arg\min_{\boldsymbol{\beta},\beta_۰,b_{\tau_۱},\ldots,b_{\tau_K}} \sum_{k=۱}^{K}\sum_{i=۱}^{n} \tau_k\xi_{ik} + (۱-\tau_k)\zeta_{ik}$$
$$subject\ to \begin{cases} \xi_{ik} - \zeta_{ik} = y_i - b_{\tau_k} - \beta_۰ - \boldsymbol{x}_i'\boldsymbol{\beta}, \\ \xi_{ik} \geq ۰, \ \ \zeta_{ik} \geq ۰, \quad k = ۱, \ldots, K, \ \ i = ۱, \ldots, n, \end{cases}$$

تبدیل کرد (تانگ و همکاران، ۲۰۱۲). برای حل مسئلهٔ بهینه سازی خطی فوق می‌توان از بستهٔ CVX در نرم‌افزار متلب استفاده کرد.

۴ مطالعه شبیه‌سازی

در این بخش برای مقایسه کارایی روش رگرسیون چندکی ترکیبی با دو روش رگرسیونی میانگین و چندکی ارایه شده توسط چان و همکاران (۲۰۱۲) در برآورد معیارهای VaR و AVaR مطالعه شبیه‌سازی انجام می‌شود. مدل رگرسیونی برای تولید داده‌ها به‌صورت

$$Y_i = \beta_۰ + \beta_۱ x_i + \epsilon_i \quad i = ۱, \ldots, n$$

در نظر گرفته شده است. مقادیر واقعی ضرایب رگرسیونی را به‌صورت $\beta_۰ = ۰$ و $\beta_۱ = ۱$ قرار داده و مقادیر x_iها نیز از توزیع یکنواخت در بازهٔ $(-۲, ۲)$ تولید شده‌اند. از چهار توزیع نرمال استاندارد $\mathrm{N}(۰, ۱)$، تی-استودنت با ۳ درجه آزادی $t(۳)$، توزیع نرمال آمیخته $\mathrm{CN}(۱, ۹)$، به‌صورت $۰/۸\mathrm{N}(۰, ۱) + ۰/۲\mathrm{N}(۱, ۹)$ و توزیع لگ-نرمال $\mathrm{LN}(۰, ۱)$، برای تولید مولفهٔ خطا استفاده می‌شود. توجه کنیم که

سه توزیع آخر نسبت به توزیع نرمال استاندارد دارای دم‌های پهن‌تری هستند که معمولا بهتر به داده‌های مالی برازش می‌یابند. پس از تولید نمونه به حجم n طبق مدل رگرسیونی، با استفاده از سه روش رگرسیونی کمترین توان‌های دوم، چندکی و چندکی ترکیبی پارامترها برآورد شده و سپس معیارهای VaR و AVaR متغیر Y، برای ۵۰۰ مقدار متفاوت x که به فاصله‌های مساوی از بازهٔ $[-۲,۲]$ انتخاب شده‌اند، برآورد می‌شوند. همچنین مقدار واقعی این دو معیار با توجه به مدل رگرسیونی، مقادیر عنوان شده برای پارامترها و توزیع مولفه خطاها برای ۵۰۰ مقدار مختلف x محاسبه می‌شوند. برای مقایسه کارایی سه روش عنوان شده و به منظور لحاظ کردن تغییرپذیری نتایج حاصل از شبیه‌سازی، ۵۰۰ مرتبه روند شبیه‌سازی را تکرار کرده ($k = ۱, \ldots, ۵۰۰$) و برای هر مقدار x، معیار متوسط قدر مطلق خطا[۸] (MAE) را برای VaR و AVaR از روابط

$$\begin{aligned} MAE_{VaR}(x_k) &= \frac{۱}{۵۰۰}\sum_{j=۱}^{۵۰۰}|V\hat{a}R_\alpha(Y_j|x_k) - VaR_\alpha(Y_j|x_k)|, \\ MAE_{AVaR}(x_k) &= \frac{۱}{۵۰۰}\sum_{j=۱}^{۵۰۰}|AV\hat{a}R_\alpha(Y_j|x_k) - AVaR_\alpha(Y_j|x_k)|, \end{aligned}$$

به دست می‌آوریم. سپس متوسط مقدار MAE_{VaR}ها (EMAE_{VaR}) و متوسط مقدار MAE_{AVaR}ها ($\text{EMAE}_{\text{AVaR}}$) را برای ۵۰۰ مقدار مختلف x محاسبه می‌کنیم که نتایج آن برای مقادیر مختلف α، n و مولفهٔ خطا، برای روش‌های رگرسیونی کمترین توان‌های دوم (LS)، چندکی (QR) و چندکی ترکیبی (CQR) در جدول ۱ آمده است. واضح است هر چه مقدار EMAE برای روشی کوچکتر باشد، حاکی از عملکرد بهتر آن روش است. با توجه به مقدار EMAE برای دو معیار VaR و AVaR در جدول ۱، به وضوح روش رگرسیون چندکی ترکیبی نسبت به دو روش دیگر دارای عملکرد بهتری است. هر چند در حالتی که توزیع خطاها نرمال استاندارد هستند، روش رگرسیونی کمترین توان‌های دوم تا حدودی نسبت به روش چندکی ترکیبی عملکرد بهتری را نشان می‌دهد، اما در سه توزیع دیگر برای مولفهٔ خطا، عملکرد این روش در مقایسه با روش چندکی ترکیبی به‌طور قابل توجه‌ای پایین

[۸] Mean Absolute Error

جدول ۱: مقدار EMAE برای دو معیار VaR و AVaR

$\mathrm{EMAE}_{\mathrm{AVaR}}$			$\mathrm{EMAE}_{\mathrm{VaR}}$					
CQR	QR	LS	CQR	QR	LS	توزیع خطاها	n	α
۰/۰۴۱	۰/۰۴۸	۰/ ۰۳۹	۰/۰۴۱	۰/۰۴۱	۰/۰۳۴	N(۰, ۱)		
۰/۰۶۲	۰/۱۷۶	۰/ ۱۳۵	۰/۰۲۶	۰/۰۷۶	۰/۰۶۵	t(۳)		
۰/۰۵۲	۰/۰۹۴	۰/۰۷۵	۰/۰۳۳	۰/۰۸۲	۰/۰۶۸	CN(۱, ۹)	۲۰۰۰	
۰/۰۹۱	۰/۳۶۴	۰/۲۶۷	۰/۰۶۹	۰/۱۵۷	۰/۱۱۷	LN(۰, ۱)		
								۰/۹۹
۰/۱۲۹	۰/۱۶۷	۰/ ۱۳۰	۰/۰۹۳	۰/۱۳۱	۰/۱۰۰	N(۰, ۱)		
۰/۷۲۱	۱/۳۶۳	۱/۰۵۶	۰/۲۳۶	۰/۵۷۵۴	۰/۴۳۴	t(۳)		
۰/۳۱۲	۰/۸۶۲	۰/۵۹۳	۰/۱۶۷	۰/۲۶۱	۰/۲۰۵	CN(۱, ۹)	۱۰۰۰	
۰/۸۴۲	۲/۶۴۷	۲/۰۳۸	۰/۵۱۱	۱/۳۴۲	۰/۹۷۵	LN(۰, ۱)		
۰/۰۵۱	۰/۰۷۲	۰/۰۵۲	۰/۰۳۵	۰/۰۱۱	۰/۰۲۲	N(۰, ۱)		
۰/۰۴۶	۰/۱۵۴	۰/ ۱۲۵	۰/۰۱۴	۰/۰۶۲	۰/۰۵۸	t(۳)		
۰/۰۴۱	۰/۰۸۵	۰/۰۶۲	۰/۰۳۰	۰/۰۷۳	۰/۰۵۳	CN(۱, ۹)	۲۰۰۰	
۰/۱۴۷	۰/۳۳۷	۰/۲۵۱	۰/۱۴۶	۰/۳۲۱	۰/۲۰۱	LN(۰, ۱)		
								۰/۹۵
۰/۰۹۶	۰/۱۵۳	۰/۱۲۷	۰/۱۰۱	۰/۱۲۵	۰/۰۸۳	N(۰, ۱)		
۰/۲۹۶	۰/۷۰۱	۰/۶۵۲	۰/۰۷۶	۰/۲۹۲	۰/۱۹۳	t(۳)		
۰/۲۱۸	۰/۸۲۱	۰/۴۵۲	۰/۰۹۲	۰/۲۰۳	۰/۱۷۳	CN(۱, ۹)	۱۰۰۰	
۰/۳۶۵	۰/۹۸۹	۰/۷۴۱	۰/۴۰۳	۰/۸۵۷	۰/۶۲۸	LN(۰, ۱)		

است. اما همان‌طور که عنوان شد در اکثر مواقع داده‌های مالی از توزیع نرمال تبعیت نمی‌کنند و توزیع‌های دم‌پهن‌تری نسبت به توزیع نرمال نظیر سه توزیع دیگر، بهتر به آنها برازش می‌شود. همان‌طور که نتایج بررسی‌های چان و همکاران (۲۰۱۲) نیز نشان داده است، روش رگرسیونی کمترین توان‌های دوم نسبت به روش چندکی کارایی بالایی دارد. شبیه‌سازی برای روشهای رگرسیونی کمترین توان‌های دوم و چندکی با نرم‌افزار R و چندکی ترکیبی با نرم‌افزار متلب انجام شده است.

۵ تحلیل داده‌هایی از بازار ایران

چون هدف اصلی معیارهای ریسک VaR و AVaR برآورد و تحلیل ریسک در داده‌های واقعی است، در این بخش ریسک سرمایه‌گذاری با استفاده از یک مجموعه دادهای واقعی در بازار سهام ایران با استفاده از این دو معیار مورد بررسی قرار می‌گیرد. هدف، برآورد VaR و AVaR برای قیمت سهام روزانه شرکت مخابرات ایران با نماد اخابر[۹] است. برای این منظور از داده‌های شش سال اخیر این شرکت، از تاریخ ۱۳۸۷/۵/۱۹ تا تاریخ ۱۳۹۳/۷/۲ استفاده می‌کنیم. سه عامل مهم و تاثیرگذار

[۹] Akhaber

بر قیمت سهام نظیر شاخص کل، شاخص بازار اول و شاخص صنعت را در مدل رگرسیونی به‌عنوان متغیرهای توضیحی مورد استفاده قرار می‌دهیم. مدل رگرسیونی را به‌صورت

$$Y_t = \beta_0 + \beta_1 x_{1t} + +\beta_2 x_{2t} + \beta_3 x_{3t} + \epsilon_t \quad t = 1, \ldots, 1255, \qquad (7)$$

در نظر می‌گیریم، که در آن $Y_t = \ln(\frac{P_{t+1}}{P_t})$ (P_{t+1}: قیمت سهام در زمان $t+1$ و P_t: قیمت آن در زمان t است) لگ-بازده روزانه قیمت سهام، x_1، شاخص کل است که به آن شاخص قیمت و بازده نقدی هم گفته شده و بیانگر سطح عمومی قیمت و سود سهام شرکت‌های پذیرفته شده در بورس است. x_2، شاخص بازار اول که مربوط به شرکت‌های بازار اصلی بورس است و x_3 نیز شاخص صنعت است که عملکرد شرکت‌های صنعتی در بورس را نشان می‌دهد.

معمولا برای ارزیابی دقت برآورد VaR در داده‌های واقعی، از روشی آماری به نام نرخ تخطی[10](VRate) که مقادیر برآوردها را با بازده‌های واقعی (بازده‌های مشاهده شده در طول زمان، یعنی همان Y_tها) مقایسه می‌کند، استفاده می شود (مک نیل و فرِی، ۲۰۰۰، لوریتو و همکاران، ۲۰۱۲، چان و همکاران، ۲۰۱۲). این معیار به صورت

$$VRate = \frac{1}{M} \sum_{t=N+1}^{M+N} I(Y_t < V\hat{a}R_{1-\alpha}(Y_t))$$

تعریف می‌شود، که در آن M اندازه دورهٔ پیش‌بینی و N تعداد مشاهدات است. در واقع برای محاسبه VRate در گام اول، N داده نخست را به‌عنوان مشاهدات تاریخی برای برآورد پارامترها استفاده کرده و آنگاه اندازه ارزش در معرض خطر برای مشاهده (زمان) $(N+1)$ام، یعنی $\mathrm{VaR}_{1-\alpha}(\mathrm{Y_{N+1}})$ برآورد می‌شود. داده‌ها را یک واحد به جلو برده (به عنوان مثال در گام دوم مشاهدات ۲ تا $(N+1)$ام را به‌عنوان داده‌های تاریخی در نظر گرفته و ارزش در معرض خطر برای مشاهده (زمان) $(N+2)$ام برآورد می‌شود) و فرایند برآورد را تکرار می‌کنیم. واضح است روشهایی که برای آنها $\mathrm{VRate} \leq (1-\alpha)$ است، نسبت به روشهایی که برای آنها

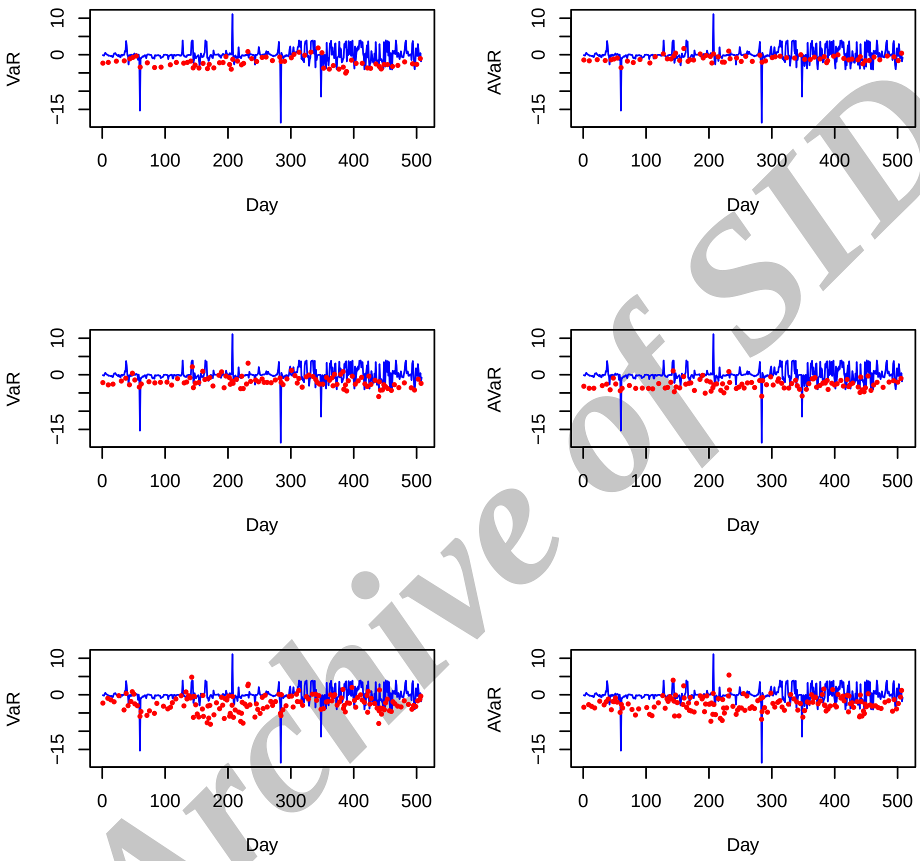
[10] Violation Rate

$\text{VRate} > (1-\alpha)$ است، کاراتراند و همچنین برای روشی که مقدار $\frac{\text{VRate}}{1-\alpha}$ نزدیک عدد یک باشد، کارایی بالایی دارد. برای داده‌های لگ-بازده قیمت سهام روزانه شرکت مخابرات ایران، اندازهٔ N را برابر ۵۰۰ در نظر گرفته و با توجه به تعداد مشاهدات که برابر ۱۲۵۵ است، به تعداد ۷۵۵ ($M = ۷۵۵$) برآورد از معیار VaR را به دست می‌آوریم. اندازهٔ $\frac{\text{VRate}}{1-\alpha}$ برای سه مقدار α و سه روش عنوان شده محاسبه شده و نتایج آن در جدول ۲ گزارش شده است. با توجه به این جدول مشخص

جدول ۲: مقادیر VRate و $\frac{\text{VRate}}{1-\alpha}$ برای سهام روزانه شرکت مخابرات ایران

مدل	α	VRate	$\frac{\text{VRate}}{1-\alpha}$
LS	۰/۹۰	۰/۱۳۴	۱/۳۴
QR	۰/۹۰	۰/۱۴۹	۱/۴۹
CQR	۰/۹۰	۰/۱۱۶	۱/۱۶
LS	۰/۹۵	۰/۰۷۲	۱/۴۴
QR	۰/۹۵	۰/۰۹۵	۱/۹
CQR	۰/۹۵	۰/۰۵۴	۱/۰۸
LS	۰/۹۹	۰/۰۲۳	۲/۳
QR	۰/۹۹	۰/۰۳۵	۳/۵
CQR	۰/۹۹	۰/۰۱۷	۱/۷

است، روش رگرسیون چندکی ترکیبی در مقایسه با دو روش دیگر برای مقادیر متفاوت α عملکرد بهتری دارد. بر خلاف مطالعه شبیه‌سازی عملکرد بهتر روش رگرسیونی چندکی ترکیبی در مقایسه با روش کمترین توان‌های دوم در داده‌های واقعی، به دلیل نرمال نبودن توزیع داده‌ها است. شکل ۱ مقادیر لگ-بازده قیمت سهام روزانه شرکت مخابرات ایران (نمودار ممتد) و مقادیر برآورد شده (نمودار نقطه‌ای) برای VaR و AVaR با استفاده از روش‌های رگرسیونی کمترین توان‌های دوم، چندکی و چندکی ترکیبی (به ترتیب از بالا به پایین) را در سطح اطمینان $\alpha = ۰/۹۵$ نشان می‌دهد. برآوردها برای روشهای رگرسیونی کمترین توان‌های دوم و چندکی در نرم‌افزار R و چندکی ترکیبی در نرم‌افزار متلب انجام شده است. نمودارها نیز با استفاده از نرم‌افزار R رسم شده‌اند.



شکل ۱: برآورد ارزش در معرض خطر و میانگین ارزش در معرض خطر برای لگ-بازده قیمت سهام روزانه شرکت مخابرات ایران با استفاده از سه مدل رگرسیونی کمترین توان‌های دوم، چندکی و چندکی ترکیبی به ترتیب از بالا به پایین، با فرض $\alpha =$ ۰/۹۵

بحث و نتیجه‌گیری

در این مقاله روشی بر اساس مدل رگرسیونی چندکی ترکیبی برای برآورد دو معیار ریسک، ارزش در معرض خطر و میانگین ارزش در معرض خطر که از معیارهای مهم ریسک سنجی در مسائل اقتصادی هستند، معرفی شد. روش پیشنهادی در واقع تعمیمی از روش رگرسیونی چندکی ارائه شده توسط چان و همکاران (۲۰۱۲) برای برآورد این دو معیار است. بر خلاف روش رگرسیونی چندکی روش ارائه شده از سطوح مختلف چندک‌ها در برآورد پارامترها استفاده می‌کند. نتایج حاصل از شبیه‌سازی و تحلیل داده‌های واقعی نشان داد، روش رگرسیونی چندکی ترکیبی در مقایسه با دو روش کمترین توان‌های دوم و چندکی دقت بالای در برآورد دو معیار VaR و AVaR دارد. نتایج حاصل از تحلیل‌های چان و همکاران (۲۰۱۲) مشخص کردند، روش رگرسیونی کمترین توان‌های دوم در کلیه توزیع‌ها عملکرد بهتری از چندکی دارد. ما نتایج حاصل از این مقاله نشان داد، هر چند روش رگرسیونی کمترین توان‌های دوم با فرض اینکه بازده داراییها دارای توزیع نرمال است، تا حدودی نسبت به روش رگرسیونی چندکی ترکیبی عملکرد بهتری دارد، اما در توزیع‌های دم پهن این روش کارایی خیلی پایینی در مقایسه با روش رگرسیونی چندکی ترکیبی دارد. همان‌طور که می‌دانیم، در بیشتر موارد داده‌های مالی از توزیع نرمال پیروی نمی‌کنند. بنابراین روش رگرسیونی چندکی ترکیبی برای برآورد دو معیار ریسک عنوان شده، در تحلیل داده‌های مالی پیشنهاد می‌شود.

تقدیر و تشکر

نویسندگان مقاله از داوران و سردبیر محترم مجله به خاطر ارائه پیشنهادات ارزنده که سبب بهبود مقاله گردید، کمال تشکر و قدردانی را دارند.

مراجع

راعی، ر. و سعیدی، ع. (۱۳۸۸)، مبانی مهندسی مالی و مدیریت ریسک، چاپ چهارم، انتشارات سمت.

شرکت ماتریس تحلیل‌گران سیستم‌های پیچیده (۱۳۸۸)، ریسک بازار با رویکرد ارزش در معرض خطر، چاپ اول، نشر آتی نگر.

مهدوی، غ. و ماجدی، ز. (۱۳۸۹)، کاربرد نظریه مقدار کرانگینی در برآورد مقدار در معرض خطر: بررسی موردی بیمه مسئولیت شرکت بیمه ایران، مجله علوم آماری، **۴**، ۵۹-۷۶.

Acerbi, C. (2002), Spectral Measure of Risk: A Coherent Representation of Subjective Risk Aversion, *Journal of Bankinf and Finance*, **26**, 1505-1518.

Artzner, P., Delbaen, F., Eber, J. E. and Heath, D. (1999), Coherent measures of Risk, *Mathematical Finance*, **26**, 203-227.

Chun, S. Y., Shapiro, A. and Uryasev, S. (2012), Conditional Value-at-Risk and Average Value-at-Risk, *Operations Research*, **60**, 739-756.

Gaglianone, W., Lima, L., Linton, O. and Smith, D. (2011), Evaluating Value-at-Risk Model via Quantile Regression, *Journal of Bussines Economics and statistics*, **29**, 150-160.

Koenker, R. and Machado, J. (1999), Goodness of Fit and Related Inference Processes for Quantile Regression, *Journal of the American Statistical Assosation*, **94**, 1296-1310.

Koenker, R. and Bassett, G. (1978), Quantile Regression, *Econometrica*, **46**, 33-50.

Leorato, S., Peracchi, F. and Tanase, A. V. (2012), Asymptoticlly Efficient Estimation of the Conditional Expected Shortfall, *Journal of Computational Statistics and Data Analysis*, **56**, 768-784.

McNeil, A. J. and Frey, R. (2000), Estimation of Tail-Related Risk Measures for for Heterosedastic Financial Time Series: An Extreme Value Approch, *Journal of Empirical Finance*, **7**, 271-300.

Rockfellar, R. T. and Uryasev, S. (2002), Conditional Value-at-Risk for General Loss Distribution, *Journal of Banking and Finance*, **26**, 1443-1471.

Tang, Y., Song, X. and zhu, Z. (2015), Variable Selection via Composite Quantile Regression with Dependent errors, *Statistica Neerlandica*, **69**, 1-20.

Tang, L., Zhou, Z. and Wu, C. (2012), Weighted Composite Quantile Estimation and Variable Selection Method for Censord Regression Model, *Statistics and Probability Letters*, **82**, 653-663.

Zou, H. and Yuan, M (2008), Composite Quantile Regression and The Oracle Model Selection Theory, *The Annals of Statistics*, **36** 1108-1126.