فصلنامه مدلسازی اقتصادسنجی سال سوم، شماره اول (پیاپی ۸)، زمستان ۱۳۹۶ (صفحات ۱۳۵ تا ۱۶۳)

ارزیابی اعتبار تجربی مدلهای قیمت‌گذاری دارایی در بازار سهام ایران: بکارگیری سطح بهینه معنی‌داری و آزمون برابری احتمال

رضا طالبلو
استادیار دانشکده اقتصاد دانشگاه علامه طباطبایی
talebloo.r@gmail.com

پریسا مهاجری
استادیار دانشکده اقتصاد دانشگاه علامه طباطبایی
parisa_m2369@yahoo.com

حسین طلاکش نایینی **(نویسنده مسئول)**
دانشجوی دکترای اقتصاد مالی دانشگاه علامه طباطبایی
talakesh@gmail.com

تاریخ دریافت: ۱۳۹۷/۰۴/۰۲ تاریخ پذیرش: ۱۳۹۷/۰۷/۰۹

چکیده

یکی از پرکاربردترین روش‌های سنجش اعتبار تجربی مدل‌های قیمت‌گذاری دارایی، استفاده از آزمون موسوم به GRS است. این مقاله درصدد است تا ضمن انجام این آزمون برای دو مدل مشهور قیمت‌گذاری دارایی (مدل قیمت‌گذاری دارایی سرمایه‌ای و مدل سه عاملی فاما و فرنچ) در بازار سهام ایران برای نخستین بار، ملاحظات مربوط به توان آزمون و سطح بهینه معنی‌داری را نیز لحاظ نماید. بر این اساس با استفاده از داده‌های بازدهی شرکت‌های حاضر در بازار بورس اوراق بهادار تهران طی سالهای ۱۳۷۹ تا ۱۳۹۵ و تشکیل سبدهای ۵*۵ موسوم به سبدهای ۲۵ تایی فاما و فرنچ، آزمون GRS با/بدون در نظر گرفتن ویژگی‌های توان آماری آزمون اجرا می-گردد. نتایج حاکی از آن است که در تمامی نمونه‌هایی که ویژگی‌های توان آزمون در نظر گرفته نمی‌شود، الگوهای قیمت‌گذاری دارایی سرمایه‌ای و سه عاملی فاما و فرنچ برقرار نیستند. لیکن در اکثر نمونه‌هایی که با در نظر گرفتن ملاحظات مربوط به توان آزمون انتخاب شده‌اند، الگوهای فوق برقرار هستند.

طبقه‌بندی JEL: *C13، C58، G12*

واژه‌های کلیدی: الگوهای قیمت‌گذاری دارایی، آزمون برابری احتمال، آزمون GRS، تحلیل توان آزمون، سطح بهینه معنی‌داری.

۱. مقدمه

با توجه به تحولات فزاینده در بازار بورس اوراق بهادار ایران و افزایش تعداد سرمایه‌گذاران و کدهای سهام‌داری طی سال‌های اخیر، بررسی عوامل مؤثر بر بازدهی سهام و آزمون الگوهای مرتبط در این زمینه، اهمیت بیشتری پیدا کرده است. یکی از مشهورترین و پرکاربردترین این الگوها، الگوهای قیمت‌گذاری دارایی‌ها هستند که در طول زمان گسترش و تکامل یافته‌اند.

در ارزیابی اعتبار تجربی مدل‌های قیمت‌گذاری دارایی، آزمون والد، به طور گسترده‌ایاستفاده شده است. این آزمون، توسط گیبونز، راس و شانکن[1] (۱۹۸۹) پیشنهاد شده و با نام آزمون GRS شناخته می‌شود. یافته‌های بسیاری از مطالعات پیشین نیز حکایت از آن دارد که فرضیه صفر آزمونGRS برای طیفی از مدل‌های قیمت‌گذاری دارایی در سطح متعارف معنی‌داری مانند ۵ درصد پذیرفته نمی‌شود.. هدف از این مقاله، بررسی اعتبار تجربی مدل‌های قیمت گذاری دارایی با استفاده از آزمون GRS و با ملاحظه دقیق و شفاف توان آماری آن در بازار بورس اوراق بهادار تهران است.

تا جایی‌که نگارندگان مقاله مطلع هستند از آزمون GRS در پژوهش‌های داخلی استفاده نشده است، از این‌رو در این مقاله تلاش می‌شود تا برای نخستین‌بار، اعتبار تجربی مدل‌های قیمت‌گذاری با استفاده از آزمون مذکور بررسی گردد. علاوه بر این، از دیگر جنبه‌های مشارکت مقاله حاضر در ادبیات موجود در ایران، بکارگیری سطح بهینه معنی‌داری و ملاحظه توان آزمون در انجام آزمون GRS می‌باشد. این موضوع از آن جهت حائز اهمیت است که در تمامی مطالعات انجام‌شده در حوزه‌های اقتصاد و مالی، سطح بهینه معنی‌داری استفاده نشده و تمامی مدل‌های بررسی شده با استفاده از سطوح متعارف معنی‌داری آزمون شده‌اند.

با عنایت به توضیحات مذکور، مقاله حاضر در پی آزمون دو فرضیه است. فرضیه نخست آن است که عبارات عرض از مبدأ معادلات یک مدل قیمت‌گذاری دارایی برای طیفی از دارایی‌ها یا سبدها بصورت مشترک برابر صفر است. پذیرش این فرضیه بدین معنا خواهد بود که مدل قیمت‌گذاری، تغییرات سیستماتیک بازدهی دارایی یا سبد را به طور مناسبی منعکس می‌کند. بر اساس فرضیه دوم، استفاده از سطح بهینه معنی‌داری و لحاظ توان آماری در بکارگیری آزمون GRS، منجر به تغییر نتایج آزمون می‌گردد. تأیید این فرضیه

---

1. Gibbons, Ross, and Shanken

نیز به معنای آن است که به جای استفاده از سطوح متعارف معنی‌داری، باید سطح بهینه معنی‌داری و توان آزمون به کار گرفته شوند.

در راستای آزمون فرضیه‌های فوق، مقاله حاضر در پنج بخش سازماندهی شده است. در بخش نخست، مبانی نظری سنجش اعتبار تجربی مدل‌های قیمت‌گذاری دارایی‌ها با تأکید بر آزمون GRS ارائه می‌گردد. بخش دوم به پیشینه تجربی پژوهش اختصاص دارد. در بخش سوم، پایه‌های آماری و متغیرهای مدل تشریح می‌شود. عملکرد مدلهای قیمت-گذاری دارایی(CAPM و سه عاملی فاما و فرنچ) با استفاده از آزمون GRS و با در نظر گرفتن ملاحظات فوق، محور اصلی مباحث بخش چهارم از مقاله را تشکیل می‌دهد. در پایان، جمع‌بندی از مهم‌ترین یافته‌های مقاله ارائه می‌شود.

۲. مبانی نظری بررسی اعتبار تجربی مدلهای قیمت‌گذاری دارایی

پس از ارائه نخستین آزمون تجربی قیمت‌گذاری دارایی به وسیله داگلاس در سال ۱۹۶۷، مدل‌های قیمت‌گذاری دارایی متعددی توسط پژوهشگران مطرح شد که از مهم‌ترین آن‌ها می‌توان به مدل CAPM توسط شارپ (۱۹۶۴) و لینتر (۱۹۶۵)[1] و مدل‌های گسترش-یافته آن به وسیله فاما و فرنچ[2] (۱۹۹۳) و کرهارت[3] (۱۹۹۷) اشاره نمود. این مدل‌ها تصریح می‌کنند که مازاد بازده یک سهم یا سبد، تابع خطی از عوامل ریسک سیستماتیک می‌باشد[4]. در مدل کرهارت، عوامل ریسک سیستماتیک شامل مازاد بازده بازار و طیفی از سایر عوامل ریسک تجربی شامل صرفه اندازه، صرفه ارزش (فاما و فرنچ، ۱۹۹۳)؛ و صرفه عامل شتاب (کرهارت، ۱۹۹۷) می‌باشند. اخیراً هو و همکاران[5] (۲۰۱۵) و فاما و فرنچ (۲۰۱۵) دو عامل ریسک اساسی دیگر(سودآوری و الگوهای سرمایه‌گذاری) را در مدل قیمت گذاری دارایی تجربی لحاظ نمودند. کوکران برای نخستین بار، عوامل مذکور را در سال‌های ۱۹۹۱ و ۱۹۹۶ معرفی نمود که در آن بازده انتظاری سهام رابطه مثبت با سودآوری مورد انتظار تنزیل شده و رابطه معکوس با نسبت سرمایه‌گذاری به دارایی شرکت دارد.

---

1. Sharpe; Lintner
2. Fama and French
3. Carhart
4. Fabozzi et al
5. Hou et al.

همانطور که در مقدمه نیز اشاره شد در ارزیابی اعتبار تجربی این نوع از مدل‌های قیمت‌گذاری دارایی، آزمون GRS (معرفی‌شده توسط گیبونز، راس و شانکن ۱۹۸۹)، بیشتر از سایر آزمون‌های آماری استفاده شده است. فرضیه صفر این آزمون آن است که عبارات عرض از مبدأ معادلات یک مدل قیمت‌گذاری دارایی برای طیفی از دارایی‌ها یا سبدها بصورت مشترک برابر صفر است. بسیاری از مطالعات گذشته گزارش می‌کنند که آزمونGRS نمی‌تواند فرضیه صفر را برای طیفی از مدل‌های قیمت‌گذاری دارایی در سطح متعارف معنی‌داری مانند ۵ درصد بپذیرد.. این موضوع سبب توسعه مدل‌ها از طریق یافتن تجربی عوامل ریسک سیستماتیک جدید گردیده است. هاروی و همکاران[1] (۲۰۱۶) جمعاً ۳۱۶ عامل را مشخص نموده‌اند که تنها ۱۲ مورد از آن‌ها منشأ نظری و سایر عوامل، منشأ تجربی دارند.

در بسیاری از پژوهش‌ها، قاعده «ارزش احتمال کمتر از ۵ درصد» مبنای تصمیم‌گیری قرار گرفته است و به نظر می‌رسد که غفلت از ملاحظات مربوط به توان آماری در به‌کارگیری آزمون GRS، از مهم‌ترین کاستی‌های اکثر مطالعات است. این نارسایی سبب شده است تا توجه بیشتر به بده-بستان سطح معنی‌داری و توان آماری در انجام آزمون فرضیه، بیش از پیش توسط برخی پژوهشگران مورد تأکید قرار گیرد که از جمله مهم‌ترین آن‌ها می‌توان به آرو، ۱۹۶۰؛ هاسمن، ۱۹۷۸؛ مَک کلوزکی و زیلیاک، ۱۹۹۶؛ مک‌کینون، ۲۰۰۲ و هاروی، ۲۰۱۷ اشاره نمود.[2]علاوه بر این، ادعای مطالعات اخیر آن است که آستانه ۵ درصد (یا ۱ درصد و ۱۰ درصد) تصادفی است و به لحاظ علمی نیز تأیید نشده است. ذکر این نکته ضروری است که آزمون‌هایی که توان آماری (یا احتمال خطای نوع دوم) را مدنظر قرار نمی‌دهند، می‌توانند نتایج استنتاجی گمراه کننده با نرخ بالایی از کشف خطا را به دنبال داشته باشند.[3] برای رفع این کاستی می‌توان از روش‌هایی استفاده نمود که سطح معنی‌داری بهینه را به عنوان تابعی از حجم نمونه یا توان آزمون در نظر می‌گیرد.

بر این اساس در این قسمت ابتدا جزئیات آزمونGRS را توضیح داده و سپس تحلیل توان آزمون، سطح معنی‌داری بهینه، و آزمون برابری احتمال را بررسی قرار خواهیم نمود.

۲-۱. آزمون GRS

[1]. Harvey et al.

[2]. Arrow, Hausman, McCloskey and Ziliak, MacKinnon, Harvey

[3]. جهت بررسی بیشتر می‌توانید به مطالعات کیم و جی (۲۰۱۵) و هاروی (۲۰۱۷) مراجعه نمایید.

یک مدل قیمت‌گذاری دارایی برای یک ورقه بهادار یا پرتفوی i $(i = 1, \ldots, N)$ را می‌توان به صورت زیر نوشت:

$$R_{i.t} - R_{F,t} = a_i + b_{iM}(R_{M.t} - R_{F.t}) + \sum_{j=2}^{K} b_{ij}X_{ji} + e_{it} \qquad (۱)$$

که در آن، $R_{i.t}$ بازدهی ورقه بهادار یا پورتفوی i در زمان t $(t = 1, \ldots, T)$، $R_{F.t}$ نرخ بازدهی بدون ریسک، a_i عرض از مبدأ، $R_{M.t}$ بازدهی وزنی ارزشی پورتفوی بازار و X_{ji} عامل ریسک علاوه بر مازاد بازده بازار $(R_{M.t} - R_{F.t})$ و e_{it} عبارت خطا است. همه بازدهی‌ها و عوامل ریسک به صورت درصد هستند. مدل قیمت‌گذاری دارایی‌های سرمایه-ای (CAPM)، یک مدل تک‌عاملی را نمایش می‌دهد که در آن مازاد بازده بازار به عنوان تنها عامل ریسک سیستماتیک می‌باشد (تعداد عوامل ریسک سیستماتیک را با K نشان دهیم، لذا در این رابطه K برابر ۱ است). فاما و فرنچ (۱۹۹۳) مدل سه عاملی را پیشنهاد دادند که بر اساس آن x ها در رابطه (۱)، شامل ، صرفه اندازه (تفاوت میان بازدهی‌های سبدهای متنوع از سهام با اندازه کوچک و سهام با اندازه بزرگ) و صرفه ارزش دفتری به ارزش بازار (تفاوت بازدهی‌های میان سبدهای متنوع‌سازی شده از سهام با ارزش دفتری به بازاری بالا و سهام با ارزش دفتری به بازاری پایین) هستند. ضمناً کرهارت (۱۹۹۷) یک مدل چهار عاملی (K=) را پیشنهاد نمود که در آن صرفه شتاب (تفاوت میان بازدهی‌های سبدهای متنوع‌سازی شده از سهام برنده گذشته و سهام بازنده گذشته) به عنوان یک عامل ریسک اضافی در نظر گرفته می‌شود. عبارت عرض از مبدأ در رابطه (۱) بازدهی تعدیل شده بر حسب ریسک برای ورقه بهادار i را نشان می‌دهد و در صورتی‌که ضرایب عاملها b_{is} ، به طور کامل تمامی منابع ریسک سیستماتیک را منعکس نمایند، انتظار می‌رود که عرض از مبدأ برابر صفر باشد.

فرضیه صفر (H_0) آزمون GRS بیان می‌دارد که a_i به طور مشترک برای تمامی i ها برابر صفر است در حالیکه فرضیه رقیب (H_1)، غیر صفر بودن حداقل یکی از a_i ها را مطرح می‌کند. تحت این فرض که عبارت خطا (e_i) با میانگین صفر و ماتریس کواریانس غیر منفرد $\sum$، دارای توزیع نرمال و مستقل است، آزمون GRS یک آزمون F نمونه محدود است که آماره آن از طریق رابطه (۲) بیان می‌شود:

$$F = \left[\frac{T(T-N-K)}{N(T-K-1)} \times \frac{\hat{a}\prime\hat{\Sigma}^{-1}\hat{a}}{1+\bar{\mu}\prime\hat{\Omega}^{-1}\bar{\mu}}\right] \qquad (۲)$$

در رابطه (۲)، T حجم نمونه، N تعداد سهام (یا سبدها) و K تعداد عوامل ریسک سیستماتیک است. از تخمین حداقل مربعات معمولی مدل (۱)، $\hat{a}$ برابر بردار $N \times 1$ برآوردگرها برای بردار عرض از مبدأها $a \equiv (a_1 . \ldots . a_n)$ است، $\hat{\Sigma}$ برآوردگر ناریب $(N \times N)$ از ماتریس کوواریانس خطا است، $\widehat{\Omega}$ ماتریس $(K \times K)$ کواریانس عوامل است و $\bar{\mu}$ بردار $(K \times 1)$ از میانگین‌های نمونه عوامل قیمت‌گذاری دارایی است. اَفلِک - گِراوز و مک دونالد (۱۹۸۹) دریافتند که آزمون GRS در برابر نرمال نبودن عبارت خطا در رابطه (۱) پایدار است. هنگامی‌که آماره F در رابطه (۲) بزرگتر از ارزش بحرانی در سطح تعیین شده معنی داری (α) است یا ارزش بحرانی آن کمتر از α است، فرضیه صفر رد می‌شود. آزمون با تعیین α در سطح ۵ یا ۱ درصد به طور گسترده‌ای انجام می‌گیرد.

بر طبق گیبونز، راس و شانکن (۱۹۸۹)، آماره F در رابطه (۲)، دارای توزیع غیر مرکزی F با درجه‌های آزادی N و $T - N - K$ است که می‌توان آن را بدین صورت بیان نمود: $F(T - N - K; \lambda)$ که λ پارامتر غیر مرکزی است و به صورت رابطه (۳) بیان می‌شود:

$$\lambda = \left[\frac{T}{1+\hat{\theta}_p^2}\right] a'^{\Sigma^{-1}} a = \left[\frac{T}{1+\hat{\theta}_P^2}\right] (\theta^{*2} - \theta_P^2) \qquad (۳)$$

$\hat{\theta}_P$ حداکثر نسبت شارپ گذشته‌نگر برای سبدهای عامل K ، θ^* شیب مرز کارایی بر مبنای تمامی دارایی‌ها و θ_P حداکثر نسبت شارپ آینده‌نگر برای سبدهای عامل K است. گیبونز، راس و شانکن (۱۹۸۹)، θ_P/θ^* را نسبت کارایی بالقوه می‌نامند. لازم به ذکر است توزیع F غیر مرکزی یک توزیع احتمال پیوسته است که تعمیم‌یافته توزیع F معمولی است. بر این اساس، هر گاه متغیر X دارای توزیع کای دوی غیرمرکزی با درجه آزادی n_1 و پارامتر غیر مرکزی λ و متغیر Y دارای توزیع کای دوی با درجه آزادی n_2 باشد و X و Y نیز به لحاظ آماری از یکدیگر مستقل باشند، کسر $\frac{(X/n_1)}{(Y/n_2)}$ دارای توزیع F غیرمرکزی با درجه‌های آزادی n_1 و n_2 و پارامتر غیرمرکزی λ است. تحت فرضیه H_0 تمامی a_i ها برابر صفر هستند و آماره F از یک توزیع مرکزی F با $\lambda = 0$ تبعیت می‌کند. تحت فرضیه H_1 که حداقل یکی از a_i ها برابر با صفر نیست، $\theta_P/\theta^* < 1$ و آماره F از یک توزیع غیر مرکزی با شاخص $\lambda > 0$ تبعیت می‌کند. یادآور می‌شود λ تابعی فزاینده از T است که به معنای تفاوت‌های رو به ازدیاد بین توزیع تحت فرضیه H_1 با توزیع تحت فرضیه H_0 (به عنوان تابعی از T) خواهد بود.

۲-۲. توان آزمون GRS و حساسیت آن به حجم نمونه

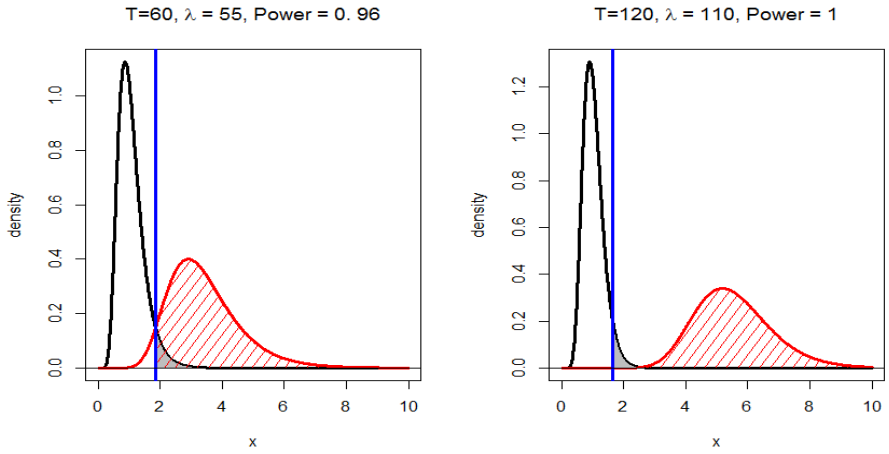
همانگونه که در بالا تعریف شد، α سطح معنی‌داری را نمایش می‌دهد که برابر احتمال رد فرضیه صفر صحیح است (خطای نوع اول). احتمال خطای نوع دوم (پذیرش فرضیه صفر نادرست) به عنوان β شناخته می‌شود و بر این اساس $(1-\beta)$ توان آزمون است. توان آزمون را می‌توان با استفاده از توزیع $F(N.T-N-K;\lambda)$ و با مقداری از λ که به وسیله مقادیری از T، θ_P و θ^* (با فرض $\hat{\theta}_P=\theta_P$ مطابق با گیبونز، راس و شانکن، ۱۹۸۹) کنترل می‌شود، محاسبه نمود. طبق رابطه (۳)، حجم نمونه T عامل تعیین‌کننده غیر مرکزی برای توزیع F در فرضیه H_1 است در حالیکه θ_P و θ_P/θ^* نیز نقش مهمی بر عهده دارند. مقادیر N، توان آزمون را از طریق درجه آزادی $F(N.T-N-K;\lambda)$ تحت تأثیر قرار می‌دهند. در حالیکه مقدار N می‌تواند نقش مهمی در توان آزمون داشته باشد، مقدار K اثر ناچیزی دارد چراکه در دامنه ۱ تا ۵ قرار می‌گیرد. بر مبنای این یافته‌ها، برای تحلیل توان آزمون، K را برابر ۳ قرار می‌دهیم.

نمودار ۱ تابع چگالی $F(N.T-N-K;\lambda)$ تحت فرضیه H_1 و H_0 (به ترتیب در رنگ‌های سیاه و قرمز) را برای مقادیر متفاوت T، θ_P و θ^* هنگامی‌که $N=25$ است، نشان می‌دهد. اگر H_0 به طور صحیح برقرار باشد، آماره F از منحنی مشکی بدست می‌آید؛ اما اگر آن فرضیه نقض شود، از نمودار قرمز رنگ تبعیت می‌کند. خط عمودی آبی رنگ به مقدار بحرانی در سطح ۵ درصد اشاره دارد به نحوی که ناحیه زیر منحنی مشکی رنگ که از این مقدار بیشتر باشد (خاکستری سایه‌دار) بیانگر احتمال ۵ درصد است.

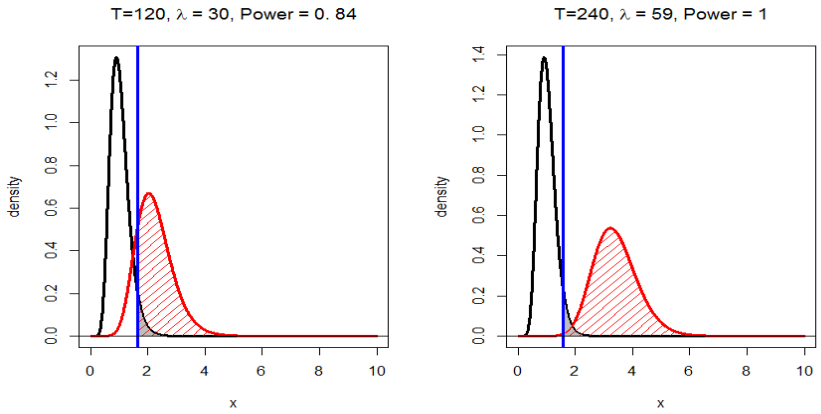
توان آزمون به وسیله ناحیه‌ای در زیر منحنی قرمز رنگ که بزرگتر از مقدار بحرانی است (قرمز سایه‌دار) نشان داده می‌شود. در بخش اول نمودار که حجم نمونه برابر ۶۰، θ_P برابر ۰/۲ و θ^* برابر ۱ است، توان آزمون برابر ۰/۹۶ می‌باشد؛ اما با افزایش حجم نمونه به ۱۲۰، توان آزمون به مقدار ۱ افزایش می‌یابد. بخش دوم نمودار موردی را که در آن $\theta_P=./3$ و $\theta^*=./6$ است نشان می‌دهد. در حالیکه توان آزمون با حجم نمونه ۱۲۰ برابر ۰/۸۴ به دست می‌آید، هنگامی‌که حجم نمونه ۲۴۰ است، توان برابر ۱ خواهد شد. این نتایج به طور واضح نشان می‌دهد هنگامی‌که حجم نمونه، ۲۴۰ و یا بیشتر باشد (همانگونه که در بسیاری از مطالعات پیشین که در ادامه بحث خواهد شد) توان آزمون GRS برابر با ۱ یا بسیار نزدیک به ۱خواهد بود. پیامدِ این موضوع آن است که فرضیه صفر هرگز پذیرفته نمی‌شود، چرا که آماره F از منحنی قرمز تولید نمی‌شود.

نمودار۱. مثالهایی از محاسبه توان برای آزمون GRS (K = 3, N = 25, α = ./05)

θ_p = ۰/۲, θ_p/θ^* = ۰/۲



θ_p = 0/3, θ_p/θ^* = 0/5



مأخذ: یافته‌های پژوهش

نکته کلیدی آن است که حتی هنگامی‌که H_0 به وسیله یک حاشیه ناچیز به لحاظ اقتصادی نقض می‌شود پدیده توان آزمون حداکثری، نیز اتفاق می‌افتد که این موضوع سبب می‌شود تا فرضیه صحیحِ H_0 به اشتباه رد شود. این رد فرضیه، ساختگی است چرا که H_0 به وسیله یک حاشیه ناچیز نقض می‌شود. این نکته قابل ذکر است که یک فرضیه صفر حداکثری در عمل نمی‌تواند برقرار باشد. مطابق با دلونگ و لانگ[1] (۱۹۹۲)، برای آزمون یک رابطه اقتصادی، می‌بایست فرضیه‌ها به صورت تقریبی مدنظر قرار گیرند. به

[1]. De Long and Lang

عنوان مثال یک بازار کارا به صورت کامل، غیر ممکن است. زیرا اگر قیمت‌ها به طور کامل تمامی اطلاعات در دسترس را منعکس نمایند، معامله‌گران هیچ انگیزه‌ای برای دستیابی به اطلاعات پرهزینه نخواهند داشت (گراسمن و استیگلیتز[۱]، ۱۹۸۰). در عوض، این موضوع که کارایی بازار نسبی است به طور وسیعی پذیرفته شده است (کمپل و همکاران[۲]، ۱۹۹۷)، و درجه کارایی نیز بستگی به شرایط بازار دارد (لو[۳]، ۲۰۰۴). بنابراین، هر آزمونِ کارایی بازار باید این سوال را مطرح نماید که آیا انحراف از کارایی بازار کامل به طور اقتصادی معنی‌دار است. با مدنظر قرار دادن آزمون GRS یک سبد کارای کامل با مقدار θ_P/θ^* دقیقاً برابر با ۱، عملاً وجود نخواهد داشت. بنابراین انجام آزمون باید با این ملاحظه صورت گیرد که آیا مقدار کمتر از یک برای θ_P/θ^* دارای اهمیت اقتصادی است یا خیر. ذکر این نکته ضروری است که هر گاه یک سبد کارا به لحاظ اقتصادی (که انحراف آن از $H_0: \theta_P/\theta^* = 1$ به لحاظ اقتصادی ناچیز است)، حجم نمونه‌ای بزرگ یا بسیار بزرگی را استفاده کند، ممکن نیست که آزمون GRS را با موفقیت بگذراند. این پدیده با آنچه اسپانوس (۲۰۱۷) « اشتباه رد شدن» فرضیه می‌نامد، بسیار مرتبط است.

۲-۳. تحلیل توان آزمون، سطح معنی‌داری بهینه و آزمون برابری احتمال

در این قسمت، ضمن تحلیل توان آزمون GRS، نشان داده می‌شود برای دستیابی به یک توازن منطقی میان احتمال خطای نوع اول و نوع دوم، می‌بایست حجم نمونه متوسطی در نظر گرفته شود. سپس پیشنهاد می‌شود که آزمون GRS، در سطح معنی‌داری بهینه و نه در سطح متعارف اجرا شود. در سطح بهینه، آزمون به خصوصیت مطلوبی دست می‌یابد که در آن احتمال‌های خطاهای نوع اول و دوم متوازن بوده (آزمون برابری احتمال) و همزمان، توان آزمون بالاتری به‌دست می‌آید.

گیبونز، راس و شانکن (۱۹۸۹)، توان آزمون GRS را با استفاده از توزیع $F(N, T-N-K; \lambda)$ با این فرض که نسبت شارپ ماکزیمم گذشته‌نگر برابر با نسبت شارپ ماکزیمم آینده‌نگر ($\hat{\theta}_P = \theta_P$) در رابطه (۳) است محاسبه نمودند. جهت تسهیل تحلیل توان آزمون، مطلوب است بیشترین مقادیر در نظر گرفته شده به عنوان مقادیر نماینده برای θ_P و نسبت کارایی بالقوه θ_P/θ^* را در اختیار داشته باشیم. کیم

---

[2]. Grossman and Stiglitz
[3]. Campbell et al.
[4]. Lo

و شمس‌الدین (۲۰۱۷) این مقادیر را از طریق به‌کارگیری پنجره‌های نمونه‌گیری غلتان[1] برای نمونه کامل برآورد نموده‌اند. هر ۳ مدل قیمت گذاری دارایی (CAPM، ۳ عاملی و ۴ عاملی) برای سبدهای ۲۵ تایی اندازه و ارزش دفتری به ارزش بازار برآورد شده‌اند. طول پنجره مورد استفاده برابر ۲۴۰ و ۳۶۰ ماه است که تعداد کافی مشاهدات برای تخمین صحیح را فراهم می‌آورد. جدول ۱ خلاصه آماره‌ها را (میانگین ،میانه، چارک اول، چارک سوم) برای مدل‌ها نشان می‌دهد.

در مدل CAPM بیشترین مقادیر معرف (نماینده) برای θ_P و θ_P/θ^* به ترتیب برابر ۰/۱۰ و ۰/۲۰ هستند. برای مدل سه عاملی، این مقادیر بطور تقریبی برابر ۰/۲۵ و ۰/۴۰ و در مدل چهار عاملی ۰/۳۵ و ۰/۵۰ هستند. نتایج بیان می‌دارند که هر دوی مقادیر θ_P و θ_P/θ^* با افزایش تعداد عوامل مورد استفاده در مدل‌های قیمت‌گذاری دارایی، به طور یکنواختی در حال افزایش هستند.

جدول ۱. بخش الف: حداکثر نسبت شارپ آینده‌نگر برای مدلهای عاملی (θ_P)

	چارک اول	میانه	میانگین	چارک سوم
T=				
مدل CAPM	0/083	0/126	0/121	0/151
مدل سه عاملی	0/211	0/265	0/253	0/286
مدل ۴ عاملی	0/332	0/378	0/356	0/408
T=				
مدل CAPM	0/108	0/124	0/122	0/140
مدل سه عاملی	0/205	0/245	0/243	0/267
مدل ۴ عاملی	0/289	0/373	0/346	0/387

مأخذ: گیبونز، راس و شانکن (۱۹۸۹)

جدول ۱. بخش ب: نسبت کارایی بالقوه (θ_P/θ^*)

	چارک اول	میانه	میانگین	چارک سوم
T=				

[1]. Rolling sub-sample window

CAPM مدل	0/167	0/192	0/19	0/218
مدل سه عاملی	0/294	0/394	0/419	0/533
مدل ۴ عاملی	0/440	0/527	0/530	0/642
T=				
مدل CAPM	0/195	0/217	0/215	0/234
مدل سه عاملی	0/306	0/494	0/431	0/512
مدل ۴ عاملی	0/416	0/62	0/56	0/655

مأخذ: گیبونز، راس و شانکن (۱۹۸۹)

۲-۳-۱- توابع توان آزمون GRS

همانگونه که در نمودار ۱ دیده می‌شود، توان آزمون به نحو چشمگیری با حجم نمونه T افزایش می‌یابد. برای مدل CAPM (K= 1) که مقادیر متعارف θ_P و θ_P/θ^* به ترتیب برابر ۰/۱ و ۰/۲ هستند، هنگامی‌که حجم نمونه به بزرگی ۳۴۲ و ۶۰۶ ماه است توان آزمون برابر ۱ می‌باشد. هنگامیکه حجم نمونه مابین ۶۰ و ۹۰ ماه است، توان آزمون متوسط می‌باشد. زمانی که حجم نمونه ۱۲۰ است، توان آزمون بسیار بزرگتر است و بسته به مقدار θ_P/θ^* می‌تواند برابر ۱ باشد. برای مدل سه عاملی که مقادیر عادی θ_P و θ_P/θ^* به ترتیب برابر ۰/۲۵ و ۰/۴۰ هستند، هنگامی‌که حجم نمونه ۳۴۲ و ۶۰۶ ماه است، توان آزمون برابر ۱ است در حالی که توان آزمون متوسط به بالا هنگامی می‌تواند حاصل شود که حجم نمونه در محدوده ۶۰ تا ۹۰ قرار می‌گیرد. مجدداً هنگامی‌که حجم نمونه برابر ۱۲۰ است، توان آزمون بالاست لیکن در محدوده‌ای از مقادیر θ_P/θ^* می‌تواند برابر ۱ باشد. مشاهدات مشابهی برای مدل ۴ عاملی وجود دارند.

این نتایج نشان می‌دهند که در حجم‌های نمونه بزرگ مانند ۱۲۰، ۳۴۲ و ۶۰۶ ماه اگر مقدار α در یک سطح متعارف همانند ۰/۰۵ قرار داشته باشد، عدم توازن شدیدی میان α و β (توان آزمون -۱) به وجود خواهد آمد. این بدان معنی است که مقدار α می‌تواند به طور فزاینده‌ای بزرگتر از مقدار β باشد. همانگونه که در نمودار ۱ بررسی شد این موضوع بدین معنی است که حتی هنگامی‌که فرضیه صفر به وسیله یک حاشیه ناچیز به لحاظ اقتصادی نقض می‌شود، احتمال رد آن در نمونه‌گیری‌های مکرر می‌تواند برابر ۱ باشد. نتیجتاً یک انحراف جزئی از فرضیه صفر که اهمیت اقتصادی اندکی دارد، در یک

سطح متعارف معنی‌داری رد می‌شود. بر این اساس هنگامی‌که توان آزمون به یک نزدیک می‌شود این موضوع برای هر آزمون آماری سازگار اتفاق می‌افتد (اسپانوس[1]، ۲۰۱۷).
هنگامی‌که N برابر ۲۵ است، تحلیل‌های قوی توان آزمون پیشنهاد می‌دهند که حجم نمونه T می‌بایست نزدیک به ۶۰ انتخاب شود. در واقع این موضوع با توصیه مطرح شده توسط گیبونز، راس و شانکن (۱۹۸۹) در خصوص انتخاب N سازگار است. بر اساس مطالعه آنان، اندازه‌های مناسب برای N به صورت تقریبی $\frac{1}{3}$ یا $\frac{1}{2}$ مقدار T است؛ یا به عبارتی هنگامی‌که $T = 60$ است، مقدار N در محدوده ۲۰ و ۳۰ مناسب خواهد بود. گرچه هیچ‌یک از مطالعات تجربی بعدی این توصیه را دنبال نکرده‌اند. در حجم نمونه ۱۲۰، توان آزمون می‌تواند بسیار بالا باشد، اما بسته به مقدار θ_P و θ_P/θ^* می‌تواند در اکثر موارد برابر ۱ یا بسیار نزدیک به ۱ باشد. تحلیل توان آزمون در این مطالعه بر مبنای قرار گرفتن N برابر ۲۵ استوار است چرا که مقداری است که بسیاری از مطالعات تجربی آن را لحاظ نموده‌اند.

۲-۳-۲- آزمون GRS برابری احتمال در سطح بهینه معنی‌داری

در آزمون فرضیه، قرار دادن مقدار α در سطح ۵ درصد مرسوم است، گرچه سایر سطوح مانند ۱ یا ۱۰ درصد نیز غالباً استفاده می‌شوند. با وجود این، همانطور که قبلاً ذکر شد، این سطوح متعارف، تصادفی هستند. در این خصوص، عنوان شده است که سطح معنی‌داری باید با توجه به حجم نمونه و توان آزمون انتخاب شود.[2] در مقاله حاضر، سطح معنی‌داری به صورت بهینه با ملاحظه توان آزمون یا مقدار β انتخاب می‌شود. در آزمون فرضیه، بده - بستان میان α و β به خوبی شناخته شده است. بر این اساس انتخاب یک مقدار بالاتر از $\alpha \in (0.1)$ لزوماً مقدار β را کاهش می‌دهد و بالعکس. برای یک مقدار داده شده α، می‌توانیم مقدار β را از توزیع آماره آزمون تحت فرضیه H_1 محاسبه کنیم. به وسیله ترسیم تمامی ترکیبات (α, β)، منحنی به دست می‌آید که لیمر (۱۹۷۸) از آن به خط قضاوت روشن‌بینانه (line of enlightened judgment) یاد می‌کند.سؤال آن است چگونه باید سطح بهینه را در میان این ترکیبات گسترده $(\alpha . \beta)$ انتخاب نمود. لیمر (۱۹۷۸) پیشنهاد می‌کند که انتخاب بهینه به وسیله حداقل نمودن زیان مورد انتظار از آزمون فرضیه انجام می‌شود. فرض کنید L_1 زیان حاصل از خطای نوع اول و L_2 زیان

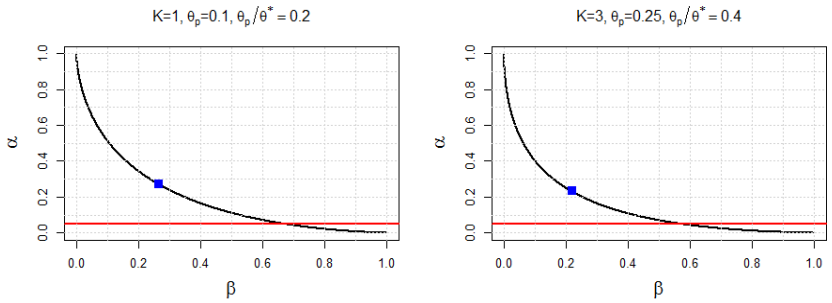
1. Spanos

۲. کیش (۱۹۵۹)، ارو (۱۹۶۰)، وینر (۱۹۶۲)، لیمر (۱۹۷۸)، مک کلوزکی و زیلیاک (۱۹۹۶)؛ و کیم و جی (۲۰۱۵) را ملاحظه کنید.

ناشی از خطای نوع دوم باشد. زیان انتظاری از آزمون فرضیه برابر است با $PaL_1 + (1-P)\beta L_2$، که $P = Prob(H_0)$ احتمال درست بودن فرضیه صفر را نشان می‌دهد. برای ساده‌سازی و هم‌جهت با آزمون برابری احتمال ارو[1] (۱۹۶۰) و نایک[2] (۱۹۶۹)، فرض می‌شود که $P = 0/5$ و L_1 برابر L_2 است. با فرض $P = 0/5$ ، H_0 و H_1 بطور برابر صحیح می‌باشند، در حالیکه فرض $L_1 = L_2$ بدین معنی است که خطای نوع اول و نوع دوم، اهمیت یکسانی دارند. تحت این فروض ساده‌شده، حداقل زیان مورد انتظار برابر $\alpha + \beta$ است.

در نمودار ۲، با فرض $T = 60$ و $N = 25$ و محدوده‌ای از مقادیر K و مقادیر نماینده θ_P و θ_P/θ^* که در جدول ۱ مشخص شده‌اند، خطوط قضاوت روشن‌بینانه برای آزمون GRS و نقاطی که زیان مورد انتظار $(\alpha + \beta)$ را حداقل می‌کنند، ترسیم شده‌اند. خط افقی قرمز رنگ، نقطه‌ای را همراه با مقدار β متناظر نشان می‌دهد که در آن سطح معنی‌داری در ۵ درصد قرار دارد. از آنجایی که توان آزمون هنگامی‌که T برابر ۶۰ است متوسط می‌باشد، سطح بهینه معنی‌داری بالاتر از ۵ درصد است. همانطور که مشاهده می‌گردد یک توازن نزدیک به کامل میان α و β در سطح بهینه حاصل شده است. این همان چیزی است که ارو (۱۹۶۰) از آن به عنوان «آزمون برابری احتمال» یاد می‌کند که در آن، وزنی برابر به خطاهای نوع اول و دوم داده می‌شود.

نمودار ۲. سطح بهینه معنی داری برای آزمون GRS (N=25, T=60)



مأخذ: یافته‌های پژوهش

[1]. Arrow
[2]. Naik

۳. مروری بر مطالعات تجربی پیشین

مبحث قیمت‌گذاری دارایی‌ها و اعتبار آنها از جمله شاخه‌های اصلی علم مالی محسوب می‌شود که مطالعات فراوانی پیرامون آن انجام شده است.بررسی‌ها حاکی از آن است که به رغم وجود مطالعات گسترده خارجی، تاکنون مطالعه‌ای در داخل کشور با تمرکز بر بکارگیری آزمون GRS جهت بررسی اعتبار تجربی مدل‌های قیمت‌گذاریانجام نشده است. با عنایت به این مشاهده و به دلیل اجتناب از طولانی شدن مقاله، در ادامه صرفاً روی آن دسته از مطالعات خارجی تمرکز می‌شود که از آزمون GRS برای بررسی اعتبار تجربی مدل‌های قیمت‌گذاری استفاده نموده‌اند.

جدول ۲. مطالعات خارجی آزمون GRS با استفاده از سبدهای اندازه و ارزش دفتری به بازار- بخش الف (الگوی تک عاملی)

محقق / محققین	تعداد سبدها N	طول دوره T	آماره GRS	P-Value	میانگین a	R^2
آهن، کوهارت و دیتمار (۲۰۰۹)	25	534	3/53	0	N/A	N/A
آرتمان، فینتر و کمف (۲۰۱۲)	16	510	3/26	0	N/A	N/A
ساکیسی، فبوزی و تان (۲۰۱۳)	25	252	2/22	0	0/34	0/69
چن و ژانگ (۲۰۱۰)	25	420	4/25	0	0/49	N/A
چن، مارکس و ژانگ (۲۰۱۰)	25	468	NR	0	0/43	N/A
کوهن، پولک و ولتناهو (۲۰۰۹)	20	180	1/34	0/14	N/A	0/22
فاما و فرنچ (۱۹۹۳)	25	342	1/91	0/004	0/26	۰/۷۸
فاما و فرنچ (۱۹۹۶)	25	366	2/76	0	0/28	0/77
فاما و فرنچ (۲۰۰۶)	6	498	9/18	0	0/35	0/52
فاما و فرنچ (۲۰۱۲)	25	245	4/07	0	0/21	0/81
گانگور و لوگر (۲۰۱۳)	100	456	2/33	0	N/A	N/A
هو، ژو و ژانگ (۲۰۱۴)	25	492	N/A	0	0/29	N/A
واندن (۲۰۰۴)	25	96	4/27	0	0/006	0/52
واندن (۲۰۰۶)	25	96	2/76	0/001	0/004	N/A

مأخذ: یافته‌های پژوهش

ادامه جدول ۲. مطالعات خارجی آزمون GRS با استفاده از سبدهای اندازه و ارزش دفتری به بازار- بخش ب (مدل سه عاملی)

محقق / محققین	تعداد سبدها N	طول دوره T	آماره GRS	P-Value	میانگین a	R^2
آهن، کوهارت و دیتمار (۲۰۰۹)	25	534	2/28	0	N/A	N/A

محقق / محققین	تعداد سبدها N	طول دوره T	آماره GRS	P-Value	میانگین a	R^2
آرتمان، فینتر و کمف (۲۰۱۲)	16	510	2/03	0/01	N/A	N/A
ساکیسی، فبوزی و تان (۲۰۱۳)	25	252	1/63	0	0/18	0/84
چن و ژانگ (۲۰۱۰)	25	420	3/08	0	0/21	N/A
چو و ژو (۲۰۰۶)	25	372	N/A	0	N/A	N/A
فاما و فرنچ (۱۹۹۳)	25	342	1/56	0/037	0/88	0/931
فاما و فرنچ (۱۹۹۶)	25	366	1/97	0/004	0/93	0/93
فاما و فرنچ (۲۰۱۲)	25	257	3/62	0	0/12	0/95
فاما و فرنچ (۲۰۱۵)	25	606	3/62	0	0/10	N/A
گری و جانسون (۲۰۱۱)	25	300	6/26	0	N/A	N/A
گریگوری، تاریان و کریستیدیس (۲۰۱۳)	25	363	1/27	0/18	0/11	0/87
گانگور و لوگر (۲۰۱۳)	100	456	2/10	0	N/A	N/A
هو، ژو و ژانگ (۲۰۱۴)	25	492	N/A	0	0/10	N/A
شانکن و ژو (۲۰۰۷)	25	480	3/15	0	N/A	N/A
واندن (۲۰۰۴)	25	96	4/27	0	0/006	0/52
واندن (۲۰۰۶)	25	96	17/08	0/001	0/003	N/A

مأخذ: یافته‌های پژوهش

همانگونه که از جدول فوق مشاهده می‌شود، مطالعات انجام شده در زمینه آزمون GRS در دو بخش مربوط به الگوی تک عاملی(CAPM) و سه عاملی(فاما و فرنچ) ارائه شده است. از جدول مشخص است که در اغلب مطالعات، تشکیل سبدهای ۲۵ تایی مبنا قرار گرفته است. همچنین هم در الگوی تک عاملی و هم سه عاملی، طول دوره‌ها بزرگ و اغلب بالاتر از ۱۰۰ ماه در نظر گرفته شده اند. بر این اساس نتایج بدست آمده از آزمون GRS، فرضیه صفر در اغلب این مطالعات پذیرفته نشده است. این بدان معنی است که الگوهای عاملی اشاره شده نتوانسته‌اند در دوره‌های زمانی مدنظر، بازدهی سهام را به خوبی تبیین نمایند.

۴. پایه‌های آماری و معرفی متغیرها

جهت ارزیابی اعتبار تجربی مدل‌های قیمت‌گذاری دارایی در بازار سهام ایران از داده‌های ماهانه مربوط به بازدهی شاخص کل و بازدهی تک تک سهام شرکت‌های حاضر در بازار بورس اوراق بهادار تهران در بازه زمانی فروردین ۱۳۷۹ تا اسفندماه ۱۳۹۵ (جمعاً ۲۰۴

ماه) استفاده شده است. این داده‌ها از طریق نرم افزار ره‌آورد نوین و سایت بورس اوراق بهادار تهران استخراج شده‌اند. پس از آن برای ایجاد سبدهای ۲۵ تایی موسوم به سبدهای فاما و فرنچ، داده‌های ماهانه بازدهی شرکتها در ۵ بخش بر اساس اندازه مرتب شده و سپس هر یک از این بخش‌ها نیز به ۵ قسمت بر اساس ارزش دفتری به بازار مرتب می‌شوند. پس از این تقسیم‌بندی، سبدهای ۵*۵ یا ۲۵ تایی بدست می‌آیند.

بر این اساس متغیرهای استفاده شده جهت انجام آزمون GRS به شرح زیر هستند:

$(R - R_F)$: مازاد بازده دارایی‌ها است که از تفاوت بازده سهام و نرخ بازده بدون ریسک بدست می‌آید. بازده سهام(R)، عبارت است از نسبت کل عایدی (ضرر) حاصل از سرمایه‌گذاری در یک دوره معین به میزان سرمایه‌ای که برای کسب این عایدی در اول همان دوره به‌کار گرفته شده است. عایدی یک سرمایه‌گذاری از دو محل است:

۱.تغییر در ارزش و قیمت اصل سرمایه مصروفه مانند قیمت سهام خریداری شده.

۲. سودهایی که در نتیجه سرمایه‌گذاری به اصل سرمایه تعلق می‌گیرد مانند سود نقدی سهام.

بازده سهام در این مقاله از طریق فرمول زیر محاسبه می‌گردد:

$$\text{بازده سهام} = \frac{\text{مزایای سهام جایزه + مزایای حق تقدم + (تفاوت قیمت سهم در اول و آخر سال مالی) + سود نقدی ناخالص هر سهم}}{\text{آخرین قیمت سهم در آخر سال مالی}}$$

نرخ سود سپرده (RF): با توجه به اینکه نرخ سود بانکی به صورت سالانه است و سایر متغیرهای مورد استفاده در این تحقیق به صورت ماهانه در نظر گرفته شده‌اند، برای همگن‌سازی داده‌ها و دستیابی به نتایج قابل اتکا، نرخ سود سپرده‌های سرمایه‌گذاری سالانه بانکی(استخراج شده از سایت بانک مرکزی) به نرخ‌های سود ماهیانه تبدیل و استفاده شده‌اند.

R_m: بازده کل شاخص بورس اوراق بهادار است، و از تغییرات قیمت و بازدهٔ نقدی پرداختی متأثر می‌شود[۱].

---

[۱]. این شاخص، تمامی شرکت‌های پذیرفته شده در بورس را دربردارد و شیوهٔ وزن‌دهی و محاسبهٔ آن همانند شاخص کل قیمت (TEPIX) است و تنها تفاوت میان آن دو در شیوهٔ تعدیل آن‌ها است. شاخص قیمت و بازدهٔ نقدی بورس تهران با استفاده از فرمول $TEDPIX_t = \frac{\sum_{i=1}^{n} p_{it} q_{it}}{RD_t} \times 100$ محاسبه می‌شود که در آن، p_{it} = قیمت شرکت iام در زمان t، q_{it} = تعداد سهام منتشره شرکت iام در زمان RD_t = پایهٔ شاخص قیمت و بازدهٔ نقدی در زمان t که در زمان مبدأ برابر $\sum p_{i0} q_{i0}$ بوده است.

($R_m - R_f$): عامل بازار یا مازاد بازده شاخص بازار است و از تفاوت بازده بازار و نرخ بازده بدون ریسک به دست می‌آید.

SMB: صرفه (عامل) اندازه است که از تفاوت بین میانگین بازده‌های پرتفوی سهام شرکت‌های کوچک (به لحاظ اندازه) و پرتفوی سهام شرکت‌های بزرگ محاسبه می‌شود.

HML: صرفه (عامل) ارزش است و از تفاوت میان میانگین بازده‌های پرتفوی سهام شرکت‌های با ارزش دفتری به بازار بالا و ارزش دفتری به بازار پایین محاسبه می‌شود.

بر این اساس فرضیه نخست تحقیق آن است که عبارات عرض از مبدأ معادلات مربوط به مدل قیمت‌گذاری دارایی‌های سرمایه‌ای و مدل سه عاملی فاما و فرنچ در بازار بورس اوراق بهادار ایران برای سبدهای ۲۵ تایی فاما و فرنچ، بصورت مشترک برابر صفر است. پذیرش این فرضیه بدین معنا خواهد بود که مدل قیمت‌گذاری، تغییرات سیستماتیک بازدهی سبدها را به طور مناسبی منعکس می‌نماید. طبق فرضیه دوم نیز استفاده از سطح بهینه معنی‌داری و لحاظ توان آماری در بکارگیری آزمون GRS، منجر به تغییر نتایج آزمون می‌گردد.

لازم به ذکر است که برای انجام آزمون GRS، از نرم افزار R استفاده شده است.

۵. نتایج آزمون‌های تجربی بررسی اعتبار مدل‌های قیمت‌گذاری

در این بخش با توجه به یافته‌های بخش‌های پیشین، آزمون GRS را یک مرتبه بدون ملاحظات توان آزمون و مرتبه دیگر با در نظر گرفتن ملاحظات مربوط به توان آزمون برای مدلهای CAPM و سه عاملی فاما و فرنچ در بازار بورس اوراق بهادار اجرا می‌نماییم. بر این اساس برای بخش اول، طول دوره‌های بزرگ مانند ۱۲۰، ۱۶۸ و ۲۰۴ ماه و برای بخش دوم طول دوره مناسب (متوسط) یعنی ۶۰ ماهه را انتخاب و آزمون می‌کنیم. جهت دستیابی به استدلالی قوی‌تر، چندین دوره ۶۰ ماهه از کل حجم نمونه (۲۰۴ماه) را انتخاب و آزمون می‌نماییم. دوره‌های بزرگ شامل ۱۲۰ ماه ابتدای دوره، ۱۲۰ ماه انتهای دوره، ۱۶۸ ماه ابتدای دوره، ۱۶۸ ماه انتهای دوره و کل دوره (۲۰۴ ماه) می‌باشند. دوره‌های ۶۰ ماهه نیز عبارتند از: ۶۰ ماه اول از کل حجم نمونه، ۶۰ ماه دوم از کل حجم نمونه، ۶۰ ماه سوم از کل حجم نمونه، ۶۰ ماه انتهایی از کل حجم نمونه و یک دوره ۶۰ ماهه دیگر از انتهای حجم کل نمونه.

قسمت اول نتایج مربوط به نتایج آزمون GRS بدون لحاظ توان آزمون و قسمت دوم با لحاظ توان آزمون هستند. برای هر یک از طول دوره‌های بزرگ و متوسط دوره ۵ نمونه

متفاوت در نظر گرفته شده است که برای الگوی CAPM و سه عاملی فاما و فرنچ جداگانه برآورد می‌گردند. بدین ترتیب ۱۰ نمونه برای هر یک از طول دوره‌های بزرگ و متوسط آزمونمی‌شوند. جداول ۳ و ۴ خلاصه‌ای از نتایج مربوط به هر یک از آنها را نشان می‌دهد. همانطور که از جدول ۳ ملاحظه می‌شود آزمون GRS در حجم نمونه بزرگ، برای ۵ طول دوره متفاوت و ۵ بازه زمانی مختلف بر روی الگوهای قیمت‌گذاری دارایی CAPM و سه عاملی فاما و فرنچ اجرا شده است. به عنوان مثال نتایج مربوط به طول دوره ۱۲۰ ماهه در الگوی CAPM نشان می‌دهند که آماره آزمون GRS برابر ۲/۷۰ است. این در حالی است که مقدار احتمال بدست آمده برابر ۰/۰۰۰۳ است. لذا فرضیه H_0 در سطح معنی-داری متعارفپذیرفته نمی‌شود. این بدان معنی است که مدل CAPM در بازه زمانی مورد نظر قادر به تبیین عامل ریسک سیستماتیک نیست. با توجه به مقدار β، توان آزمون تقریباً برابر ۱ است و دلیل رد فرضیه صفر نیز ریشه در بالا بودن توان آزمون دارد.

مثالی دیگر از این جدول، طول دوره ۱۶۸ ماهه در الگوی سه عاملی فاما و فرنچ است. با توجه به مقدار احتمال مربوط به آماره GRS، فرضیه صفر پذیرفته نمی‌شود. این بدان معنی است که مدل سه عاملی فاما و فرنچ در بازه زمانی مورد نظر قادر به تبیین تمامی عوامل ریسک سیستماتیک نیست. با توجه به مقدار β، توان آزمون تقریباً برابر ۱ است. انتخاب حجم نمونه بزرگ منجر به توان آزمون بالا شده است که در نهایت، رد فرضیه صفر را به دنبال داشته است.

جدول ۳. نتایج آزمون GRS برای حجم نمونه بزرگ بدون ملاحظات مربوط به توان آزمون

نوع الگو	دوره	طول دوره	GRS.stat	GRS.pval	میانگینa	R^2	Power	نتیجه آزمون فرضیه
الگوی CAPM	1379-1388	120	2/70	3e-04	1/49	0/11	0/999	عدم پذیرش H0
	1379-1392	168	2/24	0.002	1/50	0/17	0/998	عدم پذیرش H0
	1386-1395	120	1/70	0.03	1/49	0/11	0/987	عدم پذیرش H0
	1382-1395	168	2/28	0.001	1/26	0/22	0/998	عدم پذیرش H0
	1379-1395	204	2/91	3.55e-05	1/40	0/20	0/9999	عدم پذیرش H0
الگوی سه عاملی	1379-1388	120	3/50	9e-05	1/50	0/20	0/9999	عدم پذیرش H0
	1379-1392	168	2/86	7.32e-05	1/30	0/29	0/9999	عدم پذیرش H0
	1386-1395	120	2/19	0.004	0/75	0/48	0/998	عدم پذیرش H0
	1382-1395	168	3/02	3.03e-05	1/05	0/37	0/9999	عدم پذیرش H0
	1379-1395	204	3/61	5.82e-07	1/22	0/32	0/9999	عدم پذیرش H0

مأخذ: یافته‌های پژوهش

جدول ۴. نتایج آزمون GRS برای حجم نمونه‌های متوسط با ملاحظات مربوط به توان آزمون

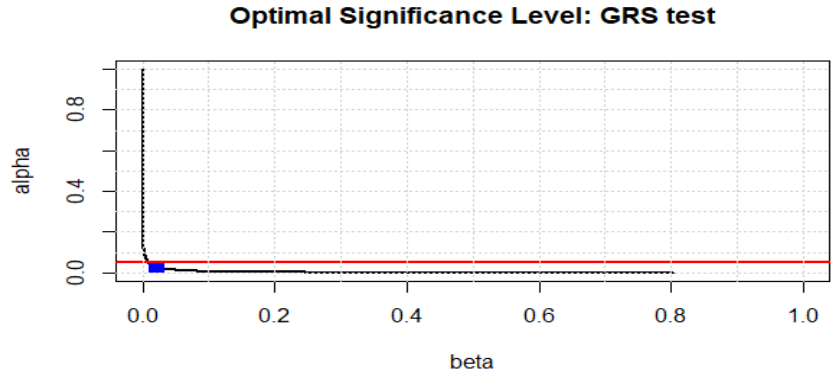
نوع الگو	دوره	طول دوره	GRS stat	GRS pval	opt sig	opt crit	میانگین a	R^2	Power	نتیجه آزمون فرضیه
الگوی CAPM	1379-1383	60	2/1	0/02	0/01	2/23	2/12	0/08	0/998	پذیرش H0
	1384-1388	60	2/74	0/003	0/005	2/55	1/16	0/17	0/99	عدم پذیرش H0
	1389-1393	60	1/16	0/33	0/07	1/73	1/13	0/36	0/91	پذیرش H0
	1390-1394	60	1/42	0/17	0/04	1/86	1/30	0/41	0/96	پذیرش H0
	1391-1395	60	1/72	0/07	0/03	2/03	1/67	0/42	0/99	پذیرش H0
الگوی سه عاملی	1379-1383	60	2/08	0/02	0/01	2/25	1/99	0/17	0/99	پذیرش H0
	1384-1388	60	3/05	0/001	0/003	2/74	1/28	0/30	0/99	عدم پذیرش H0
	1389-1393	60	1/21	0/29	0/06	1/77	0/63	0/52	0.92	پذیرش H0
	1390-1394	60	1/43	0/17	0/04	1/89	0/79	0/55	0/99	پذیرش H0
	1391-1395	60	2/06	0/02	0/01	2/23	0/91	0/55	0/99	پذیرش H0

مأخذ: یافته‌های پژوهش

بطور کلی نتایج بدست آمده در خصوص الگوی CAPM نشان می‌دهند که در تمامی ۵ دوره منتخب، فرضیه صفر در سطح معنی‌داری متعارف (۵ درصد) رد می‌شود. همچنین مقادیر اندک میانگین عرض از مبدأ بدست آمده، حاکی از آن است که فرضیات رد شده با حاشیه اندکی به لحاظ اقتصادی نقض شده‌اند. این موضوع با مبانی نظری کاملاً منطبق است. نتایج مربوط به الگوی سه عاملی نیز نشان می‌دهند که در هر ۵ دوره، فرضیه صفر در سطح معنی‌داری متعارف رد می‌شود. همانطور که مشاهده می‌شود قدرت توضیح‌دهندگی در الگوی سه عاملی، افزایش محسوسی نسبت به الگوی CAPM دارد. نتایج بدست آمده از این قسمت با یافته‌های بخش‌های پیشین کاملاً منطبق بوده و نشان می‌دهد که عدم توجه به ملاحظات مرتبط با توان آزمون منجر به رد فرضیه صفر در تمامی دوره‌ها می‌شود.

جدول ۴ نتایج آزمون GRS را برای حجم نمونه‌های متوسط با ملاحظات مربوط به توان آزمون نشان می‌دهد. بخش اول جدول مربوط به نتایج الگوی CAPM است. به عنوان نمونه، نتایج مربوط به ۶۰ ماهه انتهایی نشان می دهد که آماره آزمون GRS برابر ۱/۷۲ است. لذا با توجه به کمتر بودن مقدار آماره از مقدار بهینه آن یعنی ۲/۰۳، فرضیه صفر رد نمی‌شود. این بدان معنی است که مدل CAPM در بازه زمانی مورد نظر قادر به تبیین عامل ریسک سیستماتیک است. همچنین سطح بهینه معنی‌داری (α) تقریباً برابر ۰/۰۲ است و مقدار β نیز برابر همین عدد است. این موضوع به معنای تأیید «آزمون برابری احتمال» می‌باشد. نمودار ۳ سطح بهینه معنی‌داری مربوط به آزمون فوق را نشان می‌دهد.

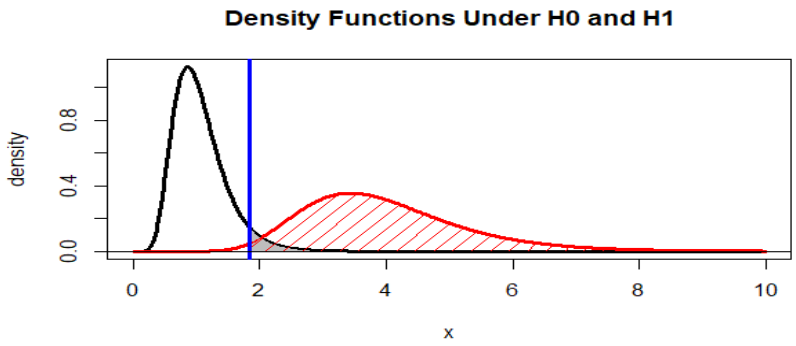
نمودار ۳. سطح بهینه معنی‌داری برای آزمون GRS طی دوره زمانی فروردین ۱۳۹۱ تا اسفند ۱۳۹۵



مأخذ: یافته‌های پژوهش

توان آزمون ۰/۹۹ است که مقدار بالایی است و نشان می‌دهد علاوه بر تأیید فرضیه صفر، آزمون GRS توان مناسبی نیز دارد. نمودار ۴ تابع چگالی مربوط به آزمون فوق را نشان می‌دهد.

نمودار ۴. تابع چگالی برای آزمون GRS طی دوره زمانی فروردین ۱۳۹۱ تا اسفند ۱۳۹۵

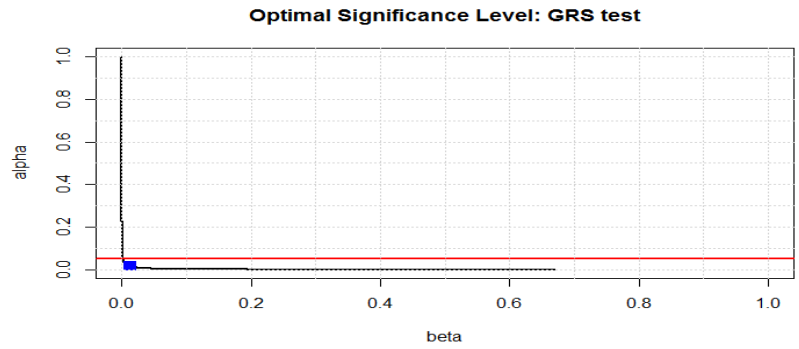


مأخذ: یافته‌های پژوهش

همانطور که ملاحظه می‌گردد تنها در یک مورد از نتایج این الگو (نمونه ۶۰ ماهه مربوط به سالهای ۱۳۸۴-۱۳۸۸)، فرضیه صفر رد شده و در ۴ مورد دیگر، فرضیه صفر را نمی‌توان رد کرد. در سطح معنی‌داری متعارف نیز در سه مورد فرضیه صفر رد نمی‌شود.

قسمت دوم جدول نتایج الگوی سه عاملی فاما و فرنچ را نشان می‌دهد. به عنوان مثال در طول دوره ۶۰ ماهه اول، با توجه به کمتر بودن مقدار آماره از مقدار بهینه، فرضیه صفر رد نمی‌شود. این بدان معنی است که مدل سه عاملی فاما و فرنچ در بازه زمانی مورد نظر قادر به تبیین تمامی‌عوامل ریسک سیستماتیک است. همچنین سطح بهینه معنی‌داری (α) برابر ۰۱/. است. نمودار ۶ سطح بهینه معنی‌داری مربوط به آزمون فوق را نشان می‌دهد.

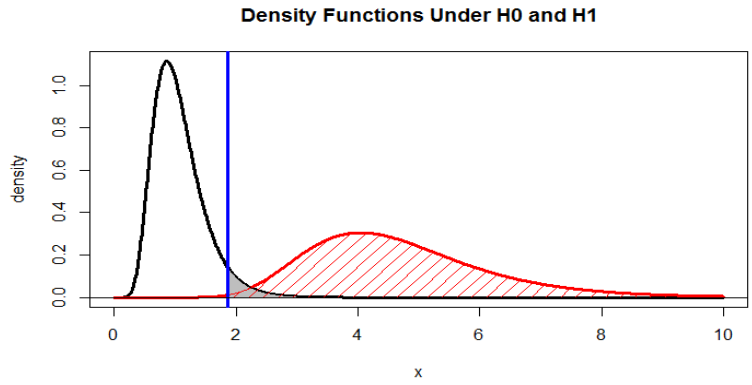
نمودار ۵. سطح بهینه معنی‌داری برای آزمون GRS طی دوره زمانی فروردین ۱۳۷۹ تا اسفند ۱۳۸۳



مأخذ: یافته‌های پژوهش

با توجه به مقدار β، توان آزمون برابر ۹۹/. می‌باشد که مقدار بالایی است و نشان می‌دهد علاوه بر تأیید فرضیه صفر، آزمون GRS دارای توان مناسبی نیز می‌باشد. نمودار ۶ تابع چگالی آزمون را نشان می‌دهد.

نمودار ۶. تابع چگالی برای آزمون GRS طی دوره زمانی فروردین ۱۳۷۹ تا اسفند ۱۳۸۳



مأخذ: یافته‌های پژوهش

در این الگو، تنها فرضیه صفر مربوط به دوره ۶۰ ماهه دوم رد شده و در سایر موارد فرضیه صفر رد نمی‌شود. همچنین در سطح معنی‌داری متعارف نیز فرضیه‌های مربوط به دو حجم نمونه (۱۳۸۹-۱۳۹۳ و ۱۳۹۴-۱۳۹۰) پذیرفته می‌شوند. در مقایسه با الگوی CAPM، مدل سه عاملی قدرت توضیح‌دهندگی به مراتب بیشتر و مقادیر میانگین عرض از مبدأ کمتری را دارند. نکته قابل توجه در خصوص هر دو الگو، توان بسیار بالای آزمون در تمامی

فرضیاتی است که مورد پذیرش قرار گرفته‌اند. نتایج حاصل شده در این قسمت نیز با یافته‌های مطرح شده توسط گیبونز، راس و شانکن (۱۹۸۹) در خصوص انتخاب حجم نمونه متناسب با تعداد سبدها و همچنین تحلیل‌های مرتبط با توان آزمون کاملاً منطبق است.

همچنین جداول ۵ و ۶ به ترتیب حداکثر نسبت شارپ برای حجم نمونه بزرگ بدون ملاحظات مربوط به توان آزمون و حداکثر نسبت شارپ برای حجم نمونه‌های متوسط با ملاحظات مربوط به توان آزمون را نشان می‌دهند.

جدول ۵. حداکثر نسبت شارپ برای حجم نمونه بزرگ ب

دون ملاحظات مربوط به توان آزمون

نوع الگو	دوره	طول دوره	$\widehat{\theta}_P$
الگوی CAPM	1379-1388	120	0/06556
	1379-1392	168	0/18792
	1386-1395	120	0/10736
	1382-1395	168	0/09142
	1379-1395	204	0/11689
الگوی سه عاملی	1379-1388	120	0/30472
	1379-1392	168	0/31373
	1386-1395	120	0/34015
	1382-1395	168	0/29189
	1379-1395	204	0/27516

مأخذ: یافته‌های پژوهش

جدول ۶. حداکثر نسبت شارپ برای حجم نمونه‌های متوسط

با ملاحظات مربوط به توان آزمون

نوع الگو	دوره	طول دوره	$\widehat{\theta}_P$
الگوی CAPM	1379-1383	60	0/34653
	1384-1388	60	0/22553
	1389-1393	60	0/25857
	1390-1394	60	0/15555
	1391-1395	60	0/12003
الگوی سه عاملی	1379-1383	60	0/43087
	1384-1388	60	0/46634
	1389-1393	60	0/.32875
	1390-1394	60	0/24811
	1391-1395	60	0/31335

مأخذ: یافته‌های پژوهش

متوسط حداکثر نسبت شارپ برای حجم نمونه‌های بزرگ در الگوی CAPM برابر ۰,۱۱ و در الگوی سه عاملی برابر ۰,۳۰ است. این اعداد بسیار نزدیک به مقادیر در نظر گرفته شده به عنوان مقادیر نماینده برای حداکثر نسبت شارپ در الگوهای فوق می‌باشند. همانطور که در بخش ۲-۳ اشاره شد در مدل CAPM بیشترین مقادیر معرف (نماینده) برای θ_P برابر ۰/۱۰ و در مدل سه عاملی، بطور تقریبی برابر ۰/۲۵ است.

بطور کلی نتایج آزمون‌ها نشان می دهند که در طول دوره های بزرگ و بدون در نظر گرفتن ملاحظات مربوط به توان آزمون، هیچ‌یک از مدل‌های قیمت‌گذاری دارایی (مدل قیمت گذاری دارایی‌های سرمایه‌ای و مدل سه عاملی فاما و فرنچ) در بازه‌های زمانی مختلف در بازار سهام ایران پذیرفته نمی‌شوند. لیکن هنگامی که ملاحظات مربوط به توان آزمون در خصوص آنها مدنظر قرار گیرند، اعتبار تجربی هر دو مدل قیمت گذاری دارایی در بازار سهام ایران، در بازه‌های زمانی مختلف تأیید می گردد.

۶. جمع‌بندی و نتیجه‌گیری

بورس اوراق بهادار به عنوان یکی از ارکان بازار متشکل سرمایه، نقش مهمی در تجمیع پس‌اندازها و هدایت آنها در راستای تأمین مالی بلندمدت بنگاه‌ها ایفا می‌کند. موفقیت در این اهداف نیز مستلزم عملکرد کارای بورس اوراق بهادار است. کارایی نیز به نوبه خود در گرو آن است که ابعاد مختلف بازار و مکانیزم‌های اثرگذار بر آن شناسایی شده و در اختیار مشارکت‌کنندگان در بازار قرار گیرد تا سطح دانش و آگاهی آنان ارتقا یابد.

از مهم‌ترین موضوعاتی که از دیرباز حساسیت بسیاری از پژوهشگران و محققین حوزه اقتصاد مالی را برانگیخته است، قیمت‌گذاری دارایی‌ها است. بر کسی پوشیده نیست که قیمت دارایی‌ها در بازار بورس، عمدتاًبه اخبار اقتصادی واکنش نشان می‌دهد و تحت تأثیر طیف گسترده‌ای از رویدادهای پیش‌بینی نشده قرار می‌گیرد که تأثیر برخی از این رویدادها، نسبت به دیگر رویدادها نیز بیشتر است. مروری بر ادبیات قیمت‌گذاری نشان می‌دهد که مدل قیمت‌گذاری دارایی سرمایه‌ای (CAPM) و مدل‌های چند عاملی (که مشهورترین آنها مدل ۵ عاملی فاما و فرنچ (۲۰۱۵) است)، بیش از سایر مدل‌ها مورد استقبال قرار گرفته‌اند. در صورت تأیید مدل‌های عاملی به لحاظ تجربی در بازار بورس اوراق بهادار تهران می‌توان به این نتیجه مهم دست یافت که عوامل ریسک سیستماتیک، قابلیت توضیح‌دهندگی بازدهی را داشته و لذا سرمایه‌گذاران می‌توانند با توجه به این عوامل ریسک، اقدام به تشکیل سبد سهام نموده و بازدهی مورد انتظار خود را بدست آورند.

به رغم آنکه این مدل‌ها شهرت فراوانی یافته و بطور گسترده‌ای در تحقیقات آکادمیک و کاربردهای تجربی مورد پذیرش قرار گرفته‌اند، یافته‌های برخی مطالعات حکایت از آن دارد که این مدل‌ها نمی‌توانند بصورت کامل و مطلوب، ریسک سیستماتیک دارایی‌ها را منعکس نمایند.بررسی دقیق‌تر این مطالعات نشان می‌دهد که از آزمون GRS که یک آزمون والد است برای آزمون فرضیه‌ها استفاده شده است. اما نکته‌ای که در اکثر مطالعات مورد غفلت واقع شده است به‌کارگیری آزمون GRS بدون در نظر گرفتن ویژگیهای توان آماری است که در این مقاله، تلاش شده است تا این کاستی مرتفع گردد.

در این راستا به منظور دقیق‌تر و قابل فهم‌تر شدن آزمون فرضیه، در این مقاله پیشنهاد گردید که آزمون GRS با ملاحظه صریح توان آن اجرا شود. لذا سطح معنی‌داری می‌بایست به عنوان تابعی کاهنده از حجم نمونه تعدیل گردد. به علاوه، تحلیل توان آزمون آشکار می‌سازد هنگامی‌که تعداد سبد دارایی‌های مورد آزمون برابر ۲۵ است، با در نظر

گرفتن حجم نمونه برابر ۶۰، توان آزمون مطلوب بوده و توازن منطقی میان احتمال‌های خطای نوع اول و دوم فراهم می‌آورد. در این اندازه حجم نمونه، آزمون GRS در سطح معنی‌داری بهینه اجرا می‌شود که به آنچه ارو (۱۹۶۰) آزمون برابری احتمال نامیده است، منجر می‌شود. آزمون برابری احتمال، میان خطاهای اول و دوم بی‌تفاوت است و توان آزمون بالایی را نیز ارائه می‌دهد. تحلیل تجربی در این مقاله ابتدا به وسیله برآورد مدل‌های CAPM و سه عاملی فاما و فرنچ برای سبدهای ۲۵ تایی اندازه و ارزش دفتری به بازار با استفاده از داده‌های ماهانه بازار بورس اوراق بهادار تهران از سالهای ۱۳۷۹ تا ۱۳۹۵ انجام شد. در ابتدا طول دوره‌های بزرگ انتخاب شده‌اند. یافته‌های این مقاله در خصوص بکارگیری آزمون GRS در بازار سهام ایران حاکی از آن است که فرضیه صفر در هر دو مدل قیمت‌گذاری دارایی در سطح معنی‌داری متعارف توسط آزمون GRS رد می‌شود.

با در نظر گرفتن ویژگی‌های مربوط به توان آماری آزمون، حجم نمونه‌های ۶۰ تایی مورد استفاده قرار گرفت و آزمون GRS در سطح بهینه معنی‌داری اجرا گردید. بر این اساس مشخص شد که فرضیه صفر در اکثر موارد برای هر دو مدل قیمت‌گذاری دارایی CAPM و سه عاملی فاما وفرنچ رد شود.

بر این اساس می‌توان نتیجه‌گیری نمود هنگامی‌که آزمون GRS با ملاحظات صریح در خصوص توان آماری اجرا شود، شواهد قوی اعتبار تجربی را برای مدل‌های قیمت‌گذاری دارایی فراهم می‌آورد. لذا ممکن است نیاز به تحقیقات مداوم در خصوص شناسایی عوامل ریسک سیستماتیک جدید وجود نداشته و در عوض تلاش‌ها در جهت توجه به توان آماری آزمون در مدلهای قیمت‌گذاری دارایی معطوف شوند.

فهرست منابع:

فبوزی، فرانک، اقتصاد مالی، ترجمه رضا طالبلو و بهاره عریانی (۱۳۹۴)، چاپ اول، انتشارات سمت.

Arrow, K. (1960), Decision theory and the choice of a level of significance for the t-test, *In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, *Standford University Press*, 2: 70-78.

Campbell, J.Y., Lo, A.W. & MacKinlay, A.C. (1997), The Econometrics of Financial Markets, *Princeton University Press*, Princeton.

Carhart, M. (1997), On persistence in mutual fund performance, *Journal of Finance*, 52: 57–82.

Cochrane, J. (1991), Production-Based Asset-pricing and the Link Between Stock Returns and Economic Fluctuations, *The Journal of Finance*, 46: 209–237.

Cochrane, J. (1996), A Cross-Sectional Test of an Investment-Based Asset-pricing Model, *Journal of Political Economy*, 104: 572-621.

De Long, J.B. & Lang, K. (1992), Are all Economic Hypotheses False? *Journal of Political Economy*, 100 (6): 1257-1271.

Douglas, G.W. (1967), Risk in the equity markets: An empirical appraisal of market efficiency, *Yale Economic Essays*, 9: 3-48.

Fama, E. F. & French, K. R. (1993), Common Risk Factors in the Returns on Stocks and Bonds. *Journal of financial economics*, 33: 3-56.

Fama, E. F. & French, K. R. (2015a), A five-factor asset-pricing model, *Journal of Financial Economics*, 116: 1-22.

Gibbons, M. R., Stephen A. R. and Shanken, J. (1989), A Test of the Efficiency of a Given Portfolio, *Econometrica*, 57: 1121–1152.

Grossman, S.J. & Stiglitz, J.E. (1980), On the impossibility of informationally efficient markets, *The American Economic Review*, 70 (3): 393–408.

Harvey, C. R. (2017), Presidential Address: The Scientific Outlook in Financial Economics, *Journal of Finance*, 72(4): 1399-1440.

Harvey, C. R., Y. Liu. & H. Zhu. (2016), …and the Cross-Section of Expected Returns, *Review of Financial Studies*, 29:5–68.

Hausman, J.A. (1978), Specification tests in econometrics, *Econometrica*, 46(6): 1251–1271.

Hou, K., Xue, C. & Zhang, L. (2015), Digesting anomalies: An investment approach, *Review of Financial Studies*, 28: 650-705.

Kim, J. H. & Ji, P. (2015), Significance Testing in Empirical Finance: A Critical Review and Assessment, *Journal of Empirical Finance*, 34: 1-14.

Kish, L. (1959), Some statistical problems in research design, *American Sociological Review*, 24: 328-338.

Leamer, E. (1978), *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley, New York.

Lintner, J. (1965), The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *Review of Economics and Statistics*, 47: 13-37.

Lo, A.W. (2004), The adaptive markets hypothesis: Market efficiency from an evolutionary perspective, *Journal of Portfolio Management*, 30: 15–29.

MacKinnon, J. G. (2002), Bootstrap inference in Econometrics, *Canadian Journal of Economics*, 35 (4): 615-644.

McCloskey, D. & Ziliak, S. (1996), The Standard Error of Regressions, *Journal of Economic Literature*, 34: 97–114.

Naik, U. D. (1969), The Equal Probability Test and its Applications to Some Simultaneous Inference Problems, *Journal of the American Statistical Association*, 64(327) : 986-998.

Sharpe, W. F. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, 19(3): 425-442.

Spanos, A. (2017), Mis-specification Testing in Retrospect, *Journal of Economic Surveys*, forthcoming, doi: 10.1111/joes.12200.

Team, R. C. (2016), R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Winer, B. J. (1962), *Statistical Principles in Experimental Design*, New York, McGraw-Hill.