بررسی قانون بن‌فورد در بورس اوراق بهادار تهران

محمدرضا رستمی[1*]، اعظم جعفری درگیری[2]

۱. استادیار مدیریت مالی دانشگاه الزهرا، تهران، ایران

۲. کارشناس ارشد مدیریت مالی دانشگاه الزهرا، تهران، ایران

دریافت: ۹۰/۵/٤ پذیرش: ۹۱/۹/۸

چکیده

یکی از الزامات رقابت سالم در بازار سرمایه دسترسی همه مشارکت‌کنندگان به اطلاعات صحیح می‌باشد. کیفیت اطلاعات در تصمیم‌گیری مفید بوده وموجب می‌شود تصمیماتی برای تخصیص بهینه منابع اتخاذ شود که این مسأله به کارایی هر چه بیشتر اقتصادی منتهی می‌شود. اطلاعات صریح منافعی چون اتخاذ افق بلندمدت از سوی سرمایه‌گذاران، بهبود دسترسی به سرمایه جدید، هزینه سرمایه کمتر، مدیریت معتبر و مسئولانه‌تر و درنهایت قیمت بالاتر سهام و منفعت بیشتر سهام‌داران را به‌دنبال خواهد داشت.

این تحقیق به بررسی قانون بن‌فورد به‌عنوان شاخص ارزیابی کیفیت داده در بورس اوراق بهادار تهران می‌پردازد. داده‌های مورد استفاده، رقم اول قیمت‌های پایانی سهام و مقادیر مربوط به شاخص‌های موجود می‌باشد که از آغاز سال ۱۳۷۸ تا پایان سال ۱۳۸۸ مورد مطالعه قرار گرفته‌اند. شرکت‌هایی به‌عنوان نمونه موضوع پژوهش انتخاب شدند که در چارک بالای شرکت‌های پر معامله به لحاظ تعداد روزهای معامله قرار داشتند. آزمون‌های آماری مورد استفاده در این پژوهش،آزمون کولموگروف- اسمرینوف و آزمون نیکویی برازش بود.

نتیجه تحقیق نشان داد که داده‌های مورد بررسی از قانون رقم اول تبعیت نمی‌کند. همچنین نتیجه مشابهی در بررسی رقم اول داده‌ها در دوره‌های قبل و بعد از حجم مبنا و درخصوص بررسی توزیع رقم اول قیمت‌های پایانی روزانه تک‌تک شرکت‌های موجود در نمونه حاصل شد. عدم مطابقت داده‌های موجود در بورس از این قانون، نتیجه متأثر شدن از تبدیلاتی است که روی آنها صورت می‌گیرد. همچنین می‌تواند به نابهنجاری‌های رفتاری بازار سرمایه نیز مربوط باشد.

کلیدواژه‌ها: قانون رقم اول- قیمت‌های پایانی روزانه سهام - شاخص کیفیت داده.

---

* نویسندهٔ مسؤول مقاله: E-mail: rostami1973@yahoo.com

۱- مقدمه

بورس اوراق بهادار به‌عنوان یک بازار رسمی، نقش مهمی را در هدایت منابع مالی و به‌دنبال آن رشد و توسعه اقتصاد ایفا می‌کند. مهم‌ترین ابزار تصمیم‌گیری در این نهاد اقتصادی، داده‌ها و اطلاعات می‌باشند. وجود اطلاعاتی صحیح و قابل اتکا تأثیر به‌سزایی در ارزیابی شرایط بازار دارد. در نتیجه افزایش توان و سرعت وسایل پردازش اطلاعات، تمایل به استفاده از مدل ریاضی قدیمی به نام قانون بن‌فورد افزایش یافته است. این قانون بیان می‌کند رقم اول داده‌هایی که به طور طبیعی به وقوع می‌پیوندند، از توزیع یکنواخت تبعیت نمی‌کنند بلکه توزیع لگاریتمی خاصی را دنبال می‌کنند. به این ترتیب احتمال وقوع ارقامی با ارزش‌های پایین، بیشتر از سایر ارقام می‌باشد. از این رو قانون حاضر چارچوب ارزشمندی را برای تخمین صحت ارزش‌های داده ارائه می‌کند و به شناسایی الگوهای ناهنجار در مجموعه‌های داده می‌پردازد.

۲- اهمیت مسأله پژوهش

بازار سرمایه به‌عنوان جزئی از نظام مالی در اقتصاد با هدایت مازاد وجوه به سمت سایر بخش‌های اقتصاد، نقش تعیین کننده‌ای در رشد و توسعه اقتصادی یک کشور دارد. یکی از الزامات رقابت سالم در بازار سرمایه دسترسی همه مشارکت‌کنندگان به اطلاعات صریح می‌باشد. اطلاعات منتقل شده در تصمیم‌گیری مفید بوده و اگر انتقال آن به شکل درستی باشد، آن گاه سبب می شود تا تصمیماتی برای تخصیص بهینه منابع اتخاذ شود که این مسأله به کارایی هر چه بیشتر اقتصادی منتهی می‌شود [٤]. وجود بازارهای سرمایه کارا شرایطی را فراهم می‌کند که سرمایه‌گذاران در محیطی مطمئن‌تر فعالیت کنند که این امر می‌تواند خطرپذیری آنها را تا حدودی پوشش دهد. اطلاعات صریح منافع متعددی برای شرکت‌ها نیز به دنبال خواهد داشت. منافعی چون سرمایه‌گذاری بلندمدت‌تر از سوی سرمایه‌گذاران، بهبود دسترسی به سرمایه جدید، هزینه سرمایه کمتر، مدیریت معتبرتر، مسئول‌تر و درنهایت قیمت بالاتر سهام و منفعت بیشتر سهام‌داران. زمانی که افزایش قیمت سهام براساس خلق واقعی ارزش باشد نه براساس انتظارات مدیریت از سود یا سود گزارش شده، اندازه بازارها و قدرت نقدشوندگی آنها افزایش پیدا کرده و نه فقط برای سهام‌داران

بلکه برای جامعه نیز به طور کلی ثروت بیش‌تری ایجاد می‌شود. با توجه به مراتب فوق، بررسی شفافیت داده‌ها و اطلاعات به‌عنوان پایه‌های تصمیم‌گیری در بازار سرمایه ضروری به نظر می‌رسد که تحقیق حاضر به این مهم می‌پردازد.

۲-۱- قانون بن‌فورد

قانون بن‌فورد[۱] یا قانون رقم معنی‌دار[۲] مشاهده تجربی است که بیان می‌کند ارقام معنی‌دار در بسیاری از مجموعه داده‌های عددی - که به طور طبیعی- به وقوع می‌پیوندند، به طور یکنواخت توزیع نشده‌اند. این ارقام توزیع لگاریتمی خاصی را دنبال می‌کنند، به این صورت که احتمال وقوع رقم یک به‌عنوان رقم اول در بسیاری از مجموعه داده‌ها بیشتر از سایر ارقام است و احتمال وقوع رقم‌های ۲ تا ۹ به تناسب اینکه ارزش آن بیشتر می‌شود، کاهش پیدا می‌کند. این قانون احتمال وقوع رقم اول و رقم دوم را به وسیله عبارت‌های زیر بیان می‌کند:

$$P(d_i) = \log(۱ + d_i^{-})$$
$$i = ۱,۲,\dots,۹$$
$$p(d_i) = \sum_{k=۱}^{k=۹} \log(۱ + (۱۰k + d)^{-۱})$$
$$i = ۰\dots۹$$

با تعمیم توابع ذکر شده، شکل کلی قانون بن‌فورد به شرح زیر بیان می‌شود:

$$p(D_۱ = d_۱,\dots, D_K = d_k) = \log[۱ + (\sum_{i=۱}^{k} d_i \times ۱۰^{k-i})]$$
$$\forall k \in z$$
$$d_۱ \in \{۱,۲,\dots,۹\}$$
$$d_j \in \{۰,\dots,۹\}$$
$$J = ۲,\dots,k$$

این قانون برای اولین بار در سال ۱۸۸۱ توسط سیمون نیوکمپ[۳] ریاضیدان و اخترشناس آمریکایی معرفی شد. او مشاهده کرد صفحات اولیه جداول لگاریتمی نسبت به صفحات بعدی کهنه‌تر بود. به دنبال این مشاهده وی چنین استنباط کرد که استفاده‌کنندگان بیشتر در جستجوی اعداد چند رقمی هستند که با ارقام کوچک (۱، ۲ یا ۳) شروع می‌شود تا اعداد

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

چندرقمی که با ارقام ٤ الی۹. او از این موضوع نتیجه گرفت که توزیع رقم معنی‌دار به این صورت است که $p(firstsignificantdigit = d) = \log(1 + d^{-1})$ ، امّا از آن جایی که نیوکمپ شواهد تجری و دلایل تئوریکی در این خصوص بیان نکرد، این قانون در این مرحله از تاریخ صرفاً در حد حدس و گمان باقی ماند.

قانون رقم معنی‌دار شش دهه بعد در سال ۱۹۳۸ توسط فرانک بن‌فورد فیزیکدان مؤسسه جنرال الکتریک که در ظاهر از کار نیوکمپ آگاه نبود، دوباره کشف شد. بن‌فورد با بررسی توزیع ارقام معنی‌دار در بیست مجموعه داده- که مشتمل بر ۲۰,۰۰۰ داده بود و از منابع متعددی نیزجمع‌آوری شده بود- توانست حمایت تجربی ویژه‌ای را از این قانون فراهم کند. از آن به بعد این حجم گسترده از داده‌ها، هسته حمایت‌های تجربی را- آنچه که هم اکنون با عنوان قانون‌بن فورد شناخته شده است - شکل دادند. بعد از بیان تجربی قانون بن‌فورد، تلاش‌های چشمگیری در جهت اثبات این موضوع که قانون مذکور به‌طور ذاتی کاملاً ریاضی و قطعی است، صورت پذیرفته است. تئودور پی‌هیل در سال ۱۹۹۵ اثبات کرد که این قانون نسبت به مقیاس و پایه ثابت است[٥]. همچنین پیترونرو[٤] و همکاران در سال ۲۰۰۱ بیان داشتند که یکی از مبانی اولیه توزیع بن‌فورد به دلیل فرایندهای ضربی است. آنها همچنین بیان کردند که بسیاری از سیستم‌های عددی نظیر قیمت های بازار سهام از سیستم‌های دینامیکی به‌وسیله فرایندهای براونی[٥] پیروی نکرده بلکه براساس فرایندهای ضربی[٦] عمل می‌کنند. در ذیل قانون بن‌فورد با استفاده از فرایندهای ضربی و در مبنایی غیر از ۱۰ اثبات شده است[۱۰].

$$N\left(t + 1\right) = \varepsilon N(t)$$

ε متغیر تصادفی[۷] است. در فضای لگاریتمی داریم:

$$\ln N\left(t + 1\right) = \ln \varepsilon + \ln N(t)$$

اگر t به سمت بینهایت میل کند، توزیع احتمال $\ln(N)$ به توزیع یکنواخت متمایل می‌شود، اما با یک انتقال در فضاخطی داریم:

$$\int Prob(lnN)d(lnN) = \int Cd(lnN) = C\int \frac{۱}{N}dN$$

C عدد ثابتی است. از آن جایی $\int_{۰}^{\infty} \frac{۱}{N}dN$ تعریف نشده است، توزیع رقم اول N در مبنای ۱۰، n است که با عبارت زیر بیان می‌شود:

$$Prob(n) = \frac{\int_{n}^{n+۱} \frac{۱}{N}dN}{\int_{۱}^{b} \frac{۱}{N}dN} = \frac{ln\frac{n+۱}{n}}{lnb} = \log_b(۱+\frac{۱}{n})$$

و این توزیع احتمالی است که توسط بن‌فورد بیان شد (جدول ۱).

جدول۱ احتمال وقوع رقم یک تا نه به‌عنوان رقم اول و براساس قانون بن‌فورد

D	۱	۲	۳	٤	٥	٦	٧	٨	۹
P	٪۳۰/۱	٪۱۷/٦	٪۱۲/٥	٪۹/٧	٪٧/۹	٪٦/٧	٪٥/٨	٪٥/۱	٪٤/٦

۳- پیشینه پژوهش

لی[۸] در سال ۱۹۹٥ قانون بن‌فورد را در خصوص یازده یک روزه شاخص صنعتی داوجونز و شاخص s & p به کارگرفت. داده‌های شاخص داوجونز مربوط به سال‌های ۱۹۹۰-۱۹۹۳ و داده‌های شاخص s&p مربوط به سال‌های ۱۹۹۳- ۱۹۲٦بود. او نتیجه گرفت که داده‌های موضوع تحقیق از قانون بن‌فورد تبعیت می‌کنند و بیان کرد آنها می‌توانند به گروهی از اعداد بدون روابط شناخته شده تقسیم شوند [۷، صص۲٦۳-۲٧۹] .

لی زیپنگ[۹] و همکاران در سال ۲۰۰٤ مجموعه‌ای از داده‌های مختلف از جمله قیمت‌های سهام را مورد بررسی قرار دادند. داده‌های بررسی‌شده در این قسمت از قانون رقم اول تبعیت نکرد. آنها بیان کردند فرایندهای ضربی ساده قادر به تولید مجدد قیمت سهام - که از فرایندهای ضربی پیچیده پیروی می‌کنند- نیستند و قیمت‌های سهام ممکن است منحنی عمیق‌تر از منحنی لگاریتمی بن‌فورد را دنبال کنند. آنها همچنین بیان داشتند که توزیع های

احتمالی متنوع[10] و آشفته[11] خیلی بهتر با قانون بن‌فورد متناسب هستند و از این قانون می‌توان برای نشان دادن تصادفی[12] بودن جهان استفاده کرد [۱۳].

کرازا[13] و همکاران در سال ۲۰۰۸ م. در تحقیق مشابه تحقیق لی توزیع فراوانی رقم اول قیمت و بازده روزانه شاخص s&p500 در دوره ۱۴ آگوست ۱۹۹۵ تا ۱۷ اکتبر ۲۰۰۷ را مورد بررسی قرار دادند. با استفاده از تست کای- دو در دوسطح معنی‌داری ۱ و ۵ درصد، داده‌های موضوع تحقیق از قانون بن‌فورد تبعیت نکردند. آنها متوجه شدند بازده‌ها نسبت به قیمت‌ها تطابق نزدیک‌تری با قانون بن‌فورد داشت. آنها نزدیک بودن توزیع فراونی رقم اول بازده‌ها به توزیع بن‌فورد را به دلیل فرایندهای ضربی برشمردند که روی قیمت‌ها انجام شده است. همچنین بیان کردند که از تجزیه و تحلیل برمبنای قانون رقم اول می‌توان به منظور شناسایی ناهنجاریهای بازار سرمایه نظیر مرزهای روان‌شناختی و وجود سطوح مقاومتی استفاده کرد [۱].

نیگرینی[14] از اولین محققانی بود که از قانون بن‌فورد به‌عنوان پایه ریاضی تجزیه و تحلیل رقمی و با هدف کشف تقلب در داده‌های حسابداری استفاده کرد. نیگرینی در سال ۲۰۰۵م. همزمان با وقوع ورشکستگی شرک انرون از بزرگ‌ترین شرکت‌های فعال در زمینه انرژی در سال ۲۰۰۱م. به دلیل نقض مقررات حسابرسی و منحل شدن شرکت آرتور اندرسون- که کار حسابرسی شرکت انرون را به عهده داشت- به علت عدم افشای اطلاعات، داده‌های مربوط به درامد و سود هر سهم شرکت‌های منتشر شده در وال‌استریت ژورنال را طی سال‌های ۲۰۰۱ و ۲۰۰۲م. مورد بررسی قرارداد. آزمون رقم اول و دوم داده‌های مربوط به درامد، نشان از تمایل مدیریت به روند کردن رو به بالای آمار مربوط به این ارقام را داشت و در مورد سود هر سهم نیز نتیجه مشابهی حاصل شد [۹].

گانزالس گارسیا و پاستور[15] در سال ۲۰۰۹م قانون بن‌فورد را به‌عنوان شاخص ارزیابی کیفیت داده در خصوص داده‌های مربوط به متغیرهای اقتصادی در چهار بخش اقتصادی اعم از بانکداری، تراز پرداخت‌ها، حساب‌های ملی و آمار مالی بخش دولتی طی دوره ۲۰۰۷- ۱۹۴۵ مورد بررسی قرار دادند. کشورهای موضوع تحقیق براساس طبقه‌بندی اقتصاد جهانی در سال ۲۰۰۸ در شش گروه کشورهای پیشرفته، اروپای مرکزی و شرقی، مستقل

مشترک‌المنافع و مغولستان، خاورمیانه و آسیایی در حال توسعه و کشورهای واقع در نیمکره غربی و آفریقا بود. آزمون استفاده شده در این پژوهش، آزمون نیکویی برازش بود. نتیجه تحقیق، دوازده مورد عدم تطابق با قانون بن‌فورد را نشان داد. آنها همچنین نتایج این پژوهش را با مقیاس کیفیت داده- که در Rosc[۱۶] منتشر شده بود- مقایسه کردند. دو نتیجه‌ای که حاصل شد، نشان داد نباید تطابق نداشتن بعضی از داده‌ها را با قانون بن‌فورد، دلیل کیفیت ضعیف داده‌های اقتصاد کلان تعبیر کرد، بلکه این عدم تطابق ممکن است ناشی از انتقال‌های ساختاری باشد که در سطح اقتصاد کلان برخی کشورها روی داده است ، از این رو موضوع حاضر نیز باید مورد توجه قرار گیرد [۲].

دکتر غلامحسین مهدوی و مصطفی کاظم‌نژاد در سال ۱۳۸۸ به بیان اهمیت استفاده از قانون بن‌فورد در حسابرسی مستمر[۱۷]پرداخته‌اند. آنها استفاده از قانون بن‌فورد را هم به لحاظ کاهش خطر نمونه‌گیری، هزینه منفعت استفاده ازآن و توانایی این قانون در بهبود کیفیت اطلاعات در حسابرسی مستمر الزامی دانستند[۳ ، صص۳۸۷-۳۹۶] .

۴- روش تحقیق

جامعه آماری این تحقیق، کل شرکت‌های حاضر در بورس اوراق بهادار بین سال‌های ۱۳۷۸-۱۳۸۸ بود که براساس روزهای معاملاتی مرتب و چارک اول، پر معامله‌ترین شرکت‌ها شامل۱٤٥ شرکت انتخاب و قیمت‌های روزانه معاملاتی آنها مبنای تحلیل قرار گرفتند.

به منظور محاسبه توزیع فراوانی ارقام اول داده‌های مربوط به قیمت‌های سهام و شاخص از براوردگرهای کرنل[۱۸] و هیستوگرام[۱۹] استفاده شد.[۱۲] سپس با استفاده از آزمون‌های نیکویی برازش و کولموگروف - اسمرینوف توزیع رقم اول داده‌ها با فراوانی پیش‌بینی شده به‌وسیله قانون رقم اول مقایسه شد. آماره آزمون نیکویی برازش در خصوص آزمون رقم اول به صورت زیر نوشته می‌شود:[٦].

$$\chi_{۸}^{۲} = N\sum_{i=۱}^{k=۹}\frac{(\hat{\theta_i} - f_i)^{۲}}{f_i}$$

که در فرمول بالا $\theta_i^{\wedge}$ فراوانی مشاهده شده، f_i فراوانی مورد انتظار و N تعداد مشاهدات می باشد. همچنین درجه آزادی توزیع $k-1$ می‌باشد. از آن جایی که این آزمون نسبت به نمونه حساس می‌باشد، از آزمون کولموگروف – اسمرینوف نیز استفاده شد که آماره آن به صورت زیر نوشته می‌شود:

$$T_K = (D_N^+ + D_N^-)$$
$$D_N^+ = \sup[H_d - P_d]$$
$$D_N^- = \sup[P_d - H_d]$$

که hd بر فراوانی نسبی مشاهده شده، pd احتمال تجمعی و N تعداد مشاهدات دلالت دارد.

۴-۱- فرضیه‌های تحقیق

۱. داده‌های مالی (قیمت‌های سهام) شرکت‌های حاضر در بورس اوراق بهادار تهران از قانون بن‌فورد تبعیت می‌کند.

۲.داده‌های شاخص‌های بورس اوراق بهادار تهران (شاخص قیمت، شاخص قیمت و بازده نقدی، شاخص صنعت و شاخص پنجاه شرکت برتر) از قانون بن‌فورد تبعیت می‌کند.

به منظور آزمون فرضیه‌ها، چنان‌چه آماره آزمون نیکویی برازش از مقدار بحرانی آن در جدول بزرگ‌تر باشد، فرض H_0 رد می‌شود و اگر کوچکتر باشد، دلیلی بر رد فرض صفر وجود ندارد. همچنین در نرم‌افزار متلب آماره تست کولموگوروف – اسمیرونوف به صورت صفر و یک نشان داده می‌شود. اگر میزان این آماره صفر شود ، به این معنا است که بین فراوانی‌های مشاهده شده و فراوانی‌های پیش‌بینی شده به‌وسیله قانون بن‌فورد تفاوتی وجود ندارد؛ به عبارت دیگر دلیلی بر رد فرض صفر نیست، بالعکس چنان‌چه میزان این آماره یک شود، یعنی فراوانی‌های مشاهده شده متفاوت از فراوانی‌های پیش‌بینی شده به‌وسیله قانون رقم اول است و فرض H_0 رد می‌شود.

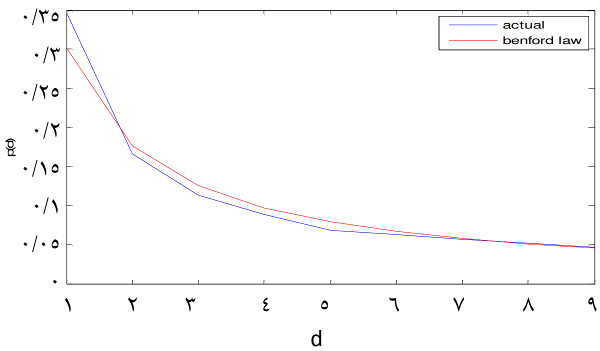
۴-۱-۱- آزمون فرضیه اول (رقم اول قیمت‌های سهام شرکت‌ها)

قیمت‌های پایانی روزانه سهام شرکت‌های موجود در نمونه – که طی سال‌های ۱۳۷۸-۱۳۸۸ مورد بررسی قرار گرفت- مشتمل بر ۲۱۲۰۱۹ مشاهده است. پس از محاسبه رقم اول

داده‌های موضوع پژوهش، با استفاده از نرم‌افزار متلب، آزمون کولموگوروف-اسمیرونوف در خصوص رقم اول داده‌های مربوط به قیمت‌های پایانی روزانه شرکت‌های موجود در نمونه انجام شد. مقدار آماره این آزمون عدد یک محاسبه شد که نشان می‌دهد فراوانی داده‌های بررسی شده متفاوت از فراوانی پیش‌بینی شده به‌وسیله قانون رقم اول می‌باشد؛ به عبارت دیگر قیمت‌های پایانی روزانه سهام موجود در نمونه از قانون بن‌فورد تبعیت نمی‌کند. همچنین مقدار آماره آزمون نیکویی برازش۲۳۳۲ تخمین زده شده است که بزرگ‌تر از مقدار بحرانی آن در سطح معنی‌داری ۵ درصد با درجه آزادی ۸ می‌باشد. نتیجه این آزمون نیز نشان می‌دهد که فرض صفر در این سطح معنی‌داری رد می‌شود. نتایج حاصل از آزمون‌ها در (جدول ۲) خلاصه شده است.

جدول ۲ آزمون رقم اول قیمت سهام

رقم اول	تعداد ارقام مشاهده شده	فراوانی ارقام مشاهده شده	فراوانی پیش بینی شده توسط قانون بن‌فورد
۱	۷۳٤۲٤	۰/۳٤٦۳	۰/۳۰۱
۲	۳٥۲٥۲	۰/۱٦٦۳	۰/۱۷٦۱
۳	۲۳۹۱٦	۰/۱۱۲۸	۰/۱۲٤۹
۴	۱۸۸۲۲	۰/۰۸۸۸	۰/۰۹٦۹
۵	۱٤٤٦۹	۰/۰٦۸۲	۰/۰۷۹۲
۶	۱۳۲٦۷	۰/۰٦۲٦	۰/۰٦٦۹
۷	۱۲۱۲۱	۰/۰٥۷۲	۰/۰٥۸
۸	۱۰۸۷۹	۰/۰٥۱۳	۰/۰٥۱۲
۹	۹۸٦۹	۰/۰٤٦٥	۰/۰٤٥۸
تعداد مشاهدات	۲۱۲۰۱۹		
KS	رد فرضیه صفر		
Z^2	۲۳۳۲		



نمودار ۱ مقایسه توزیع فراوانی رقم اول قیمت‌های سهام با توزیع بن‌فورد

۴-۱-۲- آزمون فرضیه دوم (رقم اول شاخص‌های بورس)

شاخص‌هایی که این تحقیق به مطالعه و بررسی آنها پرداخت، شامل شاخص قیمت، شاخص قیمت و بازده نقدی، شاخص صنعت و شاخص پنجاه شرکت فعال طی دوره ده ساله پژوهش بود. بررسی الگوی رقم اول این گروه از داده‌ها نشان داد در خصوص شاخص قیمت و بازده نقدی رقم سه، در مورد شاخص صنعت رقم هفت و در مورد شاخص پنجاه شرکت فعال، رقم دو بیشترین فراوانی را در میان سایر ارقام دارا بود. میزان آماره آزمون کولموگروف – اسمرینوف برای هر یک از شاخص‌های بررسی شده در سطح معنی‌داری ۵ درصد، عدد یک تخمین زده شد؛ به این معنا که فراوانی مشاهده شده رقم اول، داده‌های مربوط به هر یک از شاخص‌ها متفاوت از فراوانی پیش‌بینی به‌وسیله قانون بن‌فورد می‌باشد و فرض H_0 در سطح معنی‌داری ۵ درصد رد می‌شود. با استفاده از آزمون نیکویی برازش نیز نتیجه مشابهی حاصل شد؛ زیرا مقدار آماره آزمون نیکویی برازش در هر گروه از داده‌ها، از مقدار بحرانی آن در جدول و در سطح معنی‌داری ۵ درصد با درجه آزادی ۸ بزرگ‌تر بود. نتایج به شرح جدول ۳ است.

جدول۳ آزمون رقم اول مقادیر شاخص قیمت

فراوانی پیش‌بینی شده به‌وسیله قانون بن‌فورد	فراوانی ارقام مشاهده شده	تعداد ارقام مشاهده شده	رقم اول
۰/٤۱۸۹	۰/۳۰۱	۱۱۱۳	۱
۰/۱۱۱۸	۰/۱۷٦۱	۲۹۷	۲
۰/۰۹۵۲	۰/۱۲٤۹	۲۵۳	۳
۰/۵۸۷	۰/۰۹٦۹	۱۵٦	۴
۰/۰۳۹۹	۰/۰۷۹۲	۱۰٦	۵
۰/۰۰۹٤	۰/۰٦٦۹	۲۵	۶
۰/۰۰٦٤	۰/۰۵۸	۱٤	۷
۰/۰٤۸٦	۰/۰۵۱۲	۱۲۹	۸
۰/۲۱۱۱	۰/۰٤۵۸	۵٦۱	۹
۲٦۵۷			تعداد مشاهدات
عدم مطابقت			KS
۲۱۳۷			x^2_8

جدول۴ آزمون رقم اول مقادیر شاخص بازده نقدی و قیمت

فراوانی پیش‌بینی شده به‌وسیله قانون بن‌فورد	فراوانی ارقام مشاهده شده	تعداد ارقام مشاهده شده	رقم اول
۰/۳۰۱	۰/۰۸۵۵	۲۲٤	۱
۰/۱۷٦۱	۰/۲۷۸٦	۷۳۰	۲
۰/۱۲٤۹	۰/۳٤۹۲	۹۱۵	۳
۰/۰۹٦۹	۰/۱۳۸۲	۳٦۲	۴
۰/۰۷۹۲	۰/۰٤۲٤	۱۱۱	۵
۰/۰٦٦۹	۰/۰٤۸۵	۱۲۷	۶
۰/۰۵۸	۰/۰۲٦۳	٦۹	۷
۰/۰۵۱۲	۰/۰۱۹۱	۵۰	۸
۰/۰٤۵۸	۰/۰۱۲۲	۳۲	۹
۲٦۲۰			تعداد مشاهدات
عدم مطابقت			KS
۱۸۸۲			

جدول۵ آزمون رقم اول شاخص صنعت

فراوانی پیش‌بینی شده به‌وسیله قانون بن‌فورد	فراوانی ارقام مشاهده شده	تعداد ارقام مشاهده شده	رقم اول
۰/۴۱۸۹	۰/۰۲۶۴	۳۷۵	۱
۰/۱۱۱۸	۰/۰۲۹۸	۴۲۳	۲
۰/۰۹۵۲	۰/۰۰۹	۱۲۸	۳
۰/۰۵۸۷	۰/۰۱۳	۱۸۵	۴
۰/۰۳۹۹	۰/۰۰۲	۲۹	۵
۰/۰۰۹۴	۰/۰۰۹۴	۱۳۳	۶
۰/۰۰۶۴	۰/۰۴۳۴	۶۱۶	۷
۰/۰۴۸۶	۰/۰۲۱۹	۳۱۱	۸
۰/۲۱۱۱	۰/۰۳۲۲	۴۵۷	۹
۲۶۵۷			تعداد مشاهدات
عدم مطابقت			**KS**
۱۸۰۰			

جدول۶ آزمون رقم اول شاخص پنجاه شرکت برتر

فراوانی پیش بینی شده به‌وسیله قانون بن‌فورد	فراوانی ارقام مشاهده شده	تعداد ارقام مشاهده شده	رقم اول
۰/۴۱۸۹	۰/۰۵۶۵	۱۲۳	۱
۰/۱۱۱۸	۰/۲۱۱۹	۴۶۱	۲
۰/۰۹۵۲	۰/۱۰۳۴	۲۲۵	۳
۰/۰۵۸۷	۰/۰۵۲۴	۱۱۴	۴
۰/۰۳۹۹	۰/۱۹۳	۴۲۰	۵
۰/۰۰۹۶	۰/۱۲۵	۲۷۲	۶
۰/۰۰۶۴	۰/۱۲۹۱	۲۸۱	۷
۰/۰۴۸۶	۰/۰۷۲۲	۱۵۷	۸
۰/۲۱۱۱	۰/۰۵۶۵	۱۲۳	۹
۲۱۷۶			تعداد مشاهدات
عدم مطابقت			**KS**
۱۱۸۰			

۵- نتیجه‌گیری

قانون بن‌فورد به‌عنوان شاخص ارزیابی کیفیت داده به بررسی الگوهای نابهنجار در مجموعه‌ای از داده‌ها می‌پردازد. تحقیقات صورت گرفته در خصوص بررسی خواص این قانون بیان می‌کند که وجود قانون رقم اول به دلیل فرایندهای تصادفی می‌باشد، اما در خصوص حالت عکس آن تحلیلی صورت نپذیرفته است. تمرکز اصلی تحقیق حاضر روی قیمت‌های پایانی روزانه سهام شرکت‌های نمونه- که به طور طبیعی به وقوع می‌پیوندد- می‌باشد. همچنین با استفاده از خاصیت ثابت بودن قانون بن‌فورد نسبت به مقیاس شاخص‌های سهام نیز مورد بررسی قرارگرفت. مطاقبت نداشتن این مجموعه داده‌ها از قانون رقم اول نشان می‌دهد که قیمت‌های سهام و شاخص‌های سهام موجود در بورس، در نتیجه تبدیلاتی که روی آنها صورت می‌پذیرد، توزیع متفاوتی از توزیع لگاریتمی بن‌فورد را دنبال می‌کنند [۱۱]. از این رو می‌تواند مشابه نتیجه تحقیق کرازا، الرو وزورازی به نابهنجاری رفتاری در بازار مالی مربوط باشد که بحث در این خصوص و سایر دلایل مطابقت نداشتن این مجموعه داده‌ها از قانون رقم اول، مستلزم تحقیقات بیشتری در آینده می‌باشد.

۵-۱- پیشنهادات برای تحقیقات آتی

ارتباط بین تبعیت داده‌ها ازقانون بن‌فورد با کارایی بازار سرمایه بررسی شود.

از تجزیه و تحلیل برمبنای قانون بن‌فورد به منظور شناسایی نابهنجاری‌های رفتاری بازار سرمایه استفاده شود[۸].

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

۶- پی‌نوشت‌ها

1. Benford Law
2. First Significant Digit
3. Simon Newcomb
4. Pietronero
5. Brownian Process
6. Multiplicative Process
7. Stochastic Variable
8. Ley

9. Li Zhipeng
10. Diversified
11. Chaotic
12. Randomness
13. Coraza
14. Nigrini
15. Jesus Gonzalez-Garcia and Pastor
16. Report on the Observance of Standards and Codes
17. Continuous Auditing
18. Kernel Density
19. Histogram

۷- منابع

[۱] جمشیدی ف.، فلاح محسن‌خانی ز. ،کشاورز م.؛ براورد چگالی داده‌ها و آماره‌ها؛ مرکز آمار ایران،۱۳۸٤.

[۲] فراهانی ن.،؛ ملزومات رشد و توسعه بازار سرمایه ایران؛ مرکز مطالعات تکنولوژی دانشگاه شریف، ۱۳۸۵.

[۳] مهدوی غ. ،کاظم‌نژاد م.؛"استفاده از قانون بن‌فورد در حسابرسی مستمر"، *فصلنامه انجمن حسابداری ایران*، ۱۳۸۸.

[٤] آقایی م.، مرادی م.؛" ارائه الگویی برای بهینه‌سازی طبقه‌بندی اطلاعات دارایی‌های ثابت در ترازنامه شرکت‌های صنایع غذایی بورس اوراق بهادار تهران"؛ *فصلنامه پژوهش مدیریت در ایران*، ۱۳۸۹.

[5] Corazza M., Ellero A., Zorazi A.;"Checking financial markets via Benford's Law"; *Ca'Foscari University of Venice*,۲۰۰۸ .

[6] Gonzalez J., Gonzalo P.;" Benford's law and macroeconomic data quality"; *International Monetary Fund*, 2009.

[7] Gottwald Georg A., Nicol M.;" On the nature of Benford's law"; *Physica A303*, 2002.

[8] Hill.Theodor P.; "The first digit phenomenon"; *American Scientist*, 1998.

[9] Hill Theodore P.; " A statistical derivation of the significant-digit Law"; Statistical Science , November,Vol.10, No.4, 1995.

[10] Kreiner W.A. ;"On the newcomb-benford law"; *Z. Naturforsh*, 618-622, 2003.

[11] Ley E.;" On the peculiar distribution of the U.S indexes'digits"; *American Statistician*,Vol. 50, No. 4, 1996.

[12] Marc J.K.,De.Ceuster,Geert Dhaene,Tom.Schatteman;" On the hypothesis of psychological Barriers in Stock Market and Benford's law"; *Journal of Empirical Finance* **5** , 1998.

[13] Nigrini J. Mark;"An assessment of the change in the incidence of earnings management around the enrone-anderson episode"; *Review of Accounting and Finance*, 2005.

[14] Pietronero L.,Tossati E.,Tossati Vespignani A.;" Explaining the uneven distribution of number in nature:The laws of Benford and Zipf"; *Physica A 293*,296-304, 2001.

[15] Scott P. D. , Fasli M.; "Benford's law : An empirical investigation and a novel explanation"; *CSM Technical Report 349*.

[16] Silverman B. E.; " Density estimation for statistics and data analysis" ; *Published in Monographs on Statistics and Applied Probab*, 1986.

[17] Zhipeng Li., Lin Cong, Huajia Wang ;" *Discussion on Benford's Law and its Application*; 2004.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی