خطاهای روش‌شناسی اقتصادسنجی در بررسی نظریه‌های اقتصادی

مجتبی رستمی (نویسنده مسئول)
دکتری اقتصاد
mojtabarostami1364@yahoo.com
علی فرهادیان
استادیار گروه مدیریت دانشگاه کاشان
farhadian@kashanu.ac.ir
محمدنبی شهیکی تاش
دانشیار دانشکده اقتصاد و مدیریت دانشگاه سیستان و بلوچستان
mohammad_tash@yahoo.com
ابوطالب کاظمی
دکتری اقتصاد، سرپرست گروه سرمایه‌گذاری اداره کل امور اقتصادی و دارایی ایلام
am.kazemi1988@yahoo.com
تاریخ دریافت: ۱۳۹۹/۰۵/۲۷ تاریخ پذیرش: ۱۳۹۹/۱۲/۱۶

چکیده:
نظریه‌های علمی شیوه‌ای از شناخت انسان می‌باشند. اختلاف اساسی این شیوه با سایر شیوه‌ها آزمون‌پذیری ادعاهای آن است. آزمون‌پذیری کمک می‌کند تا شناخت علمی از خطا پیراسته شود. در اقتصاد با استفاده از روش‌ها و مدل‌های اقتصادسنجی این شناخت ایجاد می‌شود. فروض مدل‌های اقتصادسنجی فروض کمکی نامیده می‌شوند. دو مشکل اساسی در اقتصادسنجی که آزمون پذیری را با چالش مواجه می‌کنند؛ نادیده گرفتن فروض کمکی و ساده انگاشتن معنی‌داری آماری می‌باشند. دوئم (۱۹۰۴) بیان داشت که در صورت ناسازگاری فروض کمکی با واقعیت روش ابطال‌پذیری پوپر زیر سوال خواهد رفت. در پژوهش حاضر مسئله دوئم در زمینه فرضیه بازارهای کارا (EMH) و ساختار داده‌های مالی مورد بررسی قرار گرفته است. در بخشی دیگری استفاده و تفسیر معنی‌داری آماری مطالعه شده است که معمولاً با شاخصی به نام p-مقدار ارزیابی می‌شود. در حالی‌که p-مقدار می‌تواند یک متریک آماری مناسب باشد اما استفاده و تفسیر آن به شکلی نامطلوب صورت می‌گیرد. در نتیجه، برخی مجلات معتبر علمی،استفاده از p-مقدار را ممنوع کرده‌اند. برای استفاده مناسب از این معیار در روش‌شناسی اقتصادسنجی پیشنهادات انجمن آمار آمریکا مورد بحث واقع شده است. همچنین برای مقابله با نقش فروض کمکی آزمون‌های تصریح و توجه به ساختار داده‌ها ضروری است.

طبقه‌بندی JEL: A00، B00، B16
کلیدواژه‌ها: اقتصادسنجی، ابطال‌پذیری، مسئله دوئم، معنی‌داری آماری، p-مقدار

۱. مقدمه

علم نوعی از شناخت انسانی است که در جستجوی کشف روابط میان اشیاء[1] و تبیین وجود چنین روابطی با استفاده نظریه‌های علمی است. تفاوت شناخت علمی با سایر شیوه‌های شناخت در آن است که این نوع شناخت توسط نظریه‌های علمی محقق می‌شود. همچنین برخلاف سایر شیوه‌ها، ادعاهای نظریه‌های علمی به شکلی بیان می‌شوند که قابل آزمون باشند تا در نهایت با استفاده از معیار یا معیارهایی رد یا پذیرفته شوند. این شیوه کمک می‌کند تا شناخت علمی از خطا پیراسته شود[2]. برای اینکه ادعایی قابل آزمون باشد باید قابل اندازه‌گیری باشد. بنابراین، شناخت علمی و اندازه‌گیری از یکدیگر قابل تفکیک نیستند؛ به گونه‌ای که می‌توان گفت شناخت بدون اندازه‌گیری تنها درکی فلسفی از پدیده مورد بررسی است و اندازه‌گیری بدون نظریه تنها تولید یک عدد تصادفی بی معناست. شایان توجه است که منظور از اندازه‌گیری صرفا اندازه‌گیری کمی نیست بلکه هرگونه تخصیص اعداد به یک یا چند ویژگی اشیاء، رویدادها و وضعیت‌ها براساس یک قاعده با هدف تولید اطلاعات موثق در مورد این اشیاء، رویدادها یا وضعیت‌ها را اندازه‌گیری می‌نامیم که نیازمند استفاده گسترده از ریاضیات و آمار است (استیونس[3]، ۱۹۵۹).

اقتصاد، علمی اجتماعی است که به طور خلاصه پدیده‌های مربوط به تولید و توزیع کالاها و خدمات را در جامعه بررسی می‌کند. ابزار اساسی اقتصاد برای شناخت علمی، اقتصادسنجی[4] است. واژه‌ی اقتصادسنجی ترکیب یونانی کلمات Oikonomia = اقتصاد و Metron = اندازه‌گیری است. فریش[5](۱۹۳۳) در توضیح چیستی اقتصادسنجی چنین می‌گوید: "چندین جنبه از رویکرد کمی به اقتصاد وجود دارد که هیچ یک از این جنبه‌ها به طور منفرد و به خودی خود نباید با اقتصادسنجی اشتباه گرفته شود. لذا، اقتصادسنجی به هیچ وجه با آمارهای اقتصادی متناظر نیست. گرچه بخش قابل توجهی از این نظریه دارای بعدی کاملاً کمی است اما با آنچه که به طور کلی نظریه اقتصادی می‌نامیم یکسان نیست. همچنین، نباید اقتصادسنجی را مترادف استفاده از ریاضیات در اقتصاد دانست. تجربه نشان داده است که هر یک از این سه ضلع، یعنی آمار، نظریه‌ی اقتصادی و ریاضیات، شروط لازم اما به تنهایی ناکافی برای درک واقعی روابط کمی در زندگی اقتصادی مدرن

---

1. Objects
2. Erorr Free
3. Stevens
4. Econometrics
5. Frisch

است. اتحاد هر سه ضلع فوق قدرتمند است و این اتحاد است که اقتصادسنجی را تشکیل می‌دهد".

تعریف فوق، فرآیند تولید دانش در اقتصاد مدرن را توصیف می‌کند. به منظور آنکه تعریف کم نظیر فریش از اقتصادسنجی کامل‌تر توصیف شود، می‌توان آن را در یک الگوریتم به صورت زیر بیان کرد:

(۱) کشف یک نظم اقتصادی

(۲) ارائه یک نظریه برای تبیین یا پیش‌بینی (یا هر دو) چنین نظمی

(۳) ارائه مدلی ریاضی برای توصیف و تشبیه تعاریفی که نظریه از آن نظم ارائه می‌کند

(۴) کاربرد مجموعه‌ای از آزمون‌های آماری برای ارزیابی مدل نظری-آماری با داده‌ها و واقعیت

به عقیده‌ی هاسمن[1] (۱۹۹۲) با استفاده از چنین شیوه‌ای است که دانش در اقتصاد تولید می‌شود. در صورت تایید نظریه‌ی ارائه شده می‌توان از مدل مستخرج از آن برای سیاست‌گذاری استفاده نمود. الگوریتم ارائه شده‌ی فوق را در ادبیات اقتصادی، اقتصادسنجی می‌نامند.

سه نکته‌ی مهم قابل تشخیص در این الگوریتم عبارتند از:

(الف) منظور از تبیین یک نظم، توضیح علی چنین نظمی است.

(ب) به طور مستقیم نمی‌توان نظریه‌های اقتصادی را آزمون کرد و شرایط این کار را باید با واسطه مدل‌های ریاضی فراهم آورد و

(ج) نظریه در پی تعریف واقعیت مورد نظرش است و مدل در پی تشبیه و توصیف آن واقعیت به گونه‌ای است که نظریه‌ی زیربنایی‌اش قابل آزمون باشد.

مورد (الف) نتیجه روش‌شناسی مرسوم در علم اقتصاد است که تنها با استفاده از نظریه‌ها می‌توان روابط میان متغیرهای اقتصادی (که در عرف اقتصادسنجی به درون‌زا و برون‌زا تقسیم‌بندی می‌شوند) را ایجاد کرد و این روابط با استفاده از ساختارهای آماری صرفا بررسی و توصیف می‌شوند نه اینکه ساخته شوند. این موضوع پیامد استیلای روش شناسی پوپر[2] (۱۹۶۳) در علم اقتصاد است. روش‌شناسی که پوپر برای حل معضل استقراء[3] به کار

---

1. Hausman
2. Popper
3. Problem of Induction

برد تغییر فرم ساختار تبیین‌های علمی بود. به منظور توضیح بیشتر این موضوع، صورت کلی یک استدلال استقرایی که به صورت زیر قابل نمایش است را مفروض می‌گیریم:

$$\frac{\begin{array}{c}(i)h \Rightarrow e \\ (ii)e\end{array}}{(iii)(probably)h} \qquad (۱)$$

معنای اجزای این رابطه آن است (i) اگر فرضیه یا نظریه h درست باشد آنگاه شواهد e را مشاهده خواهیم کرد (که با $h \Rightarrow e$ نشان داده شده است)، (ii) شواهد e مشاهده می‌شود پس (iii) احتمالا h درست است. اینکه از بررسی جزییات e (مشاهدات محدود) چگونه می‌توان به یک قانون کلی رسید که فرضیه h آن‌را بیان می‌کند، در این صورت استدلالی یا تفسیر آن مشخص نیست. بنابراین، استدلال استقرایی هیچ استلزامی برای برقراری نتیجه ندارد. به همین دلیل، در نتیجه‌ی استدلال از کلمه *Probably* (احتمالا) استفاده شده است (اسپانوس و همکاران[1]، ۲۰۱۲). پوپر برای حل مسئله استقراء فرم استدلال را به طور کلی تغییر و به صورت زیر در آورد:

$$\frac{\begin{array}{c}(i)h \Rightarrow e \\ (ii)\acute{e}\end{array}}{(iii)\acute{h}} \qquad (۲)$$

اجزا این صورت استدلالی به صورت زیر است:

(i) اگر فرضیه یا نظریه h درست باشد آنگاه شواهد e را مشاهده خواهیم کرد (که با $h \Rightarrow e$ نشان داده شده است)، (ii) شواهد نقض کننده h که با e' نشان داده شده است مشاهده می‌شود پس (iii) قطعا h نادرست است که با h' نشان داده شده است. این رویکرد جایگزین را پوپر (۱۹۶۳) ابطال‌گرایی[2] نامیده است. در رویکرد پوپر ابتدا نظریه براساس قواعد قیاسی-فرضیه‌ای ساخته می‌شود و سپس با استفاده از مشاهدات و داده‌ها در پی بررسی تجربی پیشگویی‌های نظریه خواهیم بود. بنابراین، **در رویکرد پوپر روابط از داده‌ها و مشاهدات حاصل نمی‌شوند بلکه تنها با استفاده از آنها بررسی می‌شوند**.

[1]. Spanos
[2]. Falsifiability

این توصیف را می‌توان در مباحثه‌ی میان کوپمنز[۱] و واینینگ[۲] (۱۹۴۹) در شماره‌های اول مجله‌ی اکونومتریکا[۳] یافت.

در تعاریف جدید، نظریه به صورت خانواده‌ای از مدل‌ها تعریف می‌شود که با استفاده از اصول موضوعه‌ایی[۴] ساخته می‌شود (بالزر و دیگران[۵]، ۱۹۸۷). به همین دلیل نمی‌توان به طور مستقیم ادعاهای یک نظریه را درباره کارکردهای جهان مورد نظرش را بررسی کرد. این شیوه به اقتصاددانان کمک می‌کند تا با استفاده از تصریح مدل‌ها (بعنوان زوج نظریه‌ها در شیوه اصول موضوعه‌ایی) سوالات مطرح شده در اقتصاد را پاسخ دهند (شیارمور جی[۶]، ۱۹۹۹)[۷]. همچنین، نظریه‌های اقتصادی نحوه تبیین[۸] نتایج بدست آمده از مدل‌سازی‌های اقتصادی را نیز بیان می‌کنند. این موضوع در بخش بعدی توضیح داده شده است. شایان ذکر است که در ساخت مدل نظری روابط علی و معلولی میان متغیرهای اقتصادی (که منشا چنین روابطی نظریه است) برمبنای دو فرض عدم خطای اندازه‌گیری در کمیت‌های نظری و برقراری فرض ثبات سایر شرایط[۹] ایجاد می‌شود. اما از آنجا که برخلاف داده‌های آزمایشگاهی، داده‌های اقتصادی در آزمایشگاه بدست نمی‌آیند بلکه از درون خود فرایند اقتصادی فراهم می‌شوند، تحت تاثیر تعامل بین عوامل اقتصادی قرار دارند (آلفا سی چیانگ[۱۰]، فصل هشتم ،۲۰۰۵). این موضوع سبب می‌شود که بین متغیرهای نظری و متغیرهای تجربی اقتصاد تفاوت وجود داشته باشد. در این حالت از آنجا که نه سایر شرایط

---

1. Tinbergen
2. Vining
3. Econometrica
4. Formal-Axiomatic
5. Balzer
6. Shearmur

۷. این شیوه بررسی مسائل علمی استقراگرایی (Inductivism) نامیده می‌شود که تبیین و پیش‌بینی یک وضعیت خاص (State) با استفاده از مواجهه نظریه-داده یا نظریه-مشاهده (Hypothesis –led- Observation) صورت می‌گیرد.

۸. Explanation

۹. فرض ثبات سایر شرایط که با عبارت Ceteris Paribus نشان داده می‌شود به معنای عدم حضور شرایطی است که قانون را از عمومیت می‌اندازند. برخی اوقات این عدم حضور به معنای ثبات سایر شرایط بوده و در مواقعی هم به معنای عدم حضور سایر شرایط است. بعنوان مثال، قوانین فیزیک نیوتن در شرایطی برقرارند که سایر شرایطی همچون مقاومت هوا یا اصطکاک وجود نداشته باشند. در اقتصاد هم باید به این موضوع توجه داشت. در اقتصاد فرض ثبات سایر شرایط برخی اوقات به فرض برقراری سایر شرایط تغییر می‌یابد. برای مثال قانون سه (Say's law) که بیان می‌کند عرضه تقاضای خود را به وجود می‌آورد با فرض برقراری تولید مطابق ترجیحات مصرف‌کنندگان است.

10. Alpha C. Chiang

ثابت است و نه ابزارهای اندازه‌گیری آنچنان دقیق هستند که بدون خطا داده‌ها را تولید کنند، مواجهه تمامی مدل‌های یک نظریه با داده‌ها به عدم سازگاری آنها ختم خواهد شد. راه حل اقتصادسنجی برای این مسئله عبارت است از افزودن یک عبارت خطا به روابط نظریه به منظور لحاظ کردن خطاهای اندازه‌گیری. در این صورت می‌توان پیش‌بینی‌های بدست آمده از مدل‌های نظری در مورد روابط علی میان متغیرهای اقتصادی را با داده‌ها و شواهد موجود بررسی نمود.

با این حال، این انتهای راه نیست بلکه صرفا شیوه‌ی سیر در راه دستیابی به دانش در اقتصاد است. در این سیر مسائلی روی می‌دهند که می‌توانند تصویر تحریف شده‌ای از واقعیت‌های اقتصادی جوامع ارائه دهند. هدف این پژوهش، بررسی نقش فروض مدل‌های آماری و شیوه نگارش و تفسیر معنی‌داری مرسوم در تحقیقات اقتصادی مربوط به آزمون نظریه‌های اقتصادی است. پژوهش حاضر نشان می‌دهد که آزمون نظریه‌های اقتصادی با کمک مدل‌های ریاضی و آماری چگونه تحت تاثیر فروض مدل‌سازی قرار می‌گیرد. نادیده انگاشتن نقش فروض مدل در بررسی نظریه‌های اقتصادی، اقتصادسنجی را در بررسی نظریه‌ها یا فرضیه‌های علم اقتصاد دچار چالش ساختاری می‌کند! استنتاج در این شرایط منجر به نتایج کاذبی می‌گردد که ظاهرا صحیح هستند. اهمیت این موضوع در آن است که نادیده انگاشتن آنها تقریبا به روالی عادی در پژوهش‌های تجربی مبتنی بر اقتصادسنجی تبدیل شده است.

ساختار این پژوهش در ادامه به صورت زیر است:

در بخش دو مسئله دوئم و نقش فروض کمکی در بررسی تجربی نظریه‌های اقتصادی مورد بررسی قرار می‌گیرد، در بخش سوم ارتباط مسئله دوئم و ساختار داده‌های اقتصادی بررسی می‌شود. در بخش چهارم معنی‌داری آماری و نادیده گرفتن دستور العمل‌های ASA بررسی می شود و در بخش پنجم نتیجه‌گیری ارائه می‌گردد.

۲. مسئله دوئم و تبعات فروض کمکی در بررسی نظریه‌ها اقتصادی با استفاده از اقتصادسنجی

۲-۱. تشریح مسئله‌ی دوئم در آزمون نظریه‌ها یا فرضیه‌ها

یکی از محورهای اصلی علم اقتصاد، سیاست‌گذاری اقتصادی است. علم اقتصاد همچنین نقش مهمی در زمینه‌های دیگر همچون سیاست‌گذاری اجتماعی، مبارزه با تروریسم (مکیان و رستمی[1]، ۱۳۹۷)، تغییرات آب و هوایی و... ایفا می‌کند. سیاست‌گذاری موفق محصول شناخت علمی روابط علّی است. شناخت روابط علّی با استفاده از نظریه‌های علمی که آزمون‌های تجربی را به شکل موفقیت آمیزی پشت سر گذارده‌اند ممکن است. می‌توان گفت که تقریبا بررسی تجربی نظریه‌های اقتصادی بدون مدل‌سازی امکان پذیر نیست (جز در برخی حوزه‌های اقتصاد خرد که می‌توان از روش‌های آزمایشگاهی استفاده کرد). مدل، یک الگوی منطقی است که تفسیری قابل بررسی از نظریه ارائه می‌دهد. در منطق[2]، تعریف مدل به صورت زیر است:

مدل تعبیری از یک نظریه (بیان شده در منطق مرتبه اول) است که در آن قضایای نظریه صادق هستند (آ.گ.همیلتون[3]، منطق برای ریاضیدانان، فصل چهار).

براساس این تعریف، مدل سیستمی منطقی است که ساختار آن با ساختار نظریه در تناقض منطقی نیست. برای آنکه قضایای یک نظریه قابل آزمون باشد علاوه بر فرض‌های خود نظریه به فرض‌های دیگری نیز نیاز است که این فرض‌ها را فرض‌های کمکی[4] می‌نامند. بسته به انتخاب این نوع فرض‌ها مدل‌های مختلفی ایجاد می‌شوند. دوئم[5] (۱۹۰۴) به نقش تناقض آمیز فرض‌های کمکی در آزمون تجربی نظریه‌های علمی اشاره کرد. وی بر این اعتقاد بود که فروض کمکی در نهایت نظریه را غیرقابل آزمون می‌کنند. آنگونه که دوئم در مقاله‌اش بیان می‌کند در بررسی تجربی یک نظریه علمی، دانشمندان نمی‌توانند یک فرض منفرد [در قالب فرضیه‌های آماری] را آزمون تجربی کنند؛ بلکه کل گروه فرض‌ها را آزمون می‌کنند. بنابراین، اگر نتیجه‌ی آزمون با پش‌بینی نظریه مخالف باشد، آنچه دانشمندان می‌توانند استنتاج کنند این است که حداقل یکی از فرض‌های موجود در مجموعه فرض‌های وی غیرقابل قبول است و باید تغییر کند. علاوه بر این، آزمون نمی‌تواند مشخص کند در این مجموعه کدام فرض با واقعیت ناسازگار است (دوئم، ۱۹۰۴).

بیان دقیق ابهام دوئم بدین قرار است که؛ اگر مجموعه فرض‌های نظریه را با h و مجموعه فرض‌های مدل یا فرض‌های کمکی را با a نمایش دهیم آنگاه مجموعه فرضیات نظریه-مدل برابر با $h+a$ خواهد بود. بررسی تجربی نظریه-مدل با فرضیات $h+a$ زمانی که

[1]. Makyian & Rostami
[2]. Logic
[3]. Hamilton
[4]. Auxiliary Assumptions
[5]. Duhem

فرض‌های a در مدل رعایت شده‌اند، می‌تواند به معنای بررسی نظریه یا مجموعه فرض‌های h باشد. در صورتی که از برقراری فرض‌های a در مدل مطمئن نباشیم! در این صورت بررسی تجربی نظریه ناممکن خواهد بود زیرا در صورت عدم سازگاری نظریه با داده‌ها نمی‌توانیم به صورت مطمئن نتیجه بگیریم که این عدم سازگاری مربوط به فرضیه‌های نظریه (مجموعه‌ی h) است و یا مربوط به فرضیه‌های مدل (مجموعه‌ی a) است.

۲-۲. ابهام دوئم و تفسیر نتایج مدل‌های تجربی اقتصادسنجی

در زمینه تحقیقات اقتصادی ابهام دوئم چالش برانگیز است. زیرا هر نظریه‌ی اقتصادی برمبنای یک‌سری فرض‌ها بیان می‌شود که با استفاده از یک یا چند مدل آماری این فرض‌ها بررسی می‌شوند. رد فرض‌ها می‌تواند به دلیل تصریح نادرست مدل‌های آماری باشد (فرض‌های مدل یا فروض کمکی برقرار نیستند و یا در صورت برقراری فرض‌های سازگاری با واقعیت نیستند) که در این صورت مدل آماری ابزاری مناسب برای بررسی فرضیه‌ها نخواهد بود.

با قاطعیت می‌توان گفت در تحقیقات داخلی اساسا این ابهام مورد توجه قرار نمی‌گیرد. یکی از زمینه‌های اساسی که این موضوع مورد توجه قرار نمی‌گیرد بررسی وجود روابط بلندمدت میان سری‌های زمانی مالی است. در اینجا موردی از وجود رابطه‌ی بلندمدت میان دوبازار ارز و سهام بررسی می‌شود.

مدل آماری هم‌جمعی یا هم‌انباشتگی[1] مدلی است که به منظور مطالعه‌ی تجربی روابط بلندمدت اقتصادی در بین متغیرهای نامانا معرفی شده است که کاربردهای فراوانی در اقتصاد و مالی دارد. دلیل اهمیت این کاربرد در آن است که با وجود آنکه سری‌های زمانی نامانا به‌طور جداگانه پیش‌بینی‌ناپذیرند (به دلیل آنکه میانگین و واریانس این متغیرها در زمان متغیر است) اما ترکیب خطی آنها می‌تواند پیش‌بینی‌پذیر باشد. به طور کلی قیمت دارایی‌های مالی که مانا باشند دارای خاصیت برگشت‌پذیری به میانگین هستند (در صورت وجود). بنابراین، با استفاده از این خاصیت می‌توان استراتژی‌های مبادله‌ی سودآور را طراحی کرد. بدین صورت که زمانی‌که قیمت دارایی از میانگین آن بالاتر بود دارایی را فروخت و زمانی‌که قیمت دارایی کمتر از میانگین قرار داشت دارایی را خرید. اما قیمت دارایی‌ها معمولا نامانا‌ست و طراحی استراتژی فوق ناممکن است. با این حال چنانچه یک ترکیب خطی از دو یا بیشتر از دو قیمت دارایی نامانا وجود داشته باشد که مانا باشد می‌توان

[1]. Co-Integration

استراتژی مبادله را مشابه حالت فوق طراحی نمود. فرضیه بازارهای کارا[۱] در بازاری همچون بازار سهام یا بازار ارز، وجود چنین استراتژی را منکر می‌شود. مطابق مبانی این فرضیه وجود ارتباط کوتاه‌مدت میان بازار ارز و بازار سهام از لحاظ نظریه‌های اقتصاد مالی مشکلی را ایجاد نمی‌کند. اما ارتباط بلندمدت یا هم‌انباشتگی میان آنها با فرضیه بازارهای کارا[۲] در تقابل است. بنابراین، به دلیل آنکه می‌توان از چنین ارتباطی در جهت روندهای آتی بازار سهام استفاده نمود فرضیه بازار کارا نقض خواهد شد (موکرجی و یو[۳]، ۱۹۹۷). نتایج تحقیقات تجربی درباره رابطه میان این دو بازار (و جهت علیت) متناقض است. چنین تناقض‌هایی دو دلیل اصلی دارند که هر دو ریشه در عدم تفکیک فرض‌های نظریه‌ی بازارهای کارا و مدل‌های آماری مورد نظر برای بررسی تجربی این فرضیه می‌باشند. دلیل اول ریشه‌ی تناقض‌ها را در محدودیت‌های ایجاد شده توسط فروض کمکی در شیوه‌ی مدل‌سازی کلاسیکی یا فراوانی‌گرا می‌بینند. این محدودیت سبب عدم کفایت آزمون فرضیه‌های کلاسیک در این زمینه می‌شود. محدودیت‌های مطرح شده در این زمینه عبارتند از:

۱. در آزمون‌های کلاسیک ریشه واحد (مانند ADF و PP) و هم‌انباشتگی (مانند انگل-گرنجر و فیلیپس اولایر) نظریه توزیع مجانبی به شکل ناپیوسته بین فرضیه وجود ریشه‌ی واحد و فرضیه مانایی تغییر می‌کند و لذا آزمون فرضیه‌های کلاسیکی (مبتنی بر نظریه مجانبی) نمی‌تواند روشی منطقی برای استنباط آماری براساس نظریه مجانبی ناپیوسته را بدست دهد (سیمز[۴] (۱۹۸۸)).

۲. در آزمون‌های ADF، PP و KPSS و همچنین به طور متناظر در آزمون‌های هم‌انباشتگی کلاسیک مقادیر بحرانی در عمل برای نمونه‌های کوچک اساسا متفاوت از مقادیر مجانبی آنهاست (کوپ[۵] (۱۹۹۴)). این موضوع می‌تواند منجر به تولید نتایج متناقضی در نمونه‌های کوچک شود.

---

[۱]. فرضیه بازارهای کارا در بازار سهام بیان می‌کند که سرمایه گذاران به صورت واحدهای عقلایی رفتار می‌کنند و از تمام اطلاعات موجود در جهت کشف روند آتی قیمت‌های سهام استفاده می‌کنند. به این علت حرکات قیمت سهام به صورت تصادفی خواهد بود.

[2]. Market Efficient Hypothesis
[3]. Mookerjee and Yu
[4]. Sims
[5]. Koop

۳. استنباط وجود یا عدم وجود رابطه بلندمدت باید مشروط به نمونه‌های دسترس انجام پذیرد که در این صورت وجود نتایج متفاوت متناقض نخواهد بود. زیرا تغییر نتایج با تغییر نمونه‌ها به رسمیت شناخته شده است.

چنانچه ضعف روش‌های کلاسیک آزمون‌های هم انباشتگی موجود تنها موارد فوق باشد، با تغییر شیوه مدل‌سازی از شیوه‌ی کلاسیک به بیزی می‌توان ضعف‌های یاد شده را برطرف کرد (برای مثال به سیمز (۱۹۸۸) مراجعه کنید).

موارد فوق، فرض‌های کمکی رویکرد مدل‌سازی کلاسیک را در نتایج متناقض بدست آمده از آزمون‌های هم‌انباشتگی باقیمانده محور را دخیل می‌داند. با این وجود، حتی با استفاده از مدل‌هایی که مبتنی بر این فرض‌ها نباشند نیز مسئله کماکان باقی است.

به منظور فهم این موضوع، در ادامه فرضیه‌ی بازارهای کارا با جزییات بیشتری بررسی می‌شود. براساس فرضیه‌ی بازارهای کارا، قیمت‌های فعلی بازار حاوی کلیه اطلاعات موجود در بازار است. شرکت کنندگان در بازار کاملا عقلایی هستند. زیرا براساس همان اطلاعاتی که در دسترس همگان است و پردازش شده‌اند، عمل می‌کنند. از نظر فنی، این بدان معنی است که بهترین پیش‌بینی قیمت آینده‌ی بازار قیمت فعلی است. این موضوع معادل آن است که گفته شود تغییرات قیمت در طول زمان مستقل از یکدیگر بوده و بهترین پیش‌بینی این تغییرات برابر صفر است. براساس فرضیه این نظریه، بازار اساساً در تعادل است و نمی‌تواند حاوی حباب‌های سوداگرانه باشد (بادی[1] و همکاران، ۲۰۱۸). براساس اینکه در این تعریف منظور از اطلاعات چیست سه فرم از فرضیه کارایی خواهیم داشت که در جدول زیر نشان داده شده است:

جدول۱: فرم‌های مختلف فرضیه ی بازارهای کارا

فرم کارایی	مجموعه اطلاعات درباره قیمت‌های سهام
ضغیف	قیمت های گذشته سهام
نیمه قوی	تمام اطلاعات موجود در دسترس عموم
قوی	تمام اطلاعات اعم از عمومی و خصوصی

اگر تمامی اطلاعات در دسترس در قیمت‌های سهام خود رانشان داده باشند، در این صورت تغییرات در قیمت سهام باید کاملا تصادفی باشد. بنابراین، قیمت‌های سهام ضرورتا باید از یک فرآیند گام تصادفی[2] تبعیت نمایند تا تغییرات در قیمت‌های سهام کاملا

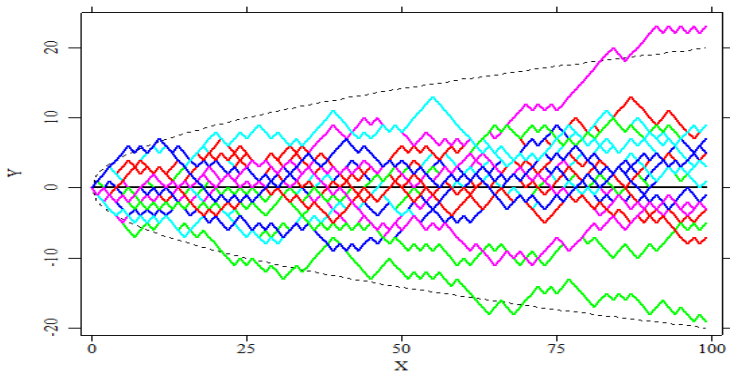
1. Bodie
2. Random Walk

تصادفی با شد یا در ا صطلاح اقتصاد سنجی معادل یک فرآیند نوفه‌ی سفید[1] با شد. مدل آماری حرکت قیمت‌های سهام در طول زمان، که با P_t نشان داده می شود به‌صورت زیر خواهد بود:

$$P_t = \mu + P_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim NIID(0,\sigma^2) \qquad (۳)$$

در این رابطه، μ تغییر قیمت مورد انتظار یا جهش[2] و عبارت $NIID(0,\sigma^2)$ نشـــان دهنده مستقل و هم‌توزیع بودن جملات خطای ε_t با میانگین صفر و واریانس σ^2 ا ست. فرض نرمال بودن ε_t بسـیاری از پیچیدگی‌های محاسـباتی با یک فرآیند گام تصـادفی را ساده می‌کند. با این حال، برخی مشکلات از جمله احتمال منفی شدن قیمت‌های سهام وجود خواهد داشـت که در نمودار ۱ با اسـتفاده از شـبیه‌سـازی ۲۰ مسـیر تصـادفی یک فرآیند گام تصادفی این موضوع نشان داده شده است.

نمودار ۱: شبیه سازی ۲۰ مسیر تصادفی با استفاده از یک فرآیند گام تصادفی



منبع:یافته‌های پژوهش

به منظور غلبه بر این مشکل از لگاریتم قیمت‌های سهام در رابطه‌ی (۱) استفاده می‌شود. در نتیجه، رابطه (۳) را می‌توان به‌صورت زیر بازنویسی کرد:

$$Ln(P_t) = \mu + Ln(P_{t-1}) + \varepsilon_t,\quad \varepsilon_t \sim NIID(0,\sigma^2) \qquad (۴)$$

[1]. White Noise
[2]. Drift

عبارت $\varepsilon_t \sim NIID(0,\sigma^2)$ در رابطه‌ی فوق به معنای استقلال تغییرات قیمت‌ها در طول زمان و تبعیت این تغییرات از توزیع نرمال است. روش فوق مدل‌سازی فرضیه‌ی بازارهای کارا با استفاده از یک فرآیند براونی هندسی پیشنهاد می‌کند.

در این روابط فرض کمکی تبعیت تغییرات قیمت دارایی از توزیع نرمال در مدلسازی فرضیه بازارهای کارا پیامدی اساسی در پی خواهد داشت. براساس قضیه‌ی حد مرکزی توزیع نرمال مناسب شرایطی است که تعداد زیادی از عوامل تصادفی مستقل با واریانس متناهی و مثبت و با اثر ناچیز بر متغیر مورد نظر تاثیرگذارند (وولدریج، ۲۰۱۷). بنابراین، در چنین وضعیتی احتمال ضرر و زیان‌های نسبتا بزرگ بسیار ناچیز خواهد بود. همچنین فرض کمکی دیگر در این شیوه مدلسازی استقلال تغییرات قیمت دارایی در طول زمان بود.

شواهد عدم برقراری این فرض‌ها در واقعیت را می‌توان در وجود حباب سوادگرانه قیمت مسکن در آستانه بحران مالی ۲۰۰۷ یا حباب دات-کام[1] در سال ۲۰۰۰ مشاهده کرد. این شواهد نشان می‌دهند که قیمت‌های واقعی چیزی بیشتر از اطلاعات در دسترس همه می‌باشند (به همین دلیل حباب دارند) و تغییرات قیمت در طول زمان مستقل نیستند. در ادبیات اقتصاد مالی، فرض تبعیت توزیع بازده دارایی از توزیع نرمال به دلیل وقوع پیشامدهای شدید اقتصادی (بحران های مالی) به شدت مورد انتقاد قرار گرفته است. اگرچه فرض توزیع نرمال در علوم طبیعی بسیار رایج و مفید است لیکن برای روش‌های علوم اجتماعی، به ویژه بخش مالی مناسب نیست. با این حال، واقعیت با فرضیه بازار کارا مغایرت دارد و شواهد نشان می‌دهند که تغییرات قیمت به طور مستقل توزیع نمی‌شود و مقادیر حدی بازده امری با احتمال اندک نیست و با یک واریانس انفجاری (نامتناهی) همراه است. توزیع نرمال به پیشامدهای حدی احتمالی نزدیک به صفر تخصیص می‌دهد. با این حال، این وقایع بسیار بیشتر از آنچه این فرض پیش بینی می‌کند، در واقعیت رخ می‌دهند.

نتیجه این موضوعات این است که توزیع نرمال که استفاده از آن در عمل بسیار رایج است، یک فرض کمکی نامناسب است.

ریشه نامناسب بودن این فرض در یکی از مهمترین واقعیت‌های آشکار شده در حوزه تلاطم[2] دارایی‌های مالی است که بیان می‌کند تلاطم این دارایی‌ها در طول زمان متغیر است (مندلبرات[3] (۱۹۶۳) ، انگل[4] (۲۰۰۴) و شورت[5] (۱۹۸۹)). به این ویژگی اصطلاحا خاصیت

---

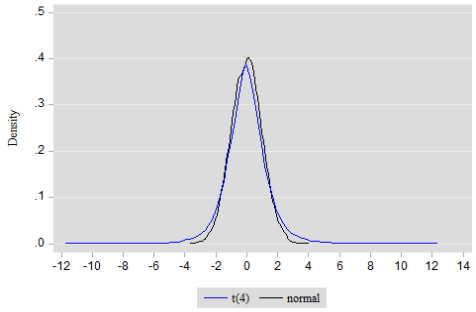
1. Dot-Com
2. Volatility Clustering
3. Mandelbrot
4. Engel
5. Schwert

خوشه‌ای تلاطم گفته می‌شود که مبین وجود خودهمبستگی مثبت دوره‌های تلاطمی است. این امر موجب می‌شود که بازده به طور شرطی ناهمسانی واریانس داشته باشد. لذا، واریانس شرطی بازده با واریانس غیر شرطی آن تفاوت داشته باشد یا به لحاظ آماری یعنی:

$$(\sigma_t^2|t-1) = Var(r_t|\Omega_{t-1}) \neq Var(r_t) \qquad (۵)$$

پیامد نابرابری فوق ضریب کشیدگی بیش از توزیع نرمال داده های بازده است.
نمودار شماره ۲ توزیع شبیه سازی شده با کشیدگی بالا (توزیع t - استیودنت) را در مقایسه با توزیع نرمال استاندارد نشان می‌دهد. نمودار ۲ نشان می‌دهد که توزیع با کشیدگی بالا سه بار از توزیع نرمال می‌گذرد. مفهوم ضمنی این پدیده آن است که در محدوده زمانی نسبتا بزرگی، در توزیع با کشیدگی بالا داده‌ها تغییرپذیری کمتری یا تقریبا مشابه‌ای با یک موقعیت نرمال دارند (حوزه‌ی میانی توزیع با کشیدگی بالا چنین وضعیتی را توصیف می‌کند.). اما موقعیت‌هایی نیز وجود دارند که تغییرپذیری داده‌ها در توزیع با کشیدگی بالا نسبت به توزیع نرمال بسیار بزرگتر است (که دنباله های توزیع با کشیدگی بالا به این موضوع اشاره می کنند).

نمودار ۲: مقایسه توزیع با کشیدگی بالا (توزیع t با ۴ درجه آزادی) با توزیع نرمال

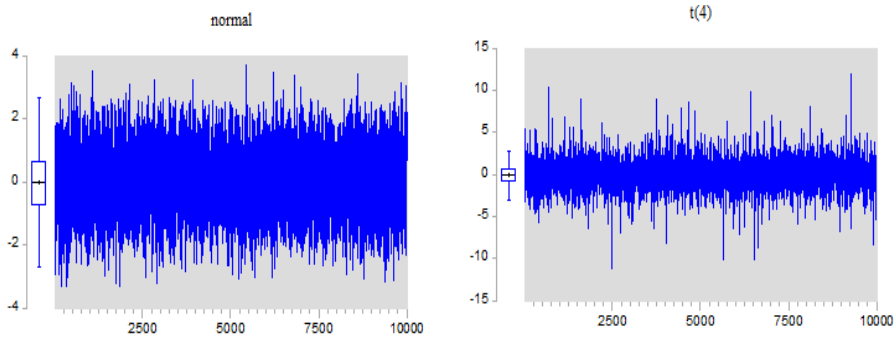


منبع: مقاله مکیان و رستمی، ۱۳۹۸

نمودار شماره ۳ سری زمانی داده‌های شبیه‌سازی شده با استفاده از توزیع t استیودنت را در مقایسه با سری زمانی داده‌های شبیه‌سازی شده با توزیع نرمال را نشان می‌دهد. این نمودار نشان می‌دهد که تغییرات داده‌های با کشیدگی بالا در مقایسه با داده‌های نرمال بسیار دندانه‌دارتر و و ناصاف‌تر (تغییرپذیرتر) هستند. این نوع تغییرپذیری در نتیجه وجود

پدیده‌ای است که در اقتصادمالی تلاطم خوشه‌ایی نامیده می شود .(رستمی و مکیان، ۱۳۹۸)

نمودار ۳: مقایسه سری زمانی داده های شبیه‌سازی شده با توزیع نرمال و توزیع با کشیدگی بالا



منبع: مقاله مکیان و رستمی، ۱۳۹۸

براساس این نتایج، مهم‌ترین ضعف آزمون‌های فرضیه‌ی بازارهای کارا در نادیده گرفتن واقعیت فرآیند مولد داده‌های قیمت دارایی‌های مالی توسط فروض کمکی مدل است. به طور کلی، فرآیند مولد داده‌های قیمت دارایی‌های مالی حاوی اثرات ARCH است که سبب می‌شود خواص مجانبی آزمون‌های ریشه واحد و هم‌انباشتگی کلاسیک و بیزی در حضور این اثرات شناخته شده نباشد (نلسون[1]، ۱۹۹۱). لذا، تنها با تغییر شیوه‌ی مدل‌سازی (از کلاسیک به بیزی) مسئله‌ی بررسی فرضیه‌ی بازارهای کارا (با توجه به ارتباط بین بازارها) حل نمی‌شود بلکه تصریح مناسب اجزاء مدل براساس واقعیت‌های تلطیف شده[2] بسیار با اهمیت است. این موضوع بدین معناست که در این زمینه ابهام دوئم بشدت اعتبار تجربی آزمون فرضیه‌ی بازارهای کارا را زیر سوال می‌برد. زیرا رد این فرضیه (یا تایید آن) می‌تواند به دلیل تصریح نادرست مدل آماری به خاطر فرضیه‌های کمکی مدل اقتصادسنجی مورد استفاده باشد.

بنابراین، فرض تبعیت توزیع حاشیه‌ای بازده از توزیع نرمال و همچنین استقلال جملات خطای معادله میانگین باید به فرض ضعیف‌تر عدم خودهمبستگی تقلیل یابد. در این صورت طبقه نوفه سفید منتخب باید تفاضل مارتینگل باشد.

---

1. Nelson
2. Stylized Facts

یک مارتینگل زمان-گسسته با استفاده از امید در زمان t برای دوره t+1 مشروط به اطلاعات زمان t تعریف می شود. این تعریف معادل با نمو انتظاری برابر با صفر از دوره‌ی t تا دوره‌ی t+1 است. بنابراین، این مفهوم به طور مکرر در فرم تفاضلی بیان شده است.
فرآیند $\{\varepsilon_t\}$ تفاضل مارتینگل نامیده می شود اگر امید شرطی آن (مشروط به گذشته‌ی خودش) برابر با صفر باشد یا:

$$E\left(\varepsilon_{t+1}\middle|\sigma(\varepsilon_t,\varepsilon_{t-1},\dots)\right)=0 \qquad (۶)$$

این شرط بیان می‌کند که گذشته تاثیری بر پیش‌بینی (امید شرطی) ندارد و پیش‌بینی در این حالت همواره صفر خواهد بود. به رغم عدم خودهمبستگی سریالی، تفاضل مارتینگل‌ها به طور کلی در طول زمان مستقل محسوب نمی شوند. این موضوع، بررسی تلاطم را بصورت متغیر در طول زمان با استفاده از این فرآیندها ممکن ساخته است.
براساس این توضیحات، رابطه‌ی (۴) را باید به‌صورت زیر بازنویسی کرد:

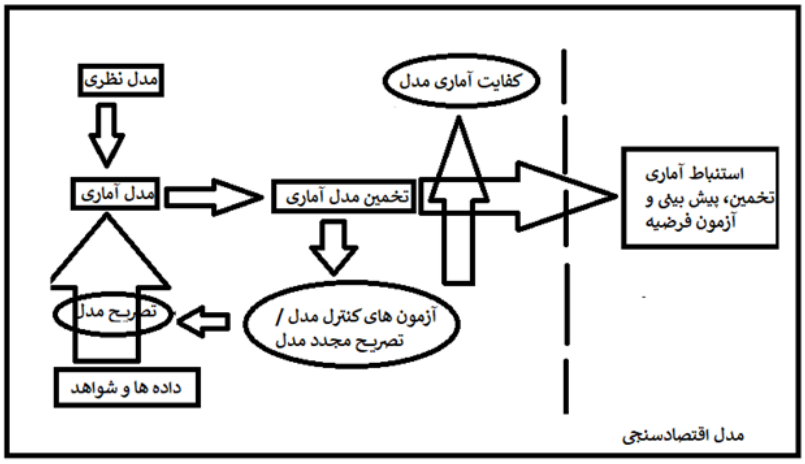
$$\begin{aligned}\mathrm{Ln}(P_t)&=\mu+\mathrm{Ln}(P_{t-1})+\varepsilon_t,\\ &\varepsilon_t\sim MD(0,\sigma^2)\end{aligned} \qquad (۷)$$

در رابطه ی فوق همانگونه که مشاهده می شود واریانس شرطی ε_t تابعی از زمان است و بنابراین فرض استقلال فرآیند نقض می‌شود. در عین حال به دلیل آنکه ε_t تفاضل مارتینگل است فاقد خودهمبستگی است. لذا می‌توان از رابطه‌ی زیر برای نشان دادن این موضوع در ε_t استفاده کرد:

$$\varepsilon_t=\sigma_t\delta\vartheta_t \qquad (۸)$$

سوال که مطرح می‌شود این است که آیا می‌توان مسئله دوئم را در آزمون فرضیه‌های اقتصادی برطرف کرد؟ پاسخ این است که بله می‌توان. برای این منظور انجام آزمون‌های تصریح مدل در مرحله قبل از استنباط آماری ضروری است. در این صورت مدل‌سازی اقتصادسنجی باید شامل مراحل زیر باشد در غیر این صورت ابهام دوئم مدل‌سازی را زیر سوال خواهد برد.

شکل ۱: تجدید نظر در شیوه‌ی عمل اقتصادسنجی با توجه به مسئله دوئم



منبع: یافته‌های پژوهش

در واقع تصریح مدل کمک می‌کند تا آزمون فرضیه نظریه ارتباطی با فروض کمکی مدل نداشته باشد. در مورد مسئله بازارهای مالی آزمون وجود اثرات ARCH در صورت تایید کمک می‌کند تا از ابتدا شیوه مدل‌سازی مورد استفاده قرار گیرد که این فرض را در بر می‌گیرد و در این صورت رد فرضیه بازارهای کارا به دلیل اتخاذ فرض‌های کمکی نادرست نخواهد بود.

۳. ارتباط مسئله دوئم و ساختار داده‌های اقتصادی

از بعدی دیگر مسئله دوئم در اقتصاد با ساختار داده‌ها علاوه بر فروض کمکی مدل‌ها نیز در ارتباط است. برای مثال پیش‌بینی ریسک‌های مالی، علاوه بر مفروضات دقیق و واقع‌گرایانه‌ی مدل‌های ریاضی و آماری، به مجموعه‌ای از داده‌ها با سطح بالایی از اطلاعات نیاز دارد. این مجموعه ازداده‌ها از نوع سری زمانی بازده، قیمت‌ها، بهره یا سود سهام هستند. داده‌های سری زمانی تنها مصداق تحقق یافته یک فرآیند تصادفی می باشند و قابل تکرار نیستند(کیم و مادلا[1]، ۲۰۰۰). بنابراین از ساختار بسیار محدوده شده‌ای دارند. علاوه بر این، داده‌های سری زمانی مالی بسیار پیچیده هستند زیرا دارای خواصی از قبیل خودهمبستگی، پایداری طولانی، ناهمسانی واریانس، خوشه‌بندی، شکست‌های ساختاری، دم سنگینی توزیع[2] و چولگی دارند. پیچیدگی مدل‌سازی با این نوع از داده‌ها به تصریحات اقتصادسنجی بسیار پیچیده و شدیدا پارامتری نیاز دارد که معمولا کاربرد آنها بسیار دشوار

[1]. Kim and Maddala
[2]. Heavy Tailed Distribution

است. یک رویکرد رایج در اقتصادسنجی مالی استفاده از مدل‌های ساده و آسان برای درک می‌باشد که مبتنی بر مفروضات ریاضی بسیار محدود (به عنوان مثال فرآیندهای گوسی) هستند. نقص این رویکرد این است که مدل‌های مورد استفاده قادر به ضبط پیچیدگی تجربی داده‌ها نیستند و از این رو از نظر اثرات علی و پیش‌بینی‌ها نتایج قابل قبولی ارائه نمی‌دهند.

یک رویکرد نسبتاً جدید برای مشکل اطلاعات محدود موجود در داده‌های مالی استفاده از داده‌های با فرکانس بالا[1] است. در دسترس بودن داده‌های مالی با فرکانس بالا از ابتدای دهه‌ی ۲۰۰۰ به دلیل پیشرفت در ظرفیت ذخیره‌سازی و قدرت محاسبه رایانه‌های جدید، امکانات مدرنی را برای اندازه‌گیری و پیش‌بینی ریسک‌های مالی و سایر اندازه‌گیری‌های مالی باز کرده است. اگرچه به نظر می‌رسد استفاده از داده‌های با فرکانس بالا وسوسه‌انگیز است، اما در عمل بسیار مشکل‌ساز است. علاوه بر خطاهای استاندارد (به عنوان مثال ورود نادرست قیمت‌ها در مجموعه داده‌ها)، اختلالات ناشی از اختلاف قیمت بازار از ارزش‌های بنیادی یک سهم که ناشی از ویژگی‌های بازار مربوطه است، باید قبل از استفاده از بین بروند.

یک مثال ساده نشان دهنده اهمیت استفاده از داده‌های حاوی اطلاعات بیشتر و عملیات دقیق ریاضی برای برآورد و پیش‌بینی ریسک‌ها در آخرین بحران مالی است. در تاریخ ۱۷ سپتامبر ۲۰۰۸ شاخص صنعتی داو جونز در نتیجه اعلام ورشکستگی برادران لمان ، ۷,۸٪ ضرر را نسبت به روز قبل از آن متحمل شد. پیش بینی VaR با سطح احتمال ۱٪ برای این روز، بر اساس مدل مقیاس-مکانی با تلاطم ثابت که با داده‌های با فرکانس پایین برآورد شده بود مقداری حدود ۳,۳٪ را پیش‌بینی می‌کرد. جایگزینی فرض تلاطم ثابت با تلاطم متغیر در زمان با استفاده از مدل‌های ARCH، پیش‌بینی را به ۴,۸٪ بهبود می‌بخشد. با توجه به داده‌های با فرکانس بالا برای اندازه‌گیری و پیش‌بینی تلاطم، پیش‌بینی VaR به ۶,۹٪ افزایش می‌یابد. فرض اضافی تبعیت داده‌های بازده از یک توزیع دم سنگین پیش بینی VaR را به بیش از ۸٪ افزایش می‌دهد. براساس این پیش‌بینی، سرمایه‌گذار باید ۸٪ از کل سرمایه‌گذاری خود را در شاخص صنعتی داو جونز ذخیره نقدی انجام دهد، که قطعاً می‌تواند زیان تحقق یافته تقریباً ۷,۸٪ را پوشش دهد (شوویکر وهمکاران[2]، ۲۰۱۷).

---

1. High Frequency Data
2. Schweiker

براساس توضیحات فوق، ابطال‌پذیری یک نظریه اقتصادی به فروض کمکی مدل‌ها و ساختار داده ها بستگی دارد. در صورت رد یک رابطه در آزمون آماری، ممکن است منبع نارسایی عدم تطابق نظریه با واقعیت یا فروض کمکی نامناسب و یا ساختار داده‌ها و یا هر سه باشد. با این حال راهی برای شناسایی آنها وجود ندارد. این را مشکل شناسایی در نتیجه مسئله دوئم به وجود می‌آید. اما در صورت موفقیت رابطه در آزمون آماری، تمام مجموعه‌های فرض‌های نظریه، فروض کمکی و مناسب بودن ساختار داده‌ها پذیرفته می‌شوند. در نتیجه، نظریه پذیرفته خواهد شد زیرا مولفه‌های این مجموعه سه عضوی به طور مستقل ساخته می‌شوند. در هر صورت، محقق باید راه‌هایی را که در نتیجه مسئله دوئم ممکن است بر نتایج آزمون تأثیرگذار باشد و ابطال‌پذیری نظریه را با چالش مواجهه کند، بررسی نماید و گزارش دهد.

۴. معنی‌داری آماری و نادیده گرفتن دستورالعمل‌های انجمن آمار آمریکا (ASA) در تفسیر نتایج

در انتشار تحقیقات تجربی مبتنی بر اقتصادسنجی معنی‌داری آماری نقش اساسی دارد، زیرا تنها یک رابطه علی از نظر آماری معنی‌دار است که یک نتیجه‌ی علمی پذیرفته شده است. با این حال، نباید فراموش کرد که اظهارات مربوط به معنی‌داری آماری مبتنی بر تصادفی بودن نمونه‌ی منتخب از داده‌ها ست. پیش از هرگونه توضیحی باید توجه داشت که در اقتصادسنجی، ساختارهای علی توسط مدل‌های رگرسیونی نشان داده می‌شوند که شامل حداقل یک متغیر وابسته[۱] (معمولا در اقتصادسنجی با حرف Y نشان داده می‌شود) است که توسط یک یا چند متغیر مستقل (با X نشان داده می‌شوند) توضیح داده می‌شود.[۲]

بخش مهمی از رگرسیون تو سط پارامترهای رگرسیون تو صیف می شود، که رابطه بین متغیر X و متغیر Y را با این فرض که سایر متغیرها ثابت بمانند (اصل ثبات سایر شرایط یا Ceteris Paribus) توصیف می‌کنند (وولدریج[۳]، فصل دوم، ۲۰۱۶). قدرت این رابطه با

---

[۱]. اگر مدل از نوع سیستم معادلات باشد در این صورت بیش از یک متغیر وابسته خواهیم داشت.

[۲]. شایان توجه است که وابسته یا مستقل بودن متغیرها امری اختیاری نیست. همچنین موضوعی نیست که از داده‌ها و مشاهدات بدست آمده باشد بلکه نظریه‌های اقتصادی و مدل‌های مبتنی بر آنها هستند که تعیین می‌کنند چه متغیری وابسته است و چه متغیری مستقل.

[3]. Wooldridge

استفاده از اندازه پارامترهای رگرسیون برآورد شده (براساس مشاهدات/داده‌ها) اندازه‌گیری می‌شود.[۱] به منظور بررسی اینکه به شراط ثبات سایر شرایط آیا رابطه‌ای واقعی میان متغیر های X و متغیر Y وجود دارد، معنی‌داری (آماری) پارامتر های مدل تخمین زده شده بررسی می‌شود. اگر یک پارامتر از نظر آماری معنی‌دار نباشد، فرض بر این قرار خواهد گرفت که رابطه‌ای بین X و Y وجود ندارد. با این وجود، چنین جمله‌ای مسئله‌ساز است. در واقع، جمله‌ی صحیح در این مورد چنین است: اگر پارامتر از نظر آماری معنی‌دار نباشد، براساس داده‌ها نمی‌توان ارتباط بین X و Y را تایید کرد. به عبارت دیگر، داده‌ها اطلاعات کافی را برای رد فرضیه صفر (عدم وجود رابطه بین X و Y) ارائه نمی‌دهند.

ارزش احتمال یا p -مقدار[۲] یکی از مفاهیم مهم در تحلیل‌های اقتصادسنجی است، که برای توصیف معنی‌داری آماری نتایج استفاده می‌شود. طبق تعریف، p-مقدار احتمال یافتن نمونه‌ای با نتیجه‌ای برابر یا بزرگتر از آنچه در آزمون آماری مشاهده شده است با فرض برقراری فرضیه‌ی صفر می‌باشد (جونسن[۳]، ۲۰۱۳). به منظور ارزیابی معنی‌داری آماری یک نتیجه، مقادیر حدی برای p-مقدار (سطح معنی‌داری) تعیین می‌شود. میزان رایج سطح معنی‌داری که در اقتصاد استفاده می‌شود در سطوح ۱٪ ، ۵٪ و ۱۰٪ تعیین می‌شود. تمام نتایج با p-مقدار کمتر از ۱٪ از نظر آماری معنی‌دار تلقی می‌شوند.

در ادبیات اقتصادی p-مقدار معمولا به طور نامناسبی استفاده و تفسیر می‌شود. برای مثال، p-مقدار معمولا به اشتباه به عنوان احتمال صحت فرضیه‌ی صفر تفسیر می‌شود. در حالی‌که فرضیه‌ی صفر متغیری تصادفی نیست و لذا هیچ احتمالی به آن تعلق نمی‌گیرد. فرضیه‌ی صفر فقط می‌تواند صحیح یا غلط باشد.[۴] همچنین مقایسه‌ی p-مقدار حاصل از نمونه‌های با اندازه‌های مختلف نادرست است. برای مثال مقادیر p-مقدار را همیشه می‌توان با بزرگ کردن نمونه افزایش داد و در نتیجه معنی‌داری نتایج را تقویت کرد.

استفاده و تفسیر نادرست p-مقدار باعث شده است که انجمن آمار آمریکا[۵] (ASA) مجموعه‌ای از قوانین را برای استفاده صحیح از p-مقدار در تحقیقات اتخاذ کند. مجموعه قوانین ASA به طور خلاصه عبارتند از:

---

[۱]. وولدریج در فصل دوم کتاب مقدمه‌ای بر اقتصادسنجی: یک رویکرد مدرن در مورد نحوه استخراج روابط علی با استفاده از رگرسیون توضیحات مفیدی ارائه کرده است.

[2]. p-Value

[3]. Johnson

[۴]. این به آن معناست که p-مقدار به طور غیرمستقیم و با واسطه نمونه‌ها درباره فرضیه صفر شواهدی را عرضه می‌کند.

[5]. American Statistical Association (ASA)

۱. مقادیر p-مقدار فقط باید به عنوان نشانه‌ی از ناسازگاری بین داده‌ها و مدل‌های پیشنهادی (تحت برقراری فرضیه صفر آزمون) استفاده شود. در واقع این ناسازگاری بین داده‌ها و مدل پیشنهادی تحت برقراری فرضیه صفر را می‌توان به عنوان شواهدی علیه فرضیه صفر تبیین کرد.

۲. مقادیر p-مقدار باید به عنوان گزاره‌ای در مورد ارتباط داده‌ها با فرضیه صفر تفسیر شود و نه به عنوان جمله‌ای در مورد خود فرضیه صفر،

۳. نتیجه‌گیری‌های علمی و تصمیمات اقتصادی یا سیاسی نباید تنها مبتنی بر این معیار که آیا p-مقدار از حد مشخصی فراتر می‌رود یا نه باشند، بلکه براساس عوامل دیگری مانند شواهد بیشتر برای پدیده مورد بررسی یا کیفیت شیوه اندازه‌گیری تفسیر شوند. در واقع تبیین علمی را نباید به اینکه آیا p-مقدار از مقدار خاصس فراتر می‌رود یا نه کاهش داد. چنین رویکردی منجر به تصمیم‌گیری ضعیف و ایجاد باورهای غلط خواهد شد. محققان به منظور استنباط علمی، باید عوامل متنوعی زیادی را وارد بازی کنند، از جمله طراحی یک مطالعه، کیفیت اندازه‌گیری‌ها، شواهد خارجی برای پدیده مورد بررسی و اعتبار فروضی که زیربنای آنالیز داده‌ها است.

۴. محققان باید اطلاعات کاملی از پژوهش خود مانند تعداد فرضیه‌های مورد بررسی، همه‌ی مقادیر p-مقدار محاسبه شده و سایر محاسبات آماری را گزارش کنند. در واقع نباید تنها براساس مقادیر p-مقدار که شرایط آستانه‌ای (برای مثال $p<0.05$) نتایج تحقیق را گزارش کنند. زیرا این موضوع منجر به تحریف روند پژوهش خواهد شد.

۵. مقادیر p-مقدار یا معنی‌داری آماری متری[1] برای اندازه یک اثر یا اهمیت آن نیست. در نتیجه علاوه بر مقادیر p-مقدار باید اندازه‌ی اثر علی را نیز گزارش نمود این امر با توجه به این واقعیت انجام می‌شود که در عمل می‌توان با تغییر حجم نمونه برای مثال بزرگ کردن حجم آن معنی‌داری را افزایش داد. همچنین اندازه یک اثر می‌تواند با وجود کوچک بودن ویا حتی بی اهمیت بودن p-مقدار کوچکی را تولید کند. چنین گزارشی از این موضوع جلوگیری می‌کند.

۶. پژوهشگران باید از محدودیت‌های p-مقدار آگاهی کامل داشته باشند، (واسرستین و لازار[2]، ۲۰۱۶).

---

[1]. Measure

[2]. Wasserstein & Lazar

توصیه‌های ASA همچنین واکنشی به استفاده روزافزون نادرست از p-مقدار در شیوه نشر نتایج علمی را نشان می‌دهد. برخلاف علوم طبیعی، هیچ تکراری از نمونه یا مشاهدات در اقتصاد نمی‌تواند انجام شود. این امر به ویژه در مورد اقتصاد کلان و امور مالی صدق می‌کند، جایی‌که داده‌های اساسی به شکل سری زمانی هستند. یک سری زمانی دنباله‌ای وابسته به زمان از داده‌هاست که قابل تکرار نیستند. این محدودیت دلالتی روشنی برای ارزش آموزنده نظریه‌های علم اقتصاد دارد.

همانطور که ذکر شد، درجه صحت اندازه‌گیری پارامترهای مدل و بنابراین معنی‌داری آماری آنها با افزایش حجم نمونه می‌تواند افزایش یابد. در حالی‌که این عمل در علم امکان‌پذیر است اما در تحقیقات اقتصاد کلان و اقتصاد مالی به دلیل کمبود اطلاعات غیرقابل انجام است (برای مثال مانند پیش‌بینی بحران مالی سال ۲۰۰۸ میلادی براساس بحران‌های قدیمی و مشابه). در نتیجه‌ی این محدودیت‌ها، باید از توصیه‌های ASA پیروی کرد و نه تنها معنی‌داری بلکه اهمیت و اندازه مقادیر عددی پارامترهای تخمین زده شده را نیز در نظر گرفت. بنابراین، محقق برای تفسیر و پیشنهادات سیاستی معنی‌داری آماری و اندازه را باید در کنار هم در نظر گیرد تا یک جمله‌ی علمی معتبر اقتصادی ارائه کند.

موضوع مهم دیگر که در مورد تفسیر p-مقدار بسیار مهم است نحوه انتخاب سطوح آستانه‌ای معنی‌داری در شرایط مختلف است. این سطوح آستانه‌ای باید به نحوی صورت گیرد که مجموعه احتمال خطای نوع اول و دوم را در یک آزمون آماری کاهش دهد.

به منظور توضیح بیشتر در این زمینه، می‌دانیم که زمانی آزمون یک فرضیه به صورت آماری؛ دو نوع خطا روی می‌دهد. خطای نوع اول در شرایطی که فرضیه صفر صحیح رد شود روی می‌دهد. این خطا بعنوان خطای نوع I ("مثبت کاذب") ذکر می‌شود. خطای نوع دوم یا خطای نوع II ("منفی کاذب")، در شرایطی روی می‌دهد که فرضیه صفر غلط رد نشود. احتمال خطای نوع I به طور مستقیم توسط محققان از طریق انتخاب سطح معنی‌داری کنترل می‌شود. هنگامی که یک آزمایش در سطح ۵٪ انجام می‌شود، احتمال ارتکاب خطای نوع I ۵٪ است. اغلب به این احتمال (سطح معنی‌داری) اندازه آزمون[1] گفته می‌شود. احتمال ارتکاب خطای نوع II به مقادیر واقعی پارامتر بستگی دارد. به طور شهودی، اگر مقادیر بیان شده در فرضیه صفر از مقدار واقعی پارامتر فاصله داشته باشد، احتمال ارتکاب خطای نوع II نسبتا اندک خواهد بود. در حالی‌که اگر مقدار بیان شده در فرضیه‌ی صفر نزدیک به مقدار واقعی پارامتر باشد، میزان احتمال ارتکاب این خطا بسیار بزرگ

[1]. Size of Test

خواهد بود. احتمال رد فرضیه صفر زمانی که نادرست است، به عنوان توان آزمون[1] شناخته می‌شود. توان آزمون نشان می‌دهد که یک آزمایش "قدرتمند" در یافتن انحراف از فرضیه صفر (بسته به مقدار پارامتر واقعی) چقدر است. بین خطاهای نوع I و نوع II یک تبادل برقرار است به طوری که کاهش اندازه یک آزمون باعث کاهش توان آن می‌شود. هرچه حجم نمونه بزرگتر باشد احتمال ارتکاب خطای نوع II کاهش خواهد یافت. در این شرایط محققان اندازه آزمون را به منظور کاهش خطای نوع I در سطح پایین ۱٪ تنظیم می‌کنند. بنابراین، برای تفسیر مناسب p-مقدار سطح آستانه‌ای معنی‌داری را باید در ۱٪ انتخاب کرد. برعکس، چنانچه حجم نمونه کم باشد برای کاهش خطای نوع II باید اندازه آزمون را افزایش داد و برای نمونه سطوح ۵٪ و ۱۰٪ را انتخاب کرد (کامین[2]، ۲۰۱۲، فصل دو).

۵. خلاصه و نتیجه‌گیری

در پژوهش حاضر دو مشکل در انجام آزمون فرضیه‌های اقتصادی بررسی گردید که چنین بررسی پیش از این در ادبیات اقتصادی ایرانی صورت نگرفته است. مورد اول که به کرات در ادبیات اقتصادی ایران نادیده گرفته می‌شود مسئله دوئم در آزمون نظریه‌ها یا فرضیه‌های علمی نامیده می‌شود. همانگونه که توضیح داده شد، مسئله دوئم ابطال پذیری نظریه‌های علمی را زیر سوال می‌برد. در نظر دوئم فروض کمکی مدل‌ها که برای بررسی نظریه‌های علمی مورد استفاده قرار می‌گیرند، رابطه ابطال پذیری توصیف شده در رابطه (۲) را به صورت زیر تغییر می‌دهند:

$$\frac{\begin{array}{c} h\&a \Rightarrow e \\ \acute{e} \end{array}}{\acute{h} \vee \acute{a}}$$

دراین رابطه علامت $\vee$ ربط منطقی یا بین دو گزاره a' و h' است. در این رابطه برخلاف رابطه‌ی (۲) مشاهده شواهد (آماری) e' دال بر رد h نیست. زیرا ممکن است که h درست و a غلط باشد. بنابراین، در حضور فروض کمکی، نتیجه آزمون فرضیه‌های اقتصادی مبهم خواهد بود.

با توجه به این توضیحات، به نظر می‌رسد که هرچه یک مدل آماری مفروضات کمکی کمتری داشته باشد مشکلات کمتری هم ایجاد کند. مفروصات بیشتر به معنای محدودیت بیشتر مدل است که در چنین وضعیتی احتمال تطابق داده‌های مشاهده شده با مفروضات

---

[1]. Power of Test
[2]. Cumming

کمتر خواهد بود. با توجه به این موضوع، به نظر می‌رسد از آنجاکه در مقایسه با آمار پارامتری تنها فرض آمار ناپارامتری این است که نمونه‌ی انتخاب شده از جامعه باید تصادفی باشد؛ لذا آمار ناپارامتری در مدل‌سازی نظریه‌های اقتصادی مفیدتر باشد. اما مدل‌سازی ناپارامتری حاوی ضعف‌هایی همچون مشقت بعدچندی[1] و دقت پایین در مقایسه با مدل‌سازی پارامتری است (پاگان و اله[2]، ۲۰۰۰). مشقت بعدچندی برای معنی‌داری آماری زمانی که تعداد متغیرهای مورد استفاده در یک رابطه افزایش می‌یابد مشکل ایجاد می‌کند. در پژوهش حاضر نقش فروض کمکی در بررسی یک نظریه در اقتصاد مالی تحت عنوان فرضیه بازارهای کارا به طور کامل و با جزییات دقیق انجام گردید که پیش از این در بستر مسئله دوئم و آزمون ناپذیری نظریه‌های اقتصادی بررسی نشده بود.

در یک تعریف گسترده‌تر همه فروضی که در هنگام استنباط ادعاهای قابل آزمایش از یک نظریه اقتصادی بکار گرفته می‌شوند را می‌توان فروض کمکی نامید. در برخی شرایط موضوع به قدری نامشخص است که بدون بررسی دقیق هیچ کس متوجه وجود یک فرض کمکی پنهان و موثر نمی‌شود و همان فرض ممکن است نتایج بررسی نظریه را نامعتبر سازد. این فروض می‌تواند حتی مربوط به ساختار داده‌ها، آزمون‌های آماری و توزیع حاشیه‌ای داده‌ها داشته باشند. پیامد نادیده گرفتن نقائص در این زمینه، منجر به زیر سوال رفتن نظریه گردد. در حالی که این فروض مدل‌های آماری است مطابق با واقعیت نیست! در برخی شرایط آزمون‌های تصریح مدل کمک می‌کنند تا آزمون فرضیه یا نظریه ارتباطی با فروض کمکی مدل نداشته باشد اما در شرایطی دیگر چنین آزمون‌های برای جداسازی فروض کمکی و فروض نظریه وجود ندارد. براساس این یافته‌ها، پیشنهاد می‌شود که هنگام تجزیه و تحلیل استدلال؛ ابتدا فروض و پیامدهای تجربی شناسایی شوند. چنانچه نظریه و فرضیه‌ای ممکن است پیامدهای عمیقی در سیاست‌گذاری اقتصادی یا اجتماعی در پی داشته باشد بهتر است که فروض کمکی مربوط به مدل‌سازی و داده‌ها به طور صریح بیان شوند تا نتایج در پرتو این اطلاعات مورد تجزیه و تحلیل قرار گیرد.

موضوع با اهمیت دیگری که در بررسی پژوهش‌های اقتصادی باید مور توجه قرار گیرد، شیوه گزارش و تفسیر معنی‌داری آماری یا p-مقدار است. در سال‌های اخیر با افزایش حجم تحقیقات علمی و داده‌های اقتصادی و اجتماعی دامنه کاربردهای روش‌های آماری گسترش پیدا کرده است. این امر راه‌های جدیدی را برای پیشرفت علمی ایجاد کرده است، اما

[1]. Curse of Dimensionality
[2]. Pagan and Ullah

نگرانی‌هایی راجع به نتیجه‌گیری‌های آماری را نیز به همراه دارد. اعتبار نتیجه‌گیری‌های علمی، علاوه بر ارتباط با تکنیک‌های مناسب انتخاب شده به تجزیه و تحلیل های درست و تفسیر درست از نتایج آماری ارائه شده توسط معیار p-مقدار نیز دارد و برای اطمینان از صحت نتیجه‌گیری‌ها باید به توصیه‌های ASA در نحوه گزارش و تفسیر p-مقدار نیز توجه داشت.

در نهایت، پژوهش آماری مطلوب به عنوان یک مؤلفه اساسی پژوهش علمی؛ بر اصول طراحی مناسب مطالعه، انواع خلاصه‌های عددی و نموداری داده‌ها، درک مناسب پدیده مورد مطالعه، تفسیر نتایج در متن، گزارش کامل، صحیح و منطقی و فهم کمی مناسب از معنای خلاصه اطلاعات تأکید دارد و هیچ شاخص واحدی همچون p-مقدار نباید جایگزین استدلال علمی شود.

منابع:

رستمی، مجتبی و مکیان، سید نظام الدین. (۱۳۹۸)، آزمون ریشه واحد بیزی با لحاظ مشاهدات پرت: مطالعه موردی بازده روزانه ۵۰ شرکت فعال بورس تهران. *مدلسازی اقتصادسنجی*،۴(۳): ۸۶-۵۹.

Bodie, Z., Kane, A. & Marcus, A. J. (2018), Essentials of investments (pp. 532-537), Taipei: McGraw-Hill/Irwin.

Duhem, P. (1904), The Aim and Structure of Physical Theory, Translated by Philip P. Wiener, 1974.

Engle R.F. (2004), Risk and Volatility: Econometric Models and Financial Practice, The American Economic Review, 94(3): 405-420.

Hamilton, A. G. (1988), Logic for mathematicians. Cambridge University Press.

Harding, S. (Ed.). (1975), Can theories be refuted? Essays on the Duhem-Quine thesis (Vol. 81). Springer Science & Business Media.

Johnson, V. E. (2013), Revised standards for statistical evidence. Proceedings of the National Academy of Sciences, 110(48): 19313-19317.

Koop, G. (1991), Cointegration Tests in Present Value Relationships," Journal of Econometrics, 49: 105-139.

Koop, G. (1994), An Objective Bayesian Analysis of Common Stochastic Trends in International Stock Prices and Exchange Rates, Journal of Empirical Finance, 1(3-4): 343-364.

Makyian, S. N. & Rostami, M. (2019), Political Stability, Corruption, Democracy and Terrorism in the Middle East and North Africa. International Journal of Business and Development Studies, 11(1): 59-82.

Mandelbrot, B. (1963), The Variation of Certain Speculative Prices, Journal of Business, 36(4): 394–419.

Mookerjee, R. & Yu, Q. (1997), Macroeconomic variables and stock prices in a small open economy: The case of Singapore. Pacific-Basin Finance Journal, 5(3): 377-388.

Pagan, A. & Ullah, A. (1999), Nonparametric econometrics, Cambridge university press.

Popper, K. (2005), The logic of scientific discovery, Routledge.

Schweiker, M., Hass, J., Novokhatko, A. & Halbleib, R. (2017), Messen und Verstehen in der Wissenschaft, Springer Fachmedien Wiesbaden.

Schwert, G. W. (1989), Why Does Stock Market Volatility Change over Time?, Journal of Finance, 44(5): 1115–1153.

Shearmur, J. (1991), Common Sense and the Foundations of Economic Theory: Duhem versus Robbins, Philosophy of the Social Sciences, 21: 64-71.

Sims, C. A. (1988), Bayesian Skepticism on Unit Root Econometrics, Journal of Economic Dynamics and Control, 12(2-3): 463-474.

Spanos, A. (2019), Probability Theory and Statistical Inference: Empirical Modeling with Observational Data, Cambridge University Press.

Wainwright, K. (2005), Fundamental methods of mathematical economics/Alpha C. Chiang, Kevin Wainwright. Boston, Mass.: McGraw-Hill/Irwin.

Wasserstein, R. L. & Lazar, N. A. (2016), The ASA Statement on p-values: Context, Process, and Purpose.

Wooldridge, J. M. (2016), Introductory econometrics: A modern approach. Nelson Education.