فصلنامه پژوهش‌های اقتصادی ایران / سال نهم/ شماره ۳۲/ پاییز ۱۳۸۶/ صفحات ۱۱۱- ۱۳۶

# رگرسیون جعلی: مفهوم و نتایج

علی‌حسین صمدی*

تاریخ ارسال: ۱۳۸۵/۵/۴ تاریخ پذیرش: ۱۳۸۶/۹/۲۷


## چکیده

*ایده رگرسیون جعلی در اقتصادسنجی توسط گرنجر و نیوبولد (۱۹۷۴) مطرح شد. این محققان نشان دادند که اگر متغیرهای مستقل و وابسته مورد استفاده در الگو $I(1)$ باشند، تخمین الگو به روش حداقل مربعات معمولی باعث پیدایش نتایج غیر واقعی( یا جعلی) خواهد شد . اما سری‌های زمانی مورد استفاده در الگوها ممکن است خاصیت جمع بستگی متفاوتی داشته باشند. بنابراین، این سوال مطرح می‌گردد که آیا پدیده رگرسیون جعلی در الگوهای دارای درجات متفاوت جمع بستگی نیز وجود خواهد داشت یا خیر. در این مقاله پس از بررسی سیر تکامل تاریخی و مفهوم رگرسیون جعلی، نتایج یافته‌های برخی از مطالعات مهم در ارتباط با داده‌های سری زمانی با الگوهای دارای متغیرهای $I(1)$ و $I(2)$ و همچنین $I(1)$ و $I(0)$مد نظر قرار گرفته و از بررسی متون آشکار گردیده است که در چنین شرایطی نیز احتمال وقوع رگرسیون جعلی وجود دارد. بنابراین پیشنهاد شده است که قبل از اقدام به تخمین هر الگویی، خواص سری‌های زمانیِ مورد استفاده، دقیقاً بررسی شود.*


**طبقه بندی JEL:** C13,C22,C52

**واژگان کلیدی:** رگرسیون جعلی، داده‌های سری زمانی.

---

* عضو هیات علمی دانشگاه شیراز

Email : asamadi @ rose.shirazu.ac.ir.

**مقدمه**

مفهوم *رگرسیون جعلی(بی معنی، دروغین یا کاذب)* [۱] یا *همبستگی‌های بی معنی*[۲] در متون آمار توسط *یول*[۳] ( ۱۹۲۶) و در متون اقتصادسنجی توسط گرنجر و نیوبولد ( ۱۹۷۴ )[۲۲] * معرفی شد. نتایج حاصل از شبیه‌سازی مونت‌کارلوی گرنجر و نیوبولد، بعدها توسط سایر محققین آزمون و بسط داده شد. نلسون و کنگ (۱۹۸۳ و ۱۹۸۱) [۴۴و ۴۳] نیز همانند گرنجر و نیوبولد با استفاده از شبیه‌سازی مونت‌کارلو نشان دادند که پدیده رگرسیون جعلی، زمانی رخ می‌دهد که رگرسیون شامل *متغیرهای نامانای ریشه واحد*[۴] باشد. فیلیپس (۱۹۸۶) [۴۸] برای اولین بار نتایج گرنجر و نیوبولد (۱۹۷۴) [۲۲] را به صورت تحلیلی و به شکل ریاضی ارائه داد. این نتایج در همیلتون (۱۹۹۴) [۲۵] نیز ارائه شده است. هندری (۱۹۹۵) [۲۷] نیز توضیح مفیدی ارائه داده است.

در کلیه مطالعات بالا، متغیرهای نامانای در نظر گرفته شده در الگوهای رگرسیونی از مرتبه (۱) می‌باشند. مارمول (۱۹۹۵) [۳۷]، نتایج فیلیپس (۱۹۸۶) را به مواردی با *فرایندهای جمع بسته*[۵] از درجات بالاتر از (۱) و در حالت کلی از درجه dام (یعنی فرایندهای $\mathrm{I(d)}$) تعمیم داد که در آن، d یک عدد صحیح می‌باشد.

مسئله رگرسیون جعلی همچنین ممکن است در موقعیت‌هایی رخ دهد که فرایندهای موجود در آن، *مرتبه‌های جمع بستگی*[۶] متفاوتی داشته باشند. به این مسئله در مطالعات هالدراپ (۱۹۹۴) [۲۴]، مارمول (۱۹۹۶) [۳۸] و هاسلر (۱۹۹۶) [۲۶] توجه شده است. هالدراپ (۱۹۹۴) این مسئله را در حالتی بررسی کرد که در آن، متغیرهای مستقل $I(۱)$ و $I(۲)$ بودند. هاسلر (۱۹۹۶) و مارمول (۱۹۹۶) نیز موقعیتی را تحلیل کردند که در آن، متغیرهای مستقل $I(۱)$ و $I(۰)$ می‌باشند.

در بررسی رگرسیون‌های جعلی، فرض بر این بوده است که رابطه همجمعی[۷] به مفهوم انگل-گرنجر بین متغیرها وجود نداشته باشد. اما چوی (۱۹۹۴) [۱۰] مسئله رگرسیون جعلی را هنگامی بررسی کرد که رگرسورهای $I(۱)$ نامانای $\mathrm{M}_2$ بعدی[۸] رابطه همجمعی داشتند.

فیلیپس (۱۹۹۸ و ۱۹۸۶) [۵۰ و ۴۸] و دارلاف و فیلیپس (۱۹۸۸) [۱۳] به بررسی *خواص مجانبی*[۹] برآوردگرها و *آماره آزمون*[۱۰] برای ضرایب رگرسیونی در رگرسیون های جعلی پرداختند. دارلاف و فیلیپس (۱۹۸۸) در رگرسیونی که در آن، *روندهای تصادفی*[۱۱] روی *رگرسورهای چندجمله‌ای معین (غیرتصادفی)*[۱۲] رگرس می‌شود، به نتایجی مشابه نتایج فیلیپس (۱۹۸۶) دست یافتند. همچنین فیلیپس (۱۹۹۸) بعضی ابزارهای جدید برای تحلیل و فهم رگرسیون های جعلی ارائه داده است.

اوگاکی و چوی (۲۰۰۱) [۴۵] برای اولین بار *خواص دقیق نمونه متناهی*[۱۳] برآوردگرها و آماره آزمون برای ضرایب رگرسیون های جعلی با متغیرهای نامانای ریشه واحد را مطالعه کرده‌اند.

---

* اعداد داخل پرانتز اشاره به سال انتشار مقاله یا کتاب و اعداد داخل کروشه اشاره به شماره مأخذ در انتهای مقاله دارند.

گرنجر و نیوبولد (۱۹۷۴) و فیلیپس (۱۹۸۶) در تحلیل خـود فـرض کـرده‌انـد کـه در یـک الگـوی رگرسیون ساده خطی، متغیرهای x و y متغیرهـای *گـام تـصادفی*(۱۴) مـستقل بـدون *جملـه رانـش*(۱۵) می‌باشند. اما انتورف (۱۹۹۷) [۱۷] ثابت کرد که اگر متغیرهای x و y متغیرهای گام تصادفی با جملـه رانش غیرصفر باشند، نتایج دیگری می‌توان بـه دسـت آورد. گرنجـر و همکـاران (۲۰۰۱) [۲۳] مـسئله رگرسیون جعلی را در رگرسیون هایی مطالعه کرده‌اند که در آنها x و y *فرایندهای مانای مستقل بـدون هرگونه مولفه روند*(۱۶) می‌باشند. کـیم و همکـاران (۲۰۰۴) [۳۱] نیـز تمرکـز خـود را روی فراینـدهای مستقل مانا در اطراف روندهای خطی قرار داده‌اند.

رگرسیون جعلی همچنین ممکن است در حالتی رخ دهد که لگاریتم یک فرایند جمع بـستهماننـد log y بر لگاریتم فرایند جمع بسته دیگری مانند logx رگرس شود. دوجانگ (۲۰۰۳) [۱۵] شـرایطی را بررسی کرده است که در آن، چنین پدیده‌ای رخ می‌دهد. تـوو و همکـاران (۲۰۰۴) [۶۱] نیـز وقـوع چنین پدیـده‌ای را در تحلیـل‌هـای همبـستگی و رگرسـیونی بـا *متغیرهـای نـسبتی*(۱۷) یـا *متغیرهـای شاخص*(۱۸) بررسی کرده‌اند و علت وقوع چنین پدیده‌ای در این‌گونه رگرسیون‌ها را به مـسائل ریاضـی و تصریح نامناسب الگو نسبت داده‌اند. در این مطالعه بررسی جامعی از چنین الگوهایی صورت گرفته است.

در کلیه مطالعات بالا فرض بر این بوده که مرتبه جمع بستگی (d) متغیرها یک عدد صحیح است و بنابراین پس از d بار تفاضل‌گیری به فرایندهای مانا دست خواهیم یافت. چنین فرضی باعث ناپیوستگی می‌شود و مشکلاتی را در تفسیر آماری نتایج روش های تخمین و یا خواص آماری آنها ایجـاد مـی‌کنـد. چنین مشکلاتی در متون دارای *فرآیندهای کسری*(۱۹) یا *الگوهای بلند حافظه*(۲۰) حل شده است. مارمول (۱۹۹۸) [۳۹] متون موجود در رابطه با رگرسیون های جعلی را بـسط داده و بـا یـک مطالعـه تحلیلـی ریاضی، وقوع چنین پدیده‌ای را با *فرایندهای نامانای جمع بسته کسری*(۲۱) FI(d) به اثبات رسانده است که در آن d یک عدد واقعی (و نه صحیح) می‌باشد.(۲۲) تسی و چیونگ (۲۰۰۰) [۶۰] نیز تحلیل نظـری رگرسیون های جعلی و *روندزدایی جعلی*(۲۳) را از فرایندهای $I(1)$ بـه فراینـدهای *جمـع بـسته کـسری بلندحافظه*(۲۴) و در واقع به کلاسی از فرایندهای FI(d) بلند حافظه بسط داده‌اند. این محققین شرایطی را تحلیل کرده‌اند که در آن، فرایندهای I(d) منجر به اثرات جعلی در یک الگوی خطی ساده مـی‌شـود. ذکر این نکته ضروری است که مرتبه جمع بستگی (d) در مطالعات مارمول (۱۹۹۸) و تسی و چیونـگ (۲۰۰۰) یک عدد واقعی است و نه عدد صحیح.

مطالعات اشاره شده در سطور بالا، پدیـده رگرسـیون جعلـی را در داده‌هـای سـری زمـانی بررسـی کرده‌اند؛ اما انتورف (۱۹۹۷) و به طـور جداگانـه کـائو (۱۹۹۹) [۲۹] چنـین پدیـده‌ای را در *داده‌هـای تلفیقی*(۲۵) و فینگلتون (۱۹۹۹) [۱۸] در *داده‌های فضایی*(۲۶) معرفی کرده‌انـد. کـائو (۱۹۹۹) چـارچوبی برای فهم رگرسیون جعلی در داده‌های تلفیقی بسط داده و خواص مجانبی *برآوردگر متغیرهای مجـازی حداقل مربعات (LSDV)*(۲۷) و سایر آماره‌های مرسوم را مانند آماره t برای رگرسیون جعلـی در چنـین داده‌هایی استخراج کرده است. همچنین با شبیه‌سازی مونت‌کارلو، خواص نمونه متنـاهی برآوردگرهـای

LSDV، ضریب تعیین ($R^2$) و آماره دوربین-واتسون (DW) را بررسی کرده است. فینگلتون (۱۹۹۹) نیز با استفاده از شبیه‌سازی مونت‌کارلو، مفهوم *رگرسیون های جعلی فضایی*[۲۸]، *ریشه واحد فضایی*[۲۹] و همجمعی *فضایی*[۳۰] را معرفی کرده است.

در قسمت بعدی این مقاله مفهوم رگرسیون جعلی ارائه می‌شود. نتایج یافته‌های برخی از مطالعات مهم بالا در ارتباط با داده‌های سری زمانی نیز در ادامه این مقاله ارائه خواهد شد.

نتایج یافته‌های برخی از مطالعات مهم دیگردر ارتباط با رگرسیون جعلی با داده‌های تلفیقی، داده‌های فضایی، داده‌های لگاریتمی، داده‌های شاخص یا نسبتی و فرایندهای جمع بسته کسری و فرایندهای مانای گه گن بایر در قسمت دوم مقاله حاضر و به صورت مقاله ای مجزا ارائه خواهد شد.

## ۱. مفهوم رگرسیون جعلی

الگوی رگرسیون خطی ساده زیر مفروض است:

$$Y = \alpha + \beta X + \varepsilon \qquad (1)$$

با این الگو می‌توان رابطه بین دو متغیر مستقل X و Y را بررسی کرد. این کار را می‌توان با تخمین ضرایب $\alpha$ و $\beta$ و انجام آزمون‌های آماری برای معنی‌داری ضرایب تخمین زده شده و معناداری کل رگرسیون و استفاده از برخی *آزمونهای تشخیصی*[۳۱] از جمله آماره باکس-پیرس، آماره d دوربین-واتسون و .... انجام داد. بنابراین ملاحظه می‌گردد که سه مرحله بسیار مهم در مطالعات کاربردی و در اقتصادسنجی کاربردی مدنظر می‌باشد[۳۲]: تصریح[۳۳]، تخمین[۳۴] و ارزیابی[۳۵] الگو[۳۶].

یکی از روش های مرسوم برای تخمین الگوی رگرسیونی شماره (۱) روش حداقل مربعات معمولی (OLS) است. استفاده از این روش، مستلزم تامین شرایطی است که یکی از مهم‌ترین آنها *مانا بودن*[۳۷] متغیرهای X و Y است. یک سری مانا دارای میانگین و واریانس ثابت در طول زمان است. به عبارت دیگر هرگاه زیرنمونه‌ای از نمونه اصلی را انتخاب و میانگین آن را حساب کنیم، میانگین این زیر نمونه نباید اختلاف معنی‌داری از میانگین کل دوره داشته باشد. احتمال وقوع خطای نوع اول با داده‌های مانا حدود ۵٪ مواقع می‌باشد.

هرگاه الگوی شماره (۱) با داده‌های نامانا و به روش OLS تخمین زده شود، ضریب تعیین الگو افزایش می‌یابد و به عدد یک نزدیک می‌گردد؛ جملات پسماند آن دچار خودهمبستگی می‌شود و آماره d دوربین-واتسون (.D.W) به سمت صفر نزدیک می‌گردد. قدر مطلق آماره t بیشتر از ۱/۹۶ می‌شود و در کل آماره‌های آزمون‌های تشخیصی، معنی‌داری آماری *بیش از حد*[۳۸] از نتایج تخمین را نشان می‌دهند. بنابراین نامانا بودن داده‌ها باعث پیدایش نتایج غیرواقعی خواهد شد اگر چنین نتایجی واقعاً وجود نداشته باشد. به چنین نتایجی، پدیده *«رگرسیون جعلی»* اطلاق شده است. این پدیده اولین بار توسط یول (۱۹۲۶) در متون آمار معرفی شد و گرنجر و نیوبولد (۱۹۷۴) [۲۲] با استفاده از آزمایشات

شبیه‌سازی مونت‌کارلو آن را در متون اقتصادسنجی دوباره معرفی کردند و سپس توسط فیلیپس (۱۹۸۶) [۴۸] در چارچوب تحلیل ریاضی تشریح شد.

یکی از معیارهای تقریبی برای تشخیص چنین پدیده‌ای آن است که مقدار ضریب تعیین الگوی تخمین زده شد، $(R^2)$ بزرگتر از آماره d دوربین-واتسون (یعنی $D.W. < R^2$) به‌طور مثال می توان به صورت تجربی در مطالعه خلیلیان و حفار اردستانی (۱۳۷۹) وقوع چنین امری را ملاحظه کرد. صمدی (۱۳۸۱) نیز با تکرار مطاله خلیلیان و حفار اردستانی (۱۳۷۹) به جعلی بودن نتایج اشاره کرده است. علت وقوع چنین امری با توجه به مطالب پیش‌گفته کاملاً نمایان است. اخیراً فیلیپس (۱۹۹۸) [۵۰] روش‌های پیشرفته و جدیدی برای فهم این پدیده ارائه داده است. اما این نابرابری را می‌توان به عنوان نشانه‌ای از فقدان هرگونه رابطه تعادلی بین متغیرها در رگرسیون تفسیر کرد و این در *رگرسیون های ناتراز یا ناسازگار*(۳۹) رخ می‌دهد.

## ۲. گرسیون جعلی با داده‌های سری زمانی

پدیده رگرسیون جعلی در ابتدا با داده‌های سری زمانی توسط گرنجر و نیوبولد (۱۹۷۴) [۲۲] معرفی گردید. سپس انتورف (۱۹۹۷) [۱۷] و به طور جداگانه کائو (۱۹۹۹) [۲۹] رگرسیون جعلی را با داده‌های تلفیقی مطرح ساختند. در همین سال فینگلتون (۱۹۹۹) [۱۸] ضمن معرفی مفهوم ریشه واحد فضایی، احتمال وقوع چنین پدیده‌ای را با داده‌های فضایی را بررسی کرد. در ادامه نتایج برخی مطالعات مربوط به داده های سری زمانی ارائه خواهد شد.

### ۱-۲. رگرسیون جعلی با متغیرهای نامانا

#### ۱-۱-۲. متغیرهای $I(1)$

الگوی رگرسیون خطی ساده زیر را در نظر بگیرید.

$$Y_t = \alpha + \beta X_t + u_t \qquad (۲)$$

فرض کنید که $Y_t$ و $X_t$ متغیرهای گام تصادفی مستقل بدون جمله رانش(۴۰) باشند و به صورت زیر تعریف گردند:

$$Y_t = Y_{t-1} + v_t \quad , \quad X_t = X_{t-1} + w_t \quad , \; t = 1, 2, 3, \ldots \qquad (۳)$$

که در آن، $v_t \sim IID(0, E_v^2)$ و $w_t \sim IID(0, E_w^2)$ است.

فرض کنید که دنباله $\{\zeta\}_1^\infty$ در فضای احتمال (π, β, ρ) تعریف شده و $S_t = \sum_{j=1}^{t} \xi_j$ *فرایند مجموع جزئی*(۴۱) و $S_0 = 0$ باشد. همچنین فرض کنید که

***فرض ۱:***

الف) برای کلیه tها، $E(\xi_t)=0$ است.

ب) برای بعضی از مقادیر $\beta > 2$ و $\varepsilon > 0$ داریم:

$$SUP_{i,t}\, E|\xi_{it}|^{\beta+\varepsilon} < \infty$$

ج) مقدار $\Sigma = \lim_{T\to\infty} T^{-1} E\left(S_T S_T'\right)$ وجود داشته و *معین مثبت(یا مثبت قطعی)*(۴۲) است.

د) دنباله $\{\xi\}_1^\infty$ *آمیخته قوی*(۴۳) با *اعداد آمیخته*(۴۴) $\alpha_m$ است و در شرط $\sum_1^\infty \alpha_m^{1-2/\beta} < \infty$ صدق می‌کند.

گرنجر و نیوبولد (۱۹۷۴) [۲۲] با استفاده از آزمایشات شبیه‌سازی مونت‌کارلو و با در نظر گرفتن شرایط اولیه $Y_0 = X_0 = 100$ و استخراج متغیرهای $v_t$ و $w_t$ از جامعه‌ای مستقل با $N(0,1)$ و با اشاره به اینکه اکثر سری های زمانی اقتصادی نامانا هستند و اغلب به صورت *نزدیک به گام تصادفی*(۴۵) ظاهر می‌گردند، نشان دادند که نتایج تخمین الگوی (۲) به روش OLS و با چنین متغیرهایی منجر به ضریب تعیین $(R^2)$ بسیار بالا و آماره d دوربین-واتسون بسیار پایین (جملات پسماند دارای خودهمبستگی بسیار زیاد) و بنابراین وقوع پدیده رگرسیون جعلی می‌گردد. گرنجر و نیوبولد (۱۹۷۴) برای ارزیابی معنی‌داری آماری ضریب $X_t$ در سطح ۵٪ پیشنهاد می‌کنند که به جای مقدار بحرانی مرسوم ۱/۹۶ باید از مقدار بحرانی جدید ۱۱/۲ استفاده گردد. اما فیلیپس (۱۹۸۶) [۴۸] اشاره می‌کند چنین پیشنهادی هیچ مبنایی بر اساس نظریه مجانبی ندارد و باید از فرمول $t_\beta' = t_\beta \big/ \sqrt{T}$ (و با مثال گرنجر و نیوبولد با T=۵۰، $11/2\big/\sqrt{50}$) استفاده گردد.

این نتایج همچنین توسط فیلیپس (۱۹۸۶) [۴۸] و در یک چارچوب تحلیلی و با فرضی ضعیف‌تر از فرض در نظر گرفته شده توسط گرنجر و نیوبولد (۱۹۷۴) درباره *عامل های ابداع*(۴۶) در معادله (۳) به دست آمده است. فیلیپس (۱۹۸۶) اشاره می‌کند که اگر n=۲ و $\xi_t' = (v_t, w_t)$ باشد، در این صورت شرایط اشاره شده در *فرض ۱* درباره عاملهای ابداع معادله شماره (۳) نسبتاً ضعیف خواهد شد.

فیلیپس (۱۹۸۶) [۴۸] نشان داده است که چنانچه دنباله‌های $\{Y_t\}_1^\infty$ و $\{X_t\}_1^\infty$ از طریق معادله (۳) تولید و دنباله‌های عامل ابداعی $\{v_t\}_1^\infty$ و $\{w_t\}_1^\infty$ مستقل باشند و $\{(v_t, w_t)\}_1^\infty$ در فرض (۱) صدق بکند، شرایطی فراهم خواهد شد(۴۷) که تحت آن شرایط با تخمین معادله (۲) به روش OLS و با افزایش T به سمت بی‌نهایت نتایج زیر را به دست خواهیم آورد(۴۸):

$$\hat{\beta} \Rightarrow \left(\frac{\sigma_v}{\sigma_w}\right) \qquad (4)$$

$$T^{-1/2}\hat{\alpha} \Rightarrow \sigma_v \{\int_0^1 v(t)dt - \xi\int_0^1 w(t)dt\} \qquad (5)$$

که در آن علامت ⇐ اشاره به مفهوم *همگرایی ضعیف*[۴۹] معیارهای احتمالی مربوط دارد. همچنین v(t) و w(t) *فرایندهای وینر (یا فرایندهای حرکت براونی)*[۵۰] مستقل در فضای $C[0,1]$[۵۱] هستند و $\sigma_v^2$ و $\sigma_w^2$ به صورت زیر تعریف می‌شود:

$$\sigma_v^2 = \lim_{T\to\infty} T^{-1}E\left(P_T^2\right)\ ,\ \sigma_w^2 = \lim_{T\to\infty} T^{-1}E\left(\phi_T^2\right) \qquad (6)$$

و

$$P_t = \sum_{j=1}^{t} v_j\ \ ,\ \ \phi_t = \sum_{j=1}^{t} w_j$$

روابط (۴) و (۵) نشان می‌دهند که برخلاف نتایج مرسوم در رگرسیون های معمولی، وقتی که T به سمت بی‌نهایت میل می‌کند، ضرایب $\hat{\alpha}$ و $\hat{\beta}$ به سمت مقدار ثابتی *همگرا در احتمال*[۵۲] نمی‌شوند. چنانچه T به سمت بی‌نهایت میل کند، $\hat{\beta}$ یک توزیع حدی *زوال‌ناپذیر*[۵۳] دارد و توزیع $\hat{\alpha}$ واگرا خواهد شد. بنابراین نااطمینانی درباره رگرسیون (۲) که از ماهیت جعلی آن ناشی می‌شود، به طور مجانبی در این‌گونه توزیع های حدی ادامه خواهد یافت.

برای $t_{\hat{\beta}}$ و $t_{\hat{\alpha}}$ نیز روابط زیر را خواهیم داشت:

$$T^{-1/2}t_{\hat{\beta}} \Rightarrow \mu/\nu^{1/2} \qquad (7)$$

که در آن:

$$\mu = \int_0^1 v(t)w(t)dt - \int_0^1 v(t)\,dt\int_0^1 w(t)dt$$

$$\nu = \{\int_0^1 v(t)^2 dt - (\int_0^1 v(t)^2 dt)^2\}\{\int_0^1 w(t)^2 dt - (\int_0^1 w(t)dt)^2\}$$
$$-\{\int_0^1 v(t)w(t)dt - \int_0^1 v(t)dt\int_0^1 w(t)dt\}^2$$

$$T^{-1/2}t_{\hat{\alpha}} \Rightarrow \frac{\{\int_0^1 v(t)\,dt - \int_0^1 w(t)\,dt\}\{\int_0^1 w(t)^2\,dt - (\int_0^1 w(t)\,dt)^2\}}{\left[\nu\int_0^1 w(t)^2 dt\right]^{1/2}} \qquad (8)$$

روابط (۷) و (۸) نیز نشان می‌دهند که با افزایش T به سمت بی‌نهایت توزیعهای $t_\alpha$ و $t_\beta$ واگرا می‌شوند، به گونه‌ای که مقادیرِ بحرانیِ به طور مجانبی صحیح برای آزمون های معناداری مرسوم ضرایب

وجود ندارد و با افزایش حجم نمونه *نرخ رد کردن*[۵۴] این آزمون ها افزایش می‌یابد. ذکر این نکته ضروری است که $S_{\alpha}^{۲}$ و $S_{\beta}^{۲}$ در حالت فقدان اثرات جعلی با نرخ $T^{-۱}$ همگرا می‌شوند.

برای آزمونهای تشخیصی $R^{۲}$، DW و آماره باکس-پیرس ($Q_k$) نیز خواهیم داشت:

$$R^{۲} \Rightarrow \frac{\xi^{۲}\{\int_{۰}^{۱} w(t)^{۲} dt - (\int_{۰}^{۱} w(t)dt)^{۲}\}}{\int_{۰}^{۱} \nu(t)^{۲} dt - (\int_{۰}^{۱} \nu(t)dt)^{۲}} \qquad (۹)$$

$$DW \rightarrow_{p} O \qquad (۱۰)$$

$$T^{-۱}Q_k = \sum_{S=۱}^{k} r_s^{۲} \rightarrow_{p} k \qquad (۱۱)$$

که در آن $r_s$ ضریب همبستگی پیاپی جملات پسماند رگرسیون بوده و به صورت زیر تعریف می‌گردد:

$$r_s = \sum_{s=۱}^{T} \hat{u}_t \hat{u}_{t-S} \bigg/ \sum_{t=۱}^{T} \hat{u}_t^{۲} \qquad (۱۲)$$

روابط (۹) تا (۱۱) نشان می‌دهند که چنانچه T به سمت بی‌نهایت میل کند، آماره DW به سمت صفر همگرا می‌شود، در حالی که ضریب تعیین $R^{۲}$ *توزیع حدی زوال‌ناپذیری* دارد و توزیع آماره باکس-پیرس ($Q_k$) نیز واگرا می‌شود. بنابراین در رگرسیون های جعلی از قبیل معادله رگرسیونی (۲) که داده‌های آن با روابط (۳) تولید شده‌اند، انتظار می‌رود که مقادیر آماره d دوربین واتسون بسیار پایین باشد و ضریب تعیین رگرسیون ($R^{۲}$) مقدار زیادی داشته باشد. فیلیپس (۱۹۸۶) خاطرنشان می‌کند که کلیه این نتایج متفاوت با رگرسیون‌های با فرایندهای مانا می‌باشد. ضریب تعیین رگرسیون در حالت نبود رابطه جعلی به سمت صفر همگرا می‌شود.

نتایج ارائه شده توسط فیلیپس (۱۹۸۶) که تحلیل ریاضی نتایج شبیه‌سازی مونت‌کارلوی گرنجر و نیوبولد (۱۹۷۴) است، با این فرض به دست آمده که $X_t$ و $Y_t$ در معادله رگرسیونی شماره (۲)، *گام تصادفی بدون جمله رانش* می‌باشند. اما انتورف (۱۹۷۷) [۱۷] با اشاره به این نکته که اغلب سریهای زمانی اقتصادی با جمله رانش ظاهر می‌شوند، نتایج زیر را در حالتی به دست آورده است که $X_t$ و $Y_t$ متغیرهای گام تصادفی با جمله رانش غیرصفر می‌باشند[۵۵]:

$$T^{۱/۲}\left(\hat{\beta} - \frac{\gamma_y}{\gamma_x}\right) \Rightarrow N\left\{۰, \frac{۶}{۵}\left(\frac{\sigma_y}{\gamma_x}\right)^{۲}\left[\left(\frac{\gamma_y}{\gamma_y}\right)^{۲} + \left(\frac{\sigma_x}{\gamma_x}\right)^{۲}\right]\right\} \qquad (۱۳)$$

$$T^{-۱/۲}\hat{\alpha} \Rightarrow N\left\{۰, \frac{۲}{۱۵}\left[\sigma_y^{۲} + \sigma_x^{۲}\left(\frac{\gamma_y}{\gamma_x}\right)^{۲}\right]\right\} \qquad (۱۴)$$

که در آن "$\Rightarrow$" به معنی *همگرایی در توزیع*(۵۶) می‌باشد. $\gamma_x$ و $\gamma_y$ به ترتیب جمله رانش و متغیرهای X، Y و $\sigma_x$ و $\sigma_y$ نیز مقادیر انحراف معیار این متغیرهاست.

روابط شماره (۱۳) و (۱۴) نشان می‌دهند که تخمین $\hat{\beta}$ در معادله (۲) با متغیرهـای گـام تـصادفی بـا جمله رانش غیرصفر (یعنی با $\gamma_y \neq ۰, \gamma_x \neq ۰$) تخمین های سازگاری از نسبت جملات رانش رگرسند و رگرسور (یعنی $\frac{\gamma_y}{\gamma_x}$) را ارائه می‌دهد و این نتیجه مخالف نتیجه بـه دسـت آمـده از رگرسـیون هـای جعلی بدون جمله رانش (نتیجه فیلیپس (۱۹۸۶)) می‌باشد. اما نتیجه تخمین مقدار عـرض از مبـدأ $\hat{\alpha}$ همانند حالت رگرسیون های با گام تصادفی بدون جمله رانش می‌باشد.

برای مقادیر آماره $t_\alpha$ و $t_\beta$ و آماره دوربین-واتسون (DW) و ضریب تعیین ($R^2$) رگرسیون هـای دارای متغیرهای گام تصادفی با جمله رانش غیرصفر نیز خواهیم داشت:

$$t_\beta = O_p(T) \tag{۱۵}$$

$$t_\alpha = O_p\left(T^{۱/۲}\right) \tag{۱۶}$$

$$\left(۱ - R^۲\right) = O_p\left(T^{-۱}\right) \tag{۱۷}$$

$$DW = O_p\left(T^{-۱}\right) \tag{۱۸}$$

روابط (۱۶) و (۱۸) نشان می‌دهند که چنانچه T به سمت بی‌نهایت میل کند، مقادیر آماره DW و $t_\alpha$ با همان سرعت نتایج متغیرهای گام تصادفی بدون جمله رانش به سمت صـفر نزدیـک مـی‌گردنـد. رابطه (۱۵) نیز نشان می‌دهند که نسبت t ضریب $\beta$ یعنی $t_\beta$ در حـضور متغیرهـای گـام تـصادفی بـا جملات رانش به طور خطی رشد می‌کند.

با استفاده از رابطه (۱۷) می‌توان ملاحظه کرد که ضریب تعیین رگرسیون بـا سـرعت T بـه سـمت کمیت واحد نزدیک می‌گردد. انتورف (۱۹۹۷) [۱۷] اشاره می‌کند که ضریب تعیـین، آمـاره تشخیـصی مفیدی برای تشخیص رگرسیون‌های جعلی با جمله رانش از رگرسیون های جعلی بـدون جملـه رانـش می‌باشد (انتورف، ص ۲۹۰). همچنانکه پیشتر اشاره شد، در رگرسیون های جعلی بـدون جملـه رانـش، ضریب تعیین رگرسیون ($R^2$) به سمت یک متغیر تصادفی زوال‌ناپذیر همگرا می‌شود.

فیلیپس (۱۹۸۶) [۴۸] علاوه بر تشریح پدیده رگرسیون جعلی بـا متغیرهـای گـام تـصادفی بـدون جملـه رانـش در قالـب یـک الگـوی رگرسـیون خطـی سـاده (معادلـه شـماره ۲)، ایـن مـسئله را بـه رگرسیون‌های چندمتغیره (مرکب) با فرایندهای جمع بسته از مرتبـه (۱) نیـز بـسط داده اسـت. بـرای تشریح این پدیده در رگرسیون های چندمتغیره (مرکب)، الگوی زیر را در نظر بگیرید:

$$y_t = \hat{\alpha} + x_t' \hat{\beta} + \hat{u}_t \quad , \quad t = ۱, ۲, \ldots, T \tag{۱۹}$$

که در آن $y_t$ و $x'_t$ برداری از فرایندهای جمع بسته از مرتبه (۱) می‌باشند. برخلاف رگرسیون های ساده با متغیرهای گام تصادفی بدون جمله رانش، در اینجا نیازی به این فرض نیست که $x_t$ و $y_t$ مستقل از هم باشند و تنها فرض ضروری این است که بردار سری زمانی $(y_t, x'_t)$ به مفهوم انگل-گرنجر *رابطه* همجمعی نداشته باشند.

برای سادگی فرض می‌کنیم کـه $Z'_t(y_t, x'_t)$ است و $Z_t$ برداری از فرایندهای جمع بسته و دارای بعد $n = m + 1$ می‌باشد و به وسیله سازوکار زیر تولید شده است:

$$Z_t = Z_{t-1} + \xi_t \ , \ t = 1, 2, \dots \qquad (۲۰)$$

که برداری از متغیرهای گام تصادفی بدون جمله رانش می‌باشند. همچنین فرض براین است که $\xi_t = I(0)$ می‌باشد. حال اگر معادله (۱۹) را به روش OLS تخمین بزنیم و دنباله عامل ابداعی $\{\xi_t\}_1^\infty$ در فرض (۱) صدق کند، در این صورت با افزایش حجم نمونه به سمت بی‌نهایت، به نتایج زیر در رگرسیون های جعلی چندمتغیره دست خواهیم یافت:[۵۷]

$$\hat{\beta} \Rightarrow A_{22}^{-1}\, a_{21} \qquad (۲۱)$$

$$T^{-1/2}\hat{\alpha} \Rightarrow b_1 - b'_2\, A_{22}^{-1}\, a_{21} \qquad (۲۲)$$

$$T^{-1/2} t_{\beta_i} \Rightarrow \left(A_{22}^{-1} a_{21}\right)_i \Big/ \{(a_{11} - a'_{21} A_{22}^{-1}\, a_{21})^{1/2} ([A_{22}^{-1}]_{ii})^{1/2}\} \qquad (۲۳)$$

$$R^2 \Rightarrow a'_{21}\, A_{22}^{-1}\, a_{21} / a_{11} \qquad (۲۴)$$

$$T^{-1} F_\beta \Rightarrow \left(\frac{1}{m}\right) a'_{21}\, A_{22}^{-1}\, a_{21} \Big/ \left\{a_{11} - a'_{21}\, A_{22}^{-1}\, a_{21}\right\} \qquad (۲۵)$$

$$TDW \Rightarrow \eta' \Sigma_\eta \Big/ \left\{a_{11} - a'_{21}\, A_{22}^{-1}\, a_{21}\right\} \qquad (۲۶)$$

که در روابط بالا

$$\begin{array}{cccc} & 1 & m & \\ A = & \begin{pmatrix} a_{11} & a'_{21} \\ a_{21} & A_{22} \end{pmatrix} & & \begin{array}{c} 1 \\ m \end{array} \end{array} \qquad (۲۷)$$

$$= \Sigma^{1/2} \{ \int_0^1 Z(t) Z(t)' dt - \int_0^1 Z(t) dt \int_0^1 Z(t)' dt \} \Sigma^{1/2}$$

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}_{m}^{1} = \Sigma^{1/2} \int_0^1 Z(t)\, dt$$

$$\eta' = \left(1, - a'_{21}\, A_{22}^{-1}\right)$$

$$\sum = \lim_{T \to \infty} T^{-1} E\left(S_T S_T'\right) \;,\; S_T = \sum_{j=1}^{t} \xi_j$$

$$\sum_{\xi} = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} E\left(\xi_t \, \xi_t'\right)$$

و $Z(t)$ برداری از *فرایندهای وینر* می‌باشد.

نتایج به دست آمده توسط فیلیپس (۱۹۸۶) برای رگرسیون های چند متغیره را می‌توان به صورت زیر خلاصه کرد:

۱- براساس معادلات شماره (۲۱) و (۲۲) می‌توان ملاحظه کرد که همانند رگرسیون خطی ساده، ضرایب $\hat{\alpha}$ و $\hat{\beta}$ با افزایش حجم نمونه $(T \to \infty)$ *همگرا* در حد به یک مقدار ثابتی نمی‌شوند. با افزایش حجم نمونه توزیع $\hat{\alpha}$ واگرا می‌شود در حالی که $\hat{\beta}$ یک توزیع حدی زوال‌ناپذیری پیدا می‌کند.

۲- $t_{\beta_i}$ نشان دهنده آماره t مرسوم برای ارزیابی معناداری $\hat{\beta}_i$ و $F_\beta$ آماره F مرسوم در ارزیابی معناداری $\hat{\beta}$ در معادله (۱۹) می‌باشند. روابط (۲۳) و (۲۵) نشان می‌دهند که با افزایش حجم نمونه به سمت بی‌نهایت $(T \to \infty)$ توزیع‌های $t_{\beta_i}$ و $F_\beta$ واگرا می‌شوند. بنابراین مقادیر بحرانی به طور مجانبی صحیح برای این آماره‌ها وجود نخواهد داشت و استفاده از مقادیر بحرانی مرسوم به نتایج جعلی منجر خواهد شد. نرخ واگرایی توزیع $F_\beta$ معادل $O(T)$ است که بیشتر از نرخ واگرایی $O(T^{1/2})$ برای آماره t انفرادی (و در حالت رگرسیون ساده $t_\alpha$ و $t_\beta$) می‌باشد. بنابراین، در رگرسیون های چندمتغیره نرخ رد فرضیه‌های صفر با آماره F به طور قابل ملاحظه‌ای افزایش خواهد یافت. این نتایج در آزمایشات شبیه‌سازی مونت‌کارلوی گرنجر و نیوبولد (۱۹۷۴) نیز تأیید شده است. گرنجر و نیوبولد (۱۹۷۴) نشان دادند که هرگاه یک رگرسور در معادله و ۵۰ مشاهده (T=۵۰) داشته باشیم نرخ رد ۷۶ درصد است در حالی که با پنج رگرسور این نرخ به ۹۶٪ افزایش می‌یابد.

۳- روابط (۲۴) و (۲۶) نشان می‌دهند که با افزایش حجم نمونه به سمت بی‌نهایت $(T \to \infty)$ آماره d دوربین-واتسون (DW) در حد همگرا به صفر شده و ضریب تعیین رگرسیون $(R^2)$ توزیع حدی زوال‌ناپذیری خواهد داشت.

در کلیه مطالعات قبلی، خصوصاً مطالعه فیلیپس (۱۹۸۶) و گرنجر و نیوبولد (۱۹۷۴)، مسئله رگرسیون جعلی در حالتی بررسی شد که بین رگرسورها و رگرسند با فرایند $I(1)$ یک رابطه همجمعی به مفهوم انگل-گرنجر وجود نداشته باشد و همچنانکه پیشتر اشاره شد، این یک فرض ضروری در مطالعه فیلیپس (۱۹۸۶) بود. اما چوی (۱۹۹۴) [۱۰] مسئله رگرسیون جعلی را با رگرسورهای دارای رابطه همجمعی تحلیل کرده است. وی اشاره می‌کند که در برخی از کاربردها می‌توان انتظار داشت که یک رابطه همجمعی بین متغیرها وجود دارد. به طور مثال، وجود رابطه همجمعی در رگرسیون هایی که

در آنها یک متغیر قیمتی (مانند نرخ ارز، نرخ بهره و ...) بر متغیرهای قیمتی دیگر رگرس می‌شود، دور از انتظار نیست (چوی، ۱۹۹۴، ص ۳۱۴)، اگرچه کوربی و اولیارس (۱۹۸۹) [۱۲] چنین نتیجه‌ای را به دست نیاورده‌اند .

بار دیگر الگوی رگرسیونی (۱۹) و سازوکار تولید کننده $Z_t$ در معادله (۲۰) را به صورت زیر در نظر بگیرید:

$$y_t = \hat{\alpha} + x_t' \hat{\beta} + \hat{u}_t \ , \ t = 1,2,\ldots,T \qquad (19)$$

$$Z_t = Z_{t-1} + \xi_t \ , \ t = 1,2,\ldots \quad \text{و} \quad Z_t' = (y_t, x_t') \qquad (20)$$

فرض بر این بود که $Z'$ برداری از فرایندهای $I(1)$ با ابعاد $m+1$ $(m \geq 2)$ است و فرایند $I(0)$ می‌باشد.

فیلیپس (۱۹۸۶) و گرنجر و نیوبولد (۱۹۷۴) فرض کردند که دنباله خطای تصادفی $\{u_t\}$ $I(1)$ است و رابطه همجمعی بین متغیرها وجود ندارد. اما چوی (۱۹۹۴) فرض می‌کند که $u_t = I(1)$ است و p (که در آن $1 \leq p \leq m-1$ است) بردار همجمعی مستقل خطی برای xها $(\gamma_1,\ldots\ldots,\gamma_p)$ وجود دارد. نتایج حاصل از مطالعه چوی (۱۹۹۴) را می‌توان به صورت زیر خلاصه کرد.[58]

۱- $\hat{\beta}_{OLS}$ توزیع حدی زوال‌ناپذیر داشته و توزیع $\hat{\alpha}_{OLS}$ واگرا می‌باشد.

۲- ضریب تعیین رگرسیون $(R^2)$ همگرا در توزیع به سمت یک متغیر تصادفی *خوش تعریف*[59] خواهد شد.

۳- آماره دوربین-واتسون همگرا در احتمال به سمت صفر می‌شود.

۴- توزیع آماره F واگرا می‌گردد ولی توزیع آماره t ضرایب به سمت متغیرهای خوش‌تعریف، همگرا می‌شود.

اما با مقایسه نتایج حاصل از مطالعات فیلیپس (۱۹۸۶) و گرنجر و نیوبولد (۱۹۷۴) و چوی (۱۹۹۴) می‌توان به نکات زیر اشاره کرد:

۱- مقدار ضریب تعیین $(R^2)$ در رگرسیون های دارای رابطه همجمعی (مطالعه چوی) معتدل‌تر از رگرسیون‌های مورد نظر در مطالعه فیلیپس (نبود رابطه همجمعی) است اما مقدار اماره دوربین-واتسون درچنین رگرسیون‌هایی کمتر از رگرسیون‌های بدون وجود رابطه همجمعی می‌باشد.

۲- فیلیپس (۱۹۸۶) نشان داد که نسبت‌های t ضرایب متغیرها واگرا می‌شوند. اما در مطالعه چوی (۱۹۹۴) یعنی در رگرسیون های دارای رابطه همجمعی نسبت‌های t واگرا نمی‌شوند و به سمت متغیرهای خوش‌تعریف همگرا می‌شوند.

۳- چوی (۱۹۹۴) نشان داد که در صورت نبود رابطه همجمعی، توزیع‌های حدی $\hat{\beta}_{OLS}$ و $\hat{\alpha}_{OLS}$ به نتایج حاصل از مطالعه فیلیپس (۱۹۸۶) تبدیل می‌شوند.

۴- متفاوت بودن توزیع مجانبی $\hat{\beta}_{OLS}$ در مطالعات چوی (۱۹۹۴) و فیلیپس (۱۹۸۶) در رگرسیونهای با رگرسورهای دارای رابطه همجمعی (مطالعه چوی) و با رگرسورهای بدون رابطه همجمعی (مطالعه فیلیپس) یک نتیجه مهمی برای آزمونهای همجمعی براساس جملات پسماند رگرسیون همجمعی دارد. زیرا توزیع فرضیه صفر در این آزمونها به توزیع مجانبی $\hat{\beta}_{OLS}$ بستگی دارد.

### ۲-۱-۲. متغیرهای $I(۱)$ و $I(۲)$

در قسمت قبل، مسئله رگرسیون جعلی با متغیرهای I(1) تشریح گردید. اما ممکن است مرتبه جمع بستگی متغیرها یکسان نبوده و ترکیبی از متغیرهای $I(۰)$، $I(۱)$ و $I(۲)$ در الگوی رگرسیونی وجود داشته باشد. چنگ و فیلیپس (۱۹۹۵) [۹] بیان می‌کنند که برخی از متغیرهای عمده کلان اقتصادی از جمله حجم پول و سطح قیمت‌ها در برخی از مطالعات فرایند $I(۲)$ بوده و این متغیرها در اکثر الگوهای کلان اقتصادی کاربرد دارند.[۶۰] بنابراین این سوال پیش می‌آید که آیا حضور متغیرهای با مرتبه جمع بستگی متفاوت، امکان وجود پدیده رگرسیون جعلی را منتفی می‌سازد یا چنین پدیده‌ای با این ترکیب متفاوت نیز وجود دارد.؟

در این قسمت الگوهای رگرسیونی دارای متغیرهای $I(۱)$ و $I(۲)$ و در بخش بعدی الگوهای رگرسیونی با متغیرهای $I(۰)$ و $I(۱)$ در نظر گرفته شده و پدیده رگرسیون جعلی و نتایج آن بررسی خواهد شد. نحوه تخمین و استنباط صحیح در رگرسیون‌های دارای رگرسورهای جمع بسته از مرتبه‌های متفاوت نیز خارج از حوصله مقاله حاضر است و به مجال دیگری نیاز دارد.

فرض کنید که دنباله $\{y_t\}_۱^\infty$ یک سری زمانی تک‌بعدی و $\{x_t\}_۱^\infty$ یک سری m- بعدی است و به صورت زیر تولید شده باشند.

$$y_t = \gamma' c_t + y_t^{\cdot} \qquad (۳۰)$$

$$x_t = \left(c_t', x_{۱t}', x_{۲t}'\right)' \qquad (۳۱)$$

که در آن:

$$x_{۱t} = \gamma_۱' c_t + x_{۱t}^{\cdot} \ , \ \Delta x_{۱t}^{\cdot} = \varepsilon_{۱t} \qquad (۳۲)$$

$$x_{۲t} = \gamma_۲' c_t + x_{۲t}^{\cdot} \ , \ \Delta x_{۲t}^{\cdot} = \varepsilon_{۲t} \qquad (۳۳)$$

و $\{c_t\}_۱^\infty$ دنباله معین (غیرتصادفی) $m_0$- بعدی است و ممکن است در اکثر کاربردها به طور ساده مرکب از یک مقدار ثابت، روند زمانی و احتمالاً روند چندجمله‌ای در نظر گرفته شود. $x_{۱t}^{\cdot}$ , $x_{۲t}^{\cdot}$ نیز به ترتیب فرایندهای تصادفی $m_۱$ و $m_۲ -$ بعدی جمع بسته از مرتبه (۱) و (۲) و $\gamma_۱, \gamma_۲$ ضرایب مولفه‌های معین (غیرتصادفی) تعریف شده در $C_t$ می‌باشند.

همچنین فرض کنید که $y_t^{\cdot}$ فرایند جمع بسته از مرتبه (۲) بوده و رابطه زیـر را بـا $x_{1t}^{\cdot}$ و $x_{2t}^{\cdot}$ داشـته باشد:

$$y_t^{\cdot} - \beta_1' x_{1t}^{\cdot} - \beta_2' x_{2t}^{\cdot} = u_t \tag{۳۴}$$

در این صورت با در نظر گرفتن رابطه (۳۰) تا (۳۴) می‌توان نوشت:

$$y_t = (\gamma_0' - \beta_1' \gamma_1' - \beta_2' \gamma_2')c_t + \beta_1' x_{1t} + \beta_2' x_{2t} + u_t = \beta' x_t + u_t \tag{۳۵}$$

حال فرض کنید که الگوی رگرسیونی (۳۵) به وسیله روش حداقل مربعات معمولی تخمین زده شود. در این صورت خواهیم داشت:

$$y_t = \hat{\beta}_0' c_t + \hat{\beta}_1' x_{1t} + \hat{\beta}_2' x_{2t} + \hat{u}_t = \hat{\beta}' x_t + \hat{u}_t \tag{۳۶}$$

که در آن

$$\hat{\beta}_{OLS} = \left(\sum_{t=1}^{n} x_t x_t'\right)^{-1} \left(\sum_{t=1}^{n} x_t y_t\right) \tag{۳۷}$$

هالدراپ (۱۹۹۴) [ ۲۴ ] نشان می‌دهد که اگر فرایند تولیدکننـده داده‌هـا (DGP) بـه صـورت روابـط (۳۰) تا (۳۵) باشد و الگوی رگرسیونی (۳۶) به روش OLS تخمین زده شود و چنانچه حجم نمونه بـه سمت بی‌نهایت میل کند $(n \to \infty)$ در این صورت برای ضرایب الگوی رگرسیونی و برخی آماره‌هـای آزمون تشخیصی می‌توان موارد زیر را ذکر کرد:[۶۱]

*۱- ضرایب متغیرهای $I(1)$ (یعنی $\hat{\beta}_1$ در الگوی رگرسیونی (۳۶)):*

- هرگاه $d = 2$ باشد، واگرا از مرتبه $O_p(n)$ خواهد شد و
- هرگاه $d = 1$ باشد، زوال‌ناپذیر می‌گردد (d مرتبه جمع بستگی $y_t$ مشروط بر $x_t$ است).

*۲- ضرایب متغیرهای $I(2)$ (یعنی $\hat{\beta}_2$ در الگوی رگرسیونی (۳۶)):*

- هرگاه الگوی شرطی $I(2)$ باشد $(d = 2)$، زوال‌ناپذیر خواهد شد و
- هرگاه $d = 1$ باشد با *نرخ ابرسازگار*[۶۲] $O_P(n-1)$، تخمین های سازگاری خواهد داشت.

*۳- آماره $F$ مربوط به ضرایب متغیرهای $I(2)$، ناسازگار می‌باشد.*

*۴- هرگاه الگوی شرطی دارای متغیرهای $I(1)$ و $I(2)$ باشد، آماره $F$ هر فرضـیه‌ای همیـشه بـا نـرخ $Op(n)$ به بی‌نهایت تمایل خواهد داشت، اگرچه فرضیه واقعاً درست باشد.*

*۵- رفتار آماره‌های تشخیصی به شدت به مرتبه جمع بستگی الگوی شرطی بستگی دارد.*

- اگر الگوی شرطی مانا بوده و یا حتی $I(1)$ باشد، ضریب همبستگی چندگانه به یـک متمایـل خواهـد شد.
- هرگاه هیچ رابطه همجمعی بین سریها وجود نداشته باشـد (یعنـی $d = 2$ باشـد)[۶۳] ضـریب تعیـین رگرسیون $(R^2)$ توزیع حدی زوال‌ناپذیری خواهد داشت.

- صرفنظر از اینکه $d=1$ یا $d=2$ باشد، آماره دوربین-واتسون (D.W.) با نرخ $O_p(n)$ به سرعت به صفر متمایل خواهد شد.

## ۲-۲. رگرسیون جعلی با متغیرهای نامانا و مانا

مسئله رگرسیون جعلی توسط گرنجر و نیوبولد (۱۹۷۴) و فیلیپس(۱۹۸۹) با متغیرهای $I(1)$ مطرح شده و سپس هالدراپ (۱۹۹۴) آن را با متغیرهای با مرتبه جمع بستگی متفاوت تحلیل کرد. وجه مشخصه این مطالعات در این است که کلیه متغیرهای مستقل و وابسته نامانا می‌باشند، اگرچه مرتبه جمع بستگی آنها متفاوت است. حال سوال این است که چنانچه متغیرهای مانایی وارد چنین الگوهای رگرسیونی گردد، توزیع حدی پارامتر تخمین‌زننده متغیرهای مانا چگونه خواهد شد و آیا مسئله رگرسیون جعلی در چنین مواردی نیز کاربرد دارد یا خیر. در این قسمت به این گونه سوالات پاسخ داده خواهد شد.

الگوی رگرسیونی زیر را در نظر بگیرید:

$$y_t = \hat{a}_1 x_{1t} + \hat{a}_2 x_{2t} + \hat{u}_t \qquad (۳۸)$$

$$y_t = \overline{\mu} + \overline{a}_1 x_{1t} + \overline{a}_2 x_{2t} + \overline{u}_t \qquad (۳۹)$$

$$y_t = \widetilde{\mu} + \widetilde{\theta} t + \widetilde{\alpha}_1 x_{1t} + \widetilde{\alpha}_2 x_{2t} + \widetilde{u}_t \qquad (۴۰)$$

در الگوهای بالا فرض بر این است که $x_{2t}, y_t$ متغیرهای نامانای ریشه واحد $I(1)$ بوده و $x_{1t}$ متغیر مانا $[I(0)]$ است. دنباله $\{y_t\}_1^\infty$ سری زمانی چندگانه n- بعدی است و $\alpha_2, \alpha_1$ به ترتیب ماتریس ضرائب $n \times m_1$ و $n \times m_2$ می‌باشند.

پارک و فیلیپس (۱۹۸۹) [۴۷] نشان داده‌اند که[۶۴] برآوردگرهای حداقل مربعات در الگوهای رگرسیونی (۳۸) تا (۴۰) به طور مجانبی معادل با ضرایب رگرسیونی حاصل از رگرسیون $y_{2t}$ روی $x_{1t}$ است که در آن $y_{2t} = y_t - \alpha_2 x_{2t}$ است.[۶۵] بنابراین نظریه متعارف رگرسیون برای برآوردگرهای حداقل مربعات معمولی متغیرهای مانا $(\hat{a}_1)$ قابل کاربرد می‌باشد. این محققان با در نظر گرفتن این مطلب که $x_{2t}, y_t$ رابطه همجمعی با هم دارند، به این نتیجه دست یافتند که کلیه برآوردگرهای حداقل مربعات معمولی برآوردگرهای سازگاری هستند، حتی اگر $\sum_{10} \neq 0$ بوده و رگرسورهای مانا با خطاهای رگرسیون به طور همزمانی همبسته باشند. ($\sum$ برآوردگر حداقل مربعات از ماتریس کوواریانس است). همچنین اضافه کردن متغیرهای جمع بسته دیگر یا متغیر روند به این رگرسیون ها تأثیری بر توزیع حدی پارامترهای $\hat{a}_1$ ندارد و نرخ همگرایی این برآوردگرها دقیقاً معادل با همان نرخی است که در رگرسیون های با رگرسورهای جمع بسته از مرتبه یکسان وجود دارد.

هاسلر (۱۹۹۶) [۲۶] با در نظر گرفتن این مطلب که $x_{2t}, y_t$ رابطه همجمعی با هم ندارند، مسئله رگرسیون‌های جعلی با متغیرهای $I(1)$ بدون جمله رانش را هنگامی تحلیل کرده است که رگرسورهای اضافی $I(0)$ به آن اضافه می‌شود. وی به نتایج زیر دست یافته است:

۱- $\hat{a}_1$ (ضرائب متغیرهای مانا در الگو) به سمت صفر همگرا نمی‌شود و توزیع حدی زوال‌ناپذیری دارد.

۲- آماره t ضرائب متغیرها مانا $(\hat{a}_1)$ همانند نتیجه چوی (۱۹۹۴) [۱۰] واگرا نمی‌شود بلکه به سمت یک متغیر خوش تعریف همگرا می‌شود. آماره F نیز واگرا نخواهد شد.

۳- آماره t ضرائب متغیرهای نامانا $(\hat{a}_2)$ و توزیع مجانبی آنها همانند حالتی است که در آن کلیه رگرسورها، متغیرهای $I(1)$ می‌باشند. این نتیجه همانند نتیجه به دست آمده در مطالعه فیلیپس (۱۹۸۶) می‌باشد.

۴- کوواریانس‌های متغیرهای مانا بر توزیع آماره دوربین- واتسون تأثیر می‌گذارد و متفاوت از نتایج به دست آمده توسط فیلیپس (۱۹۸۶) می‌باشد.

## ۳.جمع بندی

در این مقاله، سیر تکامل تاریخی و مفهوم رگرسیون جعلی *(بی معنی، دروغین یا کاذب)* و نتایج یافته‌های برخی از مطالعات مهم در ارتباط با داده های سری زمانی ارائه شده است.

مفهوم *رگرسیون جعلی(بی معنی، دروغین یا کاذب)* یا *همبستگی‌های بی معنی* در متون اقتصادسنجی توسط گرنجر و نیوبولد ( ۱۹۷۴ ) معرفی شد. این مفهوم اشاره به وضعیتی دارد که در آن نتایج حاصل از تخمین یک الگوی اقتصادسنجی غیرواقعی است . در تخمین هر الگوی اقتصادسنجی، معمولاً نوع داده ها (سری زمانی، مقطعی، تلفیقی، فضایی)، نوع فرایند (جمع بسته، جمع بسته کسری، گه گن بایر، و ...)، نوع روابط (خطی، غیر خطی)، شکل داده ها (مقادیر مطلق، لگاریتمی، مقادیر نسبی یا شاخص و ... ) و روش تخمین روشن است. مفهوم معرفی شده توسط گرنجر و نیوبولد در الگوهای سری زمانی خطی است که در تخمین آن از مقادیر مطلق متغیرها استفاده شده و همه متغیر ها از درجه جمع بستگی یکسان $I(1)$ برخوردار بوده‌اند و الگو به روش حداقل مربعات معمولی تخمین زده شده است.نتایج حاصل از شبیه‌سازی مونت‌کارلوی گرنجر و نیوبولد بعدها توسط سایر محققین مورد آزمون قرار گرفته و بسط داده شد. نلسون و کنگ (۱۹۸۳ و ۱۹۸۱) نیز همانند گرنجر و نیوبولد با استفاده از شبیه‌سازی مونت‌کارلو نشان دادند که پدیده رگرسیون جعلی زمانی رخ می‌دهد که رگرسیون شامل *متغیرهای نامانای ریشه واحد* باشد. فیلیپس (۱۹۸۶) برای اولین بار نتایج گرنجر و نیوبولد (۱۹۷۴) را به صورت تحلیلی و به شکل ریاضی ارائه داد.

در کلیه مطالعات بالا، متغیرهای نامانای در نظر گرفته شده در الگوهای رگرسیونی از مرتبه (۱) می‌باشند. مارمول (۱۹۹۵)، نتایج فیلیپس (۱۹۸۶) را به مواردی با فرایندهای جمع بسته از درجات

بالاتر از (۱) و در حالت کلی از درجه dام (یعنی فرایندهای I(d)) تعمـیم داد کـه در آن، d یـک عـدد صحیح می‌باشد.

مسئله رگرسیون جعلی همچنین ممکن است در موقعیت‌هایی رخ دهد که فرایندهای موجود در آن مرتبه‌های جمع بستگی متفاوتی داشته باشند. به این مـسئله در مطالعـات هالـدراپ (۱۹۹۴)، مـارمول (۱۹۹۶) و هاسلر (۱۹۹۶) توجه شده است. هالدراپ (۱۹۹۴) این مسئله را در حالتی بررسی کرد که در آن متغیرهای مستقل $I(۱)$ و $I(۲)$ بودند. هاسلر (۱۹۹۶) و مارمول (۱۹۹۶) نیـز مـوقعیتی را تحلیـل کردند که در آن متغیرهای مستقل $I(۱)$ و $I(۰)$ می‌باشند. نتایج یافته‌های برخی از مطالعات مهـم در مورد داده‌های سری زمانی با الگوهای دارای متغیرهای $I(۱)$ و $I(۲)$ و همچنین $I(۱)$ و $I(۰)$ نـشان داد که در چنین شرایطی نیز احتمال وقوع رگرسیون جعلی وجود دارد. بنابر این پیشنهاد شده است که قبل از اقدام به تخمین هر الگویی، خواص سری‌های زمانی مورد استفاده به طور دقیق بررسی شود.

نتایج یافته‌های برخی از مطالعات مهم دیگر مـورد رگرسـیون جعلـی بـا داده هـای تلفیقـی، داده هـای فضائی، داده‌های لگاریتمی، داده‌های شاخص یا نسبتی و فرایندهای جمـع بـسته کـسری و فراینـدهای مانای گه گن بایر در قسمت دوم مقاله حاضر ارائه خواهد شد.

**یادداشت‌ها:**

1. Spurious Regression

در کتاب واژه‌ها و اصطلاحات آماری (پژوهشکده آمار) این اصطلاح به رگرسیون دروغین ترجمه شده است.

2. Nonsense Correlations
3. Yule, G. U. (1926), why do we sometimes get nonsense Correlations between time series? A study in sampling and the nature of time series, Journal of Royal Statistical Society, 89, 1-69
4. Unit root nonstationary variables
5. Integrated processes
6. Orders of Integration
7. Cointegration

در ادبیات سری‌های زمانی برای مفهوم cointegrotion اصطلاحاتی از قبیل، همگرایی، همگرایی یکسان، هم‌انباشتگی، هم تجمعی و همگرایی بلندمدت انتخاب می‌شود. پژوهشکده آمار ایران در کتاب واژه‌ها و اصطلاحات آماری کلمه هم جمعی را آورده است، اگرچه اعتقاد نویسنده براین است که این اصطلاح نیز همانند سایر اصطلاحات به کار برده شده اصطلاح مناسبی نیست ولی با هدف یکسان شدن مفاهیم ارائه شده در سایر متون، این اصطلاح پذیرفته شده است.

هرگاه متغیرهای بردار سری‌های زمانی $\left(y_t, x_t'\right)$ به تنهایی نامانا ولی یک ترکیب خطی از آنها جمع بسته از مرتبه صفر (مانا) باشد، گفته می‌شود که بردار سری‌های زمانی $\left(y_t, x'\right)$ به مفهوم انگل-گرنجر رابطه همجمعی با هم دارند.

8. M2-dimensional I (1) vector of regerssors
9. Asymptotic properties
10. Test statistics
11. Stochastic trends
12. Deterministic polynomial regressors
13. Exact finite sample properties
14. Random walk

در کتاب واژه‌ها و اصطلاحات آماری (پژوهشکده آمار) این اصطلاح به قدم زدن تصادفی ترجمه شده است.

15. Drift term
16. Independent stationary processes without any trend components
17. Ratio variables
18. Index variables
19. Fractional processes
20. Long-Memory Models
21. Nonstationary fractionally integrated processes

۲۲) برای مطالعه این الگوها و تفاضل گیری کسری به منابع زیر مراجعه فرمایید. در قسمت دوم مقاله حاضر این مفاهیم بیشتر توضیح داده خواهد شد.

**-Granger, C.W. J. (1980), Long memory relationships and the aggregation of dynamic models, Journal of Econometrics, 14, 227-238.**

- Granger, C. W. J., and J. Joyeux (1980), An Introduction to long memory Time Series models and fractional differencing, Journal of time series Analysis, 1, 15-29.
- Hosking, J. (1981), Fractional differencing, Biometrica, 68, 165-176.
- Baillie, R. T. (1996), Long memory processes and fractional integration in econometrics, Journal of Econometrics, 73, 6-59.

23. Spurious detrending
24. Long memory fractionally integrated processes.
25. Panel Data
26. Spatial Data
27. Least Square Dummy Variable estimator
28. Spurious spatial regression
29. Spatial unit root
30. Spatial cointegration

**31. Diagnostic Tests (Checking)**

مهمترین نقش آزمونهای تشخیصی در مرحله ارزیابی الگو است.

۳۲) البته فرض بر این است که داده‌های مورد استفاده در الگو با هیچ مشکلی مواجه نیست. برخی از این مشکلات عبارتند از: مفقود بودن (Missing) مشاهدات، مشاهدات دورافتاده (Outlier)، خطای اندازه‌گیری، مشاهدات موثر و ......

33. Specification

تصریح سنجی (Specimetrics) یکی از شاخه‌های مهم اقتصادسنجی کاربردی است و موضوع آن انتخاب بهترین فرم تابعی برای تخمین الگو می‌باشد. در این قسمت برای سادگی فرض شده است که تصریح خطی برای رسیدن به هدف کفایت می‌کند.

34. Estimation

هر روش تخمینی تحت شرایط خاصی قابل کاربرد می‌باشد. توجه به شرایط و زمینه‌های لازم برای استفاده از روش تخمین مورد نظر در مطالعات تجربی بسیار حیاتی است و عدم توجه به این شرایط و استفاده نامناسب از روشهای تخمین، احتمال وقوع پدیده روابط جعلی را فراهم می‌آورد.

35. Evaluation

۳۶. برای مطالعه مفاهیم و اهمیت اصطلاحات تصریح و ارزیابی به کتاب بسیار ارزشمند گرنجر مراجعه فرمایید.

**Granger, C. W. J, Empirical Modeling in Econometrics: Specification and Evaluation, Cambridge University Press, Cambridge, 1999.**

37. Stationarity

برای اصطلاح stationarity در متون موجود کلماتی از قبیل مانایی، ایستایی، ساکن بودن و پایایی استفاده شده است. در این مقاله، اصطلاح مانایی ترجیح داده شده است. این اصطلاح در کتاب واژه‌ها و اصطلاحات آماری (پژوهشکده آمار) نیز به مانایی ترجمه شده است.

آزمونهای بسیار زیادی برای تشخیص مانایی یا نامانایی یک سری زمانی وجود دارد که توضیح آنها خارج از اهداف این مقاله است .

38. Overstate

39. Unbalanced or inconsistent regressions

گرنجر (۱۳۸۱) بیان می‌کند که «یکی از ویژگیهای اساسی هر مدل آن است که باید تراز باشد» (ص ۲۹). اگر هیچ رابطه همجمعی بین متغیرها وجود نداشته باشد، الگوهای با رگرسورهای دارای مرتبه‌های مختلف جمع بستگی ناتراز خواهند شد. اما یوهانسن (Johansen, 1988) یک معنی دیگری از عبارت تراز ارائه داده است (Marmol, 1998, p. 241). جهت مطالعه بیشتر مفهوم تراز در الگوهای رگرسیونی به صفحات ۲۹ تا ۳۱ گرنجر (۱۳۸۱) و یا منبع زیر مراجعه فرمایید:

**- Johansen, S., (1988), The Mathematical structure of error correction models, Contemporary Mathematics, 80, 359-386**.

۴۰. فرض کنید فرایند ایجاد سری زمانی نامانای $y_t$ به صورت زیر تعریف گردد:

$$y_t - \mu = p\left(y_{t-1} - \mu\right) + e_t$$

که در آن $\mu$ میانگین سری زمانی y، p درجه پایداری (Degree of persistence) انحرافات $y_t$ از $\mu$ و $e_t$ جمله خطای تصادفی با میانگین صفر واریانس ثابت و معین می‌باشد. هرگاه $p = 1$ باشد، در این صورت انحرافات دائمی بوده و خواهیم داشت:

$$y_t = y_{t-1} + e_t$$

الگوی فوق به **الگوی گام تصادفی بدون جمله رانش** معروف می‌باشد. هرگاه الگوی فوق را به صورت زیر بنویسیم، در این صورت ($\beta$) جمله رانش (drift) بوده و الگوی زیر **الگوی گام تصادفی با جمله رانش** خواهد بود:

$$y_t = \beta + y_{t-1} + e_t$$

برای درک مفهوم گام تصادفی نیز به کتاب‌هایی با عنوان "تجزیه و تحلیل سری‌های زمانی" مراجعه فرمایید. توضیح خوبی از این مفهوم در منبع زیر آورده شده است.

-کرایر، جاناتان دی، تجزیه و تحلیل سری‌های زمانی، مترجم، حسینعلی نیرومند، دانشگاه فردوسی مشهد، ۱۳۷۱ (صفحات ۱۸-۱۱۳).

41. Partial sum process

42. Positive definite

43. Strong Mixing

44. Mixing Numbers
45. Near Random Walk
46. Innovations

۴۷. این شرایط در لم شماره (۱) مقاله فیلیپس (۱۹۸۶) در صفحات ۳۱۴-۳۱۶ و اثبات آن در صفحات ۳۲۶-۳۳۰ ضمیمه ریاضی A2 آورده شده است.

۴۸. برای ملاحظه نحوه استخراج این نتایج به ضمیمه ریاضی A3 مقاله فیلیپس (۱۹۸۶) در صفحات ۳۳۰-۳۳۶ مراجعه فرمایید.

49. Weak Convergence
50. Wiener (Brownian Motion) Process

برای مطالعه بیشتر این مفهوم به کتابهای مربوط مراجعه فرمایید. تعریف و تشریح بسیار خوبی از این مفهوم در ضمیمه شماره ۲ صفحات ۲۵۱-۲۴۹ کتاب زیر ارائه شده است.

**- Hatanaka, M., Time Series-Based Econometrics: unit roots and cointegration, oxford University Press, 1996.**

۵۱. یعنی فضای کلیه توابع پیوسته دارای مقادیر حقیقی در فاصله $[۰,۱]$ است.

52. Converge in probability
53. Non-degenerate
54. Rejection rate

۵۵. برای ملاحظه اثبات این روابط و نحوه استخراج آنها به ضمیمه مقاله انتورف (Entorf, 1997) صفحات ۲۹۳-۲۹۵ مراجعه فرمایید.

56. Convergence in distribution

۵۷. برای ملاحظه اثبات این روابط و نحوه استخراج آنها به ضمیمه A4 مقاله فیلیپس (۱۹۸۶) صفحات ۳۳۷-۳۳۹ مراجعه فرمایید.

۵۸. برای ملاحظه روابط ریاضی به قضیه (۱) در صفحات ۳۱۶ و ۳۱۷ مقاله چوی (Choi, 1994) مراجعه فرمایید.

59. Well-defined

۶۰. طبق بررسیهای نگارنده، هیچ مطالعه‌ای وجود ندارد که در آن مرتبه جمع بستگی متغیرهای کلان اقتصادی بیش از ۲ باشد. بنابراین منظور از رگرسیون های مرکب با مرتبه جمع بستگی متفاوت متغیرها، ترکیبی از رگرسورهای $I(۰)$ و $I(۱)$ و $I(۲)$ می‌باشد.

۶۱. برای حفظ سادگی و پرهیز از پرداختن به نکات بسیار فنی در این مقاله، روابط مربوط به ضرایب و آماره‌های آزمونهای تشخیصی همانند سایر قسمتها در این جا آورده نشده است. خواننده علاقمند می‌تواند به این روابط به صفحات ۱۵۵- ۱۶۱ و برای ملاحظه اثبات آنها به ضمیمه مقاله هالدراپ (Haldrup, 1994) صفحات ۱۷۴- ۱۸۰ مراجعه فرماید.

62. Super consistent rate

۶۳. هرگاه $d = ۱$ باشد، بنابراین $y_t, x_t \sim CI(۲,۱)$ است و طبق تعریف تغییرات جملات پسماند مرتبه پایین‌تری در احتمال نسبت به متغیرهای $I(۲)$ اولیه خواهد داشت.

۶۴. برای ملاحظه روابط ریاضی به قضیه ۱-۳ در صفحه ۱۰۲ مقاله و اثبات قضایا در ضمیمه ریاضی مقاله پارک و فیلیپس (Park and Phillips, 1989) مراجعه فرمایید.

۶۵. پارک و فیلیپس (۱۹۸۹) همچنین الگوهای (۳۸) تا (۴۰) را به مواردی بسط داده‌اند که در آن متغیرهای نامانای $I(۲)$ را نیز شامل می‌شود. به عبارت دیگر در الگوهای (۳۸) تا (۴۰)، متغیرهای $x_{1t}$ متغیرهای مانا $I(۲)$، $x_{۲t}$ و $y_t$ متغیرهای نامانای $I(۱)$ و $x_{۳t}$ متغیرهای نامانای $I(۲)$ می‌باشند.

## منابع


پژوهشکده آمار، واژه‌ها و اصطلاحات آماری (ویرایش دوم)، مرکز آمار ایران، ۱۳۸۲.

خشادوریان، ادموند (۱۳۷۸)، بررسی وجود خواص مانایی در آمارهای سری زمانی اقتصادی کشور برای دوره ۷۴-۱۳۳۸، موسسه تحقیقات پولی و بانکی، بانک مرکزی جمهوری اسلامی ایران.

خلیلیان، صادق و مریم حفار اردستانی (۱۳۷۹)، بررسی رابطه صادرات کالاهای کشاورزی و رشد بخش کشاورزی در ایران (۷۵-۱۳۵۷)، اقتصاد کشاورزی و توسعه، سال هشتم، شماره ۳۲، زمستان.

صمدی، علی حسین (۱۳۸۱)، ارزیابی تاثیر صادرات و بی‌ثباتی درآمدهای صادراتی بخشهای مختلف اقتصادی بر رشد اقتصادی این بخشها: مطالعه موردی ایران (۷۴-۱۳۴۷). اقتصاد کشاورزی و توسعه، سال دهم، شماره ۳۸، تابستان.

گرنجر، دبلیو، جی، کلادیو، مدل‌سازی تجربی در اقتصاد: تشخیص و ارزیابی، ترجمه: اسمعیل ابونوری، انتشارات دانشگاه مازندران، بابلسر، ۱۳۸۱.

Ardeni, P.G. (1989), Does the law of one price really hold for commodity prices?, American Journal of Agricultural Economics, 661-669.

Baltagi, B.H., Econometric analysis of panel data, John Wiley and Son's Publisher Co., 1995.

Chambers, M. J. (1996), Fractional integration, trend stationary and difference stationary: Evidence from some U.K. macroeconomic time series, Economics Letters, 50: 19-29.

Chang, Y., and P. C. B. Phillips (1995), Time series regression with mixtures of integrated processes, Econometric Theory, 11: 1033-1094.

Choi, I. (1994), Spurious regression and residual-based tests for cointegration when regressors are cointegrated, Journal of Econometrics, 60: 313-320.

Cohen, J., and P. Cohen, Applied multiple regression/correlation analysis for the behavioral sciences, London: LEA, 1983.

Corbae, D., and S. Ouliaris (1989), A Random walk through the Gibson paradox, Journal of Applied Economics, 4: 295-303.

Darlauf, S.N., and P. C. B Phillips (1988), Trends versus random walks in time series analysis, Econometrica, 59: 1333-1354.

De Angelis, D., S. Fachin, and G. A. Young (1997), Bootstrapping unit root tests, Applied Economics, 29: 1155-1161.

De Jung, M. Robert (2003), Logarithmic spurious regressions, Economics Letters, 81: 13-21.

Elbers, C., and G. Ridder (1982), True and spurious duration dependence: the Identifiability of the proportional Hazard model, Review of Economic studies, XLIX: 403-409.

Entorf, H. (1997), Random walks with drifts: nonsense regression and spurious fixed-effect estimation, Journal of Econometrics, 80: 287-296.

Fingleton, B. (1999), Spurious spatial regression: some Monte Carlo results with a spatial unit root and spatial cointegration, Journal of Regional science, 39/1: 1-19.

Gil-Alana, L. A. (2003), A Fractional integration analysis of the population in some OECD countries, Journal of Applied statistics, 30/10: 1147-1159.

Granger, C. W. J. (1980), Long memory relationships and the aggregation of dynamic models, Journal of Econometrics, 14: 227-2398.

Granger, C. W. J, Empirical modeling in Economics: Specification and Evaluation, Cambridge University Press, Cambridge, 1999.

Granger, C. W. J. and P. Newbold (1974), Spurious regressions in econometrics, Journal of Econometrics, 74: 111- 120.

Granger, C. W. J., N. Hyung, and Y. Jeon (2001), Spurious regressions with stationary series, Applied Economics, 33: 899-904.

Haldrup, N. (1994), The asymptotic of single-equation cointegration regression with I (1) and I (2) variables, Journal of Econometrics, 63: 153-181.

Hamilton, J. D.,(1994) Time Series analysis, Princeton University Press, Princeton.

Hassler, O. (1996), Spurious regression when stationary regressors are included, Economics Letters, 50: 23-31.

Hendry, D., Dynamic Econometrics, (chap. 4), Oxford University Press, 1995.

Kamin, S. B., and J. H. Rongers (2000), Output and real exchange rate in developing countries: An application to Mexico, Journal of Development Economics, 61: 85-109.

Kao, C. (1999), Spurious regression and residual-based tests for cointegration in panel data, Journal of Econometrics, 90: 1-44.

Kim, T., y. Lee, P. Newbold (2003), Spurious regressions with processes around trends or drifts, Discussion Paper, University of Nottingham.

31. Kim, T. H., Y.-S. Lee, and P. Newbold (2004), Spurious regressions with stationary processes around linear trends, Economics Letters, 83: 257-262.

Kronmal, R. A. (1993), Spurious correlation and the fallacy of the ratio standard revisited, Journal of the Royal Statistical Society, Series A, 156: 379-392.

LeSage, J. P., (1999) Spatial Econometrics, University of Toledo, Department of Economics.

Leybourne S. J., B. P. M. McCabe, and A. R. Tremayne (1996), Can economic time series be differenced to stationery? Journal of Business & Economic Statistics, 14/4: 435-446.

Lubian, D. (1998), Local-to-spurious regression: solution, Econometric Theory, 608-618.

Lubian, D. (1996), Spurious regression and generalized least squares: Solution, Econometric Theory, 204-209.

Marmol, F. (1995), Spurious regressions between I(d) processes, Journal of Time series Analysis, 16: 313-321.

Mormol, F. (1996), Nonsense regression between integrated processes of different orders, Oxford Bulletin of Economics and Statistics, 58: 525-536.

Marmol, F. (1998), Spurious regression theory with nonstationary fractionally integrated processes, Journal of Econometrics, 84: 233-250.

Montanes, A. (1999), Reasonable spurious regression: solution, Econometric Theory, 760-764.

Morrison, D. G., and D. J. Gluck (01970, Spurious correlations that result form “Awareness vs. Usage” type regression, Journal of Marketing Research, 7: 381-384.

Moulton, B. R. (1990), An illustration of a pitfall in estimating the effects of aggregate variables on micro units, The Review of Economics and Statistics, 334-338.

Nelson, C. R. and H. Kang (1981), Spurious periodicity in inappropriately detrended time series, Econometrica, 49: 741-751.

Nelson, C. R. and H. Kang (1983), Pitfalls in the use of time as an explanatory variable in regression, Journal of Business & Economic Statistics, 2: 73-82.

Ogaki, M, and C. – Y. Choi (2001), The Gauss- Markov theorem and spurious regression, Ohio State University, Department of Economics, Working Papers.

Park, J. Y., and P. C. B. Phillips (1988), Statistical inference in regression with integrated processes: Part 1, Econometric theory.

Park, J. W. and P. C. B. Phillips (1989), Statistical inference in regressions with integrated processes: Part 2, Econometric Theory, 5: 95-131.

Journal of Econometrics, 33: 311-345.

Phillips, P. C. B (1995), Nonstationary time series and conitegration, Journal of Applied Econometrics, 10: 87-94.

Phillips, P. C. B. (1998), New tools for understanding spurious regression, Econometrica, 66/6: 1299- 1325.

Phillips, P. C. B. (2001), Bootstrapping spurious regression, Cowles Foundation Discussion paper, N. 1330.

Phillips, P. C. B., and D. J. Hodgson (1994), Spurious regression and generalized least squares: problem Econometric Theory, 10: 967-968.

Phillips, P. C. B., and H. R. Moon (1999), Linear regression limit theory for nonstationarty panel data, Econometrica, 67/5: 1057-1111.

Roberts, J. (2000), Spurious regression problems in the determinants of health care expenditure: A comment on Hitiris (1997), Applied Economics Letters, 7: 279-283.

Robinson, P. M. (1998), Comment, Journal of Business & Economic Statistics, 16/3: 276-279.

Smith, A., F. Sowell, and S. E. Zin (1997), Fractional integration with drift: estimation in small samples, Empirical Economics, 22: 103-116.

Snyder, D. (1990), Foreign aid and domestic savings: A spurious correlation, Economic Development and Cultural Change, 175-181.

Sun, Y. (2002a), Spurious regressions with stationary Gegenbauer processes and Harmonic processes, University of California, San Diego, UCSO Working Paper, December.

Sun, Y. (2002b), A Convergent t-statistics in spurious regressions, University of California, San Diego, UCSD Discussion Paper.

Tsay, W.-J., C.- F. Chung (2000), the spurious regression of fractionally integrated processes, Journal of Econometrics, 96: 155- 182.

TU, Y.- K., V. Clerenhugh, M. S. Gilthorpe (2004), Ratio variables in regression analysis can give rise to spurious results: illustration from two studies in periodontology, Journal of Dentistry, 32: 143- 151.

Wright, J. (1996), Local-to-spurious regression-problem, Econometric Theory, 12: 585-586.

Wasserfallen, W. (1988), Trends, random walks, and the expectations-augmented Phillips-Curve: evidence from six countries, Journal of Money, Credit, and Banking, 20/3: 306-318.