فصلنامه مدلسازی اقتصادسنجی – سال چهارم، شماره اول (پیاپی ۱۲)، زمستان ۱۳۹۷ (صفحات ۱۲۱ تا ۱۴۴)

مقایسه ای بین روش های ماکسیمم درستنمایی و بیزی برای برآورد پارامترهای سه مدل اقتصادسنجی فضایی

ناهید اشرفی (نویسنده مسئول)
دانشیار ریاضی، گروه ریاضی، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سمنان
nashrafi@semnan.ac.ir

فاطمه حسینی
استادیار آمار، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سمنان
fatemeh.hoseini@semnan.ac.ir

امید کریمی
استادیار آمار، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سمنان
omid.karimi@semnan.ac.ir

تاریخ دریافت: ۱۳۹۸/۰۲/۲۲ تاریخ پذیرش: ۱۳۹۸/۰۶/۱۰

چکیده

گاهی در اقتصادسنجی مشاهدات مورد مطالعه مستقل نیستند و وابستگی آن‌ها ناشی از موقعیت قرار گرفتن مشاهدات در فضای مورد مطالعه است. برای تحلیل این نوع از داده‌ها از مدل‌های رگرسیونی فضایی استفاده می‌شود. به دلیل وجود تعداد زیاد پارامتر در این مدل‌ها، برای به دست آوردن برآوردهای ماکسیمم درستنمایی از الگوریتم‌های تکرار شونده استفاده می‌شود که با مشکل پیچیدگی محاسبات مواجه است. علاوه بر این در مطالعات اقتصادی تعداد داده‌ها زیاد است که استفاده از رهیافت بیزی مفید به نظر می‌رسد. هدف استفاده از رهیافت‌های بیزی و ماکسیمم درستنمایی برای برآورد پارامترهای سه مدل معروف اقتصادسنجی فضایی و مقایسه عملکرد این دو رهیافت و مقایسه کارایی سه مدل و در نهایت پیاده‌سازی مدل‌ها و روش‌ها برروی دو مجموعه داده است. در هر دو مجموعه داده مشاهده می‌شود که نتایج رهیافت بیزی نسبت به رهیافت درستنمایی از دقت بهتری برخوردار است.

طبقه‌بندی JEL: C01، C11، C21
واژگان کلیدی: اقتصادسنجی فضایی، رهیافت ماکسیمم درستنمایی، رهیافت بیزی

۱. مقدمه

در مطالعاتی که داده‌ها دارای جزء مکانی هستند، دیگر به کارگیری شیوه‌های اقتصادسنجی معمولی که اغلب برمبنای استقلال متغیر پاسخ می‌باشند، معتبر نیست و استفاده از اقتصادسنجی فضایی پیشنهاد شده است. در اقتصادسنجی فضایی مدل‌های رگرسیونی فضایی برای مدل‌بندی داده‌هایی که دارای جزء مکانی هستند، پیشنهاد شده است. سه مدل معروف رگرسیون فضایی که در اقتصادسنجی فضایی مورد استفاده قرار می‌گیرد عبارتند از مدل اتورگرسیو فضایی یا مدل تاخیر فضایی[1] (SLM)، مدل دوربین فضایی[2] (SDM) و مدل خطای فضایی[3] (SEM). در مدل تاخیر فضایی اثر فضایی فقط در متغیر پاسخ منظور می‌شود، اما در مدل خطای فضایی فرض بر این است که همبستگی فضایی در جمله خطاست و در مدل دوربین فضایی اثر فضایی هم از طریق متغیر پاسخ و هم از طریق متغیرهای کمکی در مدل وارد می‌شود.

برای تحلیل این مدل‌ها معمولا انواع روش‌های برآورد ماکسیمم درستنمایی پیشنهاد شده است. به دلیل اضافه کردن همبستگی فضایی به این مدل‌ها و اضافه شدن تعداد پارامترهای مدل برای استفاده از رهیافت ماکسیمم درستنمایی تعدادی از پارامترها معلوم فرض می‌شوند و سایر پارامترها را برآورد و مجددا با استفاده از برآوردهای به‌دست آمده پارامترهایی که معلوم فرض شدند را برآورد می‌کنند. این یکی از مشکلات اساسی در استفاده از رهیافت ماکسیم درستنمایی در این مدل‌ها است. در رهیافت بیزی برای تمام پارامترها توزیع پیشین فرض می‌شود و با استفاده از تابع درستنمایی تشکیل شده و توزیع‌های پیشین فرض شده، توزیع پسین تشکیل و با استفاده از الگوریتم‌های موجود مثل الگوریتم زنجیرهای کارکوفی مونت کارلویی[4] به‌طور همزمان پارامترها برآورد می‌شوند.

ساختار مقاله به این‌صورت است که در بخش دوم سه مدل رگرسیون فضایی معرفی می‌شوند، در بخش سه تحلیل درستنمایی و در بخش چهار تحلیل بیزی سه مدل بیان می‌شود و در بخش چهارم و پنجم برای مقایسه بین دو رهیافت درستنمایی و بیزی به تحلیل دو مجموعه داده واقعی پرداخته می‌شود. در نهایت بحث و نتیجه‌گیری ارائه شده است.

---

[1]. Autoregressive Spatial Model or Spatial Lag Model
[2]. Spatial Durbin Model
[3]. Spatial Error Model
[4]. Markov Chain Monte Carlo

۲. پیشینه تحقیق

۲-۱. مرور ادبیات از دیدگاه نظری

اولین بار انسلین[۱] (۱۹۹۸) به معرفی اقتصادسنجی فضایی پرداخت و ادعا کرد که در مطالعات منطقه‌ای استفاده از اقتصادسنجی فضایی دارای قابلیت بهتری نسبت به روش-های معمول اقتصادسنجی است. انسلین و همکاران (۲۰۰۴)، لسج[۲] (۱۹۹۹)، لسج و پس[۳] (۲۰۰۹)، لسج و فیشر[۴] (۲۰۰۸) به طور مفصل‌تر روش‌های اقتصادسنجی فضایی را مورد بررسی قرار دادند. انسلین (۲۰۱۰) برای داده‌ها در طول زمان طولانی و با به‌کار بردن مدل‌های رگرسیونی فضایی به بررسی و مطالعه اقتصادسنجی فضایی پرداخت. بایوند و پورتنو[۵] (۲۰۰۴) به معرفی تکنیک‌های تحلیل داده‌های فضایی و کاربرد این تکنیک‌ها در مطالعات اقتصادسنجی با نرم‌افزار R پرداختند. اربیا[۶] (۲۰۱۶) و کلیجیان و پیراس (۲۰۱۷) انواع روش‌های اقتصادسنجی را به اقتصادسنجی فضایی تعمیم داده‌اند و آخرین روش‌های تجزیه و تحلیل داده‌های اقتصادی با وابستگی فضایی را می‌توان در این مطالعات مشاهده نمود.

۲-۲. مرور ادبیات از دیدگاه تجربی

از جمله مطالعات کاربردی که در آن‌ها از اقتصادسنجی فضایی استفاده شده است می‌توان به هولی و همکاران (۲۰۱۱)، مونکانن و همکاران (۲۰۱۲)، کیوته و پده (۲۰۱۱) اشاره نمود. لیایو و دانگ (۲۰۱۲) با استفاده از نوعی رگرسیون فضایی به مطالعه قیمت مسکن پرداخت و تیی (۲۰۱۸) اقتصادسنجی فضایی را برای بررسی تغییرات قیمت مسکن کشور هلند به کار گرفت. مطالعات داخلی پیرامون مدل‌های اقتصادسنجی فضایی که اغلب پیاده‌سازی این مدل‌ها برروی داده‌های واقعی است، می‌توان به اکبری و توسلی (۱۳۸۷) اشاره نمود که با رهیافت اقتصادسنجی فضایی تاثیر عوارض شهرداری‌ها را بر روی قیمت مسکن در شهر اصفهان بررسی نمودند. رحمانی و همکاران (۱۳۸۶) با رهیافت اقتصادسنجی فضایی به بررسی تاثیر سرمایه اجتماعی بر رشد اقتصادی در استان‌های ایران پرداختند. شهیادی (۱۳۸۴) تقاضای نیروی کار را در سطح استان‌های کشور با در

1. Anselin
2. Lesage
3. Pace
4. Lesage & Ficher
5. Bivand & Portnov
6. Arbia

نظر گرفتن همبستگی فضایی مورد بررسی قرار داد. از جمله مطالعات دیگر با رهیافت اقتصادسنجی فضایی می‌توان به مطالعات کسرایی (۱۳۸۵)، اکبری و فرهمند (۱۳۸۴)، رحمانی و حاجی رحیمی (۱۳۹۴)، طالبلو و همکاران (۱۳۹۵)، صادقی و همکاران (۱۳۹۶)، عسگری و همکاران (۱۳۹۶)، براتی و همکاران (۱۳۹۷) اشاره نمود.

۳. روش تحقیق

برای پیاده‌سازی این سه مدل رگرسیون فضایی و مقایسه بین دو روش ماکسیمم درستنمایی و بیزی برای تحلیل این مدل‌ها، دو مجموعه داده واقعی مورد بررسی قرار گرفته است. مجموعه داده اول مربوط به بررسی تاثیر قیمت مسکن و درآمد خانوار در ۴۹ منطقه و تاثیر آن‌ها بر روی متوسط سرقت خودرو و خانه در شهر کلمبوس است. برای این منظور از سه مدل رگرسیونی فضایی SLM، SDM و SEM استفاده شد و از دو رهیافت درستنمایی و بیزی، داده‌ها به‌طور کامل مورد مطالعه قرار گرفت. مجموعه داده دوم مربوط به داده‌های قیمت مسکن کمتر از یک سال ساخت در شهریور ماه سال ۱۳۹۴ مربوط به ۲۲ منطقه و ۲۶۸ محله از شهر تهران می‌باشد. با استفاده از سه مدل رگرسیون فضایی مذکور و رهیافت بیزی و ماکسیمم درستنمایی به مطالعه داده‌های مذکور پرداخته شد به طوری که متغیر مساحت مسکن به عنوان تنها متغیر کمکی در اختیار بود. با توجه به عدم نرمال بودن داده‌ها از تبدیل لگاریتمی داده‌ها استفاده شد.

۴. مدل‌های رگرسیونی در اقتصادسنجی فضایی

در این بخش سه مدل معروف رگرسیون فضایی، مدل تاخیر فضایی (SLM)، مدل دوربین فضایی (SDM) و مدل خطای فضایی (SEM) به‌طور خلاصه معرفی می‌شوند.
مدل SLM که در بسیاری از مطالعات مربوط به اقتصادسنجی و مطالعات فضایی از آن با عنوان مدل اتو رگرسیو فضایی (SAR) هم نام برده می‌شود، به صورت

$$\boldsymbol{y} = \rho W\boldsymbol{y} + X\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad \boldsymbol{\varepsilon} \sim N(0. \sigma^2 I_n) \qquad (۱)$$

بیان می‌شود،که در آن $\boldsymbol{Y}_{n\times1}$ بردار مشاهدات فضایی در n موقعیت $(s_1. s_2. \dots. s_n)$ و $X_{n\times k}$ ماتریس متغیرهای توضیحی، $\boldsymbol{\beta}_{k\times1}$ بردار پارامترهای رگرسیونی، $\boldsymbol{\varepsilon}_{n\times1}$ بردار خطاهای مستقل، ρ پارامتر تاخیر فضایی، $W_{n\times n}$ ماتریس وزن‌های فضایی و $W\boldsymbol{y}$ بردار تاخیر فضایی می‌باشد، I_n یک ماتریس واحد از بعد $n \times n$، لسج و پس (۲۰۰۹) و انسلین

(۱۹۹۸). پارامتر خود همبستگی فضایی یا همان تاخیر فضایی در بازه‌ی $(\frac{1}{\lambda_{min}}.\frac{1}{\lambda_{max}})$ محدود می‌شود که در آن λ_{min} و λ_{max} کوچک‌ترین و بزرگ‌ترین مقدار ویژه‌ی W هستند. ارتباط فضایی متغیرها به‌صورت دو به دو و به فرم عددی توسط ماتریس W مشخص می‌شود، که معمولا و نه لزوما همیشه ماتریسی متقارن در نظر گرفته می‌شود و در آن عنصر W_{ij} ارتباط فضایی مربوط به متغیر پاسخ در موقعیت iام با متغیر پاسخ در موقعیت j‌ام است. به عنوان مثال در برخی مطالعات مربوط به قیمت مسکن فاصله مرکز منطقه iام از منطقه j‌ام را با d_{ij} نشان می‌دهند و W_{ij} را از رابطه $\frac{1}{d_{ij}}$ به‌دست می‌آورند به‌طوری که با افزایش فاصله دو منطقه اثر فضایی کم می‌شود. لسج و پس (۲۰۰۹) نشان دادند که می‌توان W را به‌صورت یک ماتریس سطری تصادفی ساخت، به طوری که بردار تاخیر فضایی $W\boldsymbol{y}$ شامل مقادیر ساخته شده از متوسط مشاهدات همسایه می‌باشد. منظور از ماتریس سطری تصادفی، ماتریس نامنفی است که عناصر روی هر سطر آن نرمال‌سازی شده و بنابراین مجموع هر سطر یک می‌باشد و در این‌صورت $\lambda_{max} = 1$ و $\rho < 1$ است.
مدل SDM تعمیمی از مدل تأخیر فضایی است، به‌طوری که در مدل تأخیر فضایی علاوه بر این که همبستگی فضایی در متغیر پاسخ منظور می‌شود در متغیرهای توضیحی نیز در نظر گرفته می‌شود. این مدل به‌صورت

$$\boldsymbol{y} = \rho W\boldsymbol{y} + WX\boldsymbol{\gamma} + X\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \qquad \boldsymbol{\varepsilon} \sim N(0.\sigma^2 I_n) \qquad (۲)$$

بیان می‌شود. گاهی اوقات با در نظر گرفتن $Z = [XWX]$ و $\boldsymbol{\delta} = \begin{pmatrix}\boldsymbol{\beta}\\ \boldsymbol{\gamma}\end{pmatrix}$ مدل دوربین فضایی به صورت

$$\boldsymbol{y} = \rho W\boldsymbol{y} + Z\boldsymbol{\delta} + \boldsymbol{\varepsilon} \qquad (۳)$$

بازنویسی می‌شود که به شکل یک مدل تأخیر فضایی در آمده به‌طوری‌که $\boldsymbol{\varepsilon} \sim N(0.\sigma^2 I_n)$ و $\boldsymbol{\varepsilon} = (I - \rho W)\boldsymbol{y} - Z\boldsymbol{\delta}$ است.
مدل SEM معمولا به صورت

$$\begin{aligned} &\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u}. \ \ \boldsymbol{u} \sim N(0.\overbrace{\sigma^2 (I - \lambda W)^{-1}(I_n - \lambda W)'^{-1}}^{\Omega}) \\ &\boldsymbol{u} = \lambda W\boldsymbol{u} + \boldsymbol{\varepsilon}. \quad \boldsymbol{\varepsilon} \sim N(0.\sigma^2 I_n) \\ &\boldsymbol{y} = X\boldsymbol{\beta} + (I_n - \lambda W)^{-1}\boldsymbol{\varepsilon} \end{aligned} \qquad (۴)$$

در نظر گرفته می‌شود. برای بررسی وجود همبستگی مکانی و برای مقایسه‌ی چهار مدل رگرسیون خطی معمولی، مدل تأخیر فضایی، مدل دوربین فضایی و مدل خطای فضایی معیارهای مختلفی معرفی شده است. از جمله می‌توان به معیار مدل‌گزینی[1] AIC، آزمون ضرایب لاگرانژ[2]، آماره نسبت درستنمایی[3] و آزمون I موران[4] اشاره نمود. معمولا مدلی که در فرض صفر در نظر گرفته می‌شود مدل $\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}.\ \boldsymbol{\varepsilon} \sim N(0.\sigma^2 I_n)$ می‌باشد که با فرض استقلال مشاهدات مدل رگرسیون خطی معمولی است و فرض مقابل مدل مورد نظر قرار می‌گیرد.

۴-۱. تحلیل درستنمایی مدل‌های رگرسیونی فضایی

۴-۱-۱. تحلیل درستنمایی مدل تاخیر فضایی

برای به دست آوردن برآورد ماکسیمم درستنمایی پارامترهای مدل (۱) یعنی مدل SLM اگر فرض کنیم ρ معلوم است مدل را می‌توان به‌صورت

$$\boldsymbol{y} - \rho W \boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}.\sigma^2 I_n) \tag{۵}$$

نوشت. بنابراین $\boldsymbol{\varepsilon} = (I - \rho W)\boldsymbol{y} - X\boldsymbol{\beta}$ و داریم

$$L(\sigma^2.\boldsymbol{\varepsilon}) = (\frac{1}{2\pi\sigma^2})^{\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\right\}. \tag{۶}$$

پس تابع درستنمایی مشاهدات به‌صورت

$$\begin{aligned} L(\rho.\boldsymbol{\beta}.\sigma^2|\boldsymbol{y}) &= (\frac{1}{2\pi\sigma^2})^{\frac{n}{2}}|J| \\ &\times \exp\{-\frac{1}{2\sigma^2}[(I-\rho W)\boldsymbol{y} - X\boldsymbol{\beta}]'[(I-\rho W)\boldsymbol{y} - X\boldsymbol{\beta}]\} \end{aligned} \tag{۷}$$

خواهد شد، که در آن $|J| = |\frac{\partial \boldsymbol{\varepsilon}}{\partial \boldsymbol{y}}| = |I_n - \rho W|$ می‌باشد. با مشتق‌گیری از لگاریتم تابع درستنمایی نسبت به $\boldsymbol{\beta}$ برآورد ماکسیمم درستنمایی آن به‌صورت

$$\widehat{\boldsymbol{\beta}} = (X'X)^{-1}X'(I_n - \rho W)\boldsymbol{y} \tag{۸}$$

و با مشتق‌گیری نسبت به σ^2 برآورد ماکسیمم درستنمایی آن‌به‌صورت

[1]. Akaike Information Criteria
[2]. Lagrange Multiplier Test
[3]. Likelihood Ratio
[4]. Moran's I Test

$$\hat{\sigma}^2 = \frac{1}{n}[(I_n - \rho W)\boldsymbol{y} - X\boldsymbol{\beta}]'[(I_n - \rho W)\boldsymbol{y} - X\boldsymbol{\beta}] \\ = \frac{1}{n}(\boldsymbol{y} - X\boldsymbol{\beta})'(I_n - \rho W)'(I_n - \rho W)(\boldsymbol{y} - X\boldsymbol{\beta}) \qquad (۹)$$

به دست می‌آید. با جایگذاری $\widehat{\boldsymbol{\beta}}$ در $\hat{\sigma}^2$ و با تعریف

$$A = [\boldsymbol{y} - X(X'X)^{-1}X'\boldsymbol{y} - \rho(W\boldsymbol{y} - (X'X)^{-1}X'W\boldsymbol{y})] \qquad (۱۰)$$

داریم $\hat{\sigma}^2 = \frac{1}{n}A'A$ با قرار دادن $\boldsymbol{e}_0 = \boldsymbol{y} - X(X'X)^{-1}X'\boldsymbol{y}$ یعنی باقی‌مانده رگرسیون معمولی $\boldsymbol{y}$ روی X و $\boldsymbol{e} = W\boldsymbol{y} - (X'X)^{-1}X'W\boldsymbol{y}$ باقی‌مانده رگرسیون معمولی $W\boldsymbol{y}$ روی X می‌توان نوشت

$$\hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{e}_0 - \rho\boldsymbol{e}_1)'(\boldsymbol{e}_0 - \rho\boldsymbol{e}_1). \qquad (۱۱)$$

با جایگذاری روابط (۸) و (۱۱) در رابطه‌ی تابع درستنمایی که اکنون تابع درستنمایی متمرکز شده[1] به‌صورت

$$l_c(\rho) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\frac{1}{n}(\boldsymbol{e}_0 - \rho\boldsymbol{e}_1)'(\boldsymbol{e}_0 - \rho\boldsymbol{e}_1) + \sum_i \ln(1 - \rho\omega_i) \qquad (۱۲)$$

می‌شود، با مشتق‌گیری نسبت به ρ

$$\frac{\partial l_c(\rho)}{\partial\rho} = \frac{n(\boldsymbol{e}'_0\boldsymbol{e}_1 - \rho\boldsymbol{e}'_1\boldsymbol{e}_1)}{\boldsymbol{e}'_0\boldsymbol{e}_0 - 2\rho\boldsymbol{e}'_0\boldsymbol{e}_1 + \rho^2\boldsymbol{e}'_1\boldsymbol{e}_1} - \frac{\sum_i \omega_i}{1 - \rho\omega_i} = 0. \qquad (۱۳)$$

برآورد برای ρ که ماکسیمم کننده‌ی عبارت (۱۲) باشد باید طوری پیدا شود که در بازه‌ی $(\frac{1}{\lambda_{min}}.\frac{1}{\lambda_{max}})$ هم قرار بگیرد. برای به دست آوردن برآورد ماکسیمم درستنمایی ρ معمولا یک بردار $q \times 1$ از مقادیری که در بازه‌ی $[\rho_{min}.\rho_{max}]$ است به عنوان بردار اولیه در نظر می‌گیرند و آن‌گاه با تعریف $\boldsymbol{e}(\rho) = \boldsymbol{e}_0 - \rho\boldsymbol{e}_1$ می‌توان نوشت:

$$\begin{pmatrix} l(\rho_1) \\ \vdots \\ \vdots \\ l(\rho_q) \end{pmatrix} \propto \begin{pmatrix} \sum_i \ln(1 - \rho_1\omega_i) \\ \vdots \\ \vdots \\ \sum_i \ln(1 - \rho_q\omega_i) \end{pmatrix} - \frac{n}{2}\begin{pmatrix} \ln\boldsymbol{e}'(\rho_1)\boldsymbol{e}(\rho_1) \\ \vdots \\ \vdots \\ \ln\boldsymbol{e}'(\rho_q)\boldsymbol{e}(\rho_q) \end{pmatrix} \qquad (۱۴)$$

اکنون با به‌کار بردن یک الگوریتم تکرار شونده و بهینه‌سازی حول مقادیر $\rho_1 \dots \rho_q$، $l(\rho)$ مشخص و مقدار بهینه تعیین می‌شود، (باری و پس، ۱۹۹۷). پس از تعیین برآورد

---

[1]. Function Likelihood Concentrated

ماکسیمم درستنمایی ρ به‌صورت $\hat{\rho}$ می‌توان برآورد $\boldsymbol{\beta}$ و σ^2 را از روابط (۸) و (۱۱) به دست آورد.

۴-۱-۲. تحلیل درستنمایی مدل دوربین فضایی

برای تحلیل درستنمایی مدل دوربین فضایی تابع درستنمایی به‌صورت

$$L(\rho.\boldsymbol{\delta}.\sigma^2|\boldsymbol{y}) = (\frac{1}{2\pi\sigma^2})^{\frac{n}{2}}|I-\rho W| \times \exp\{-\frac{1}{2\sigma^2}[(I-\rho W)\boldsymbol{y}-Z\boldsymbol{\delta}]'[(I-\rho W)\boldsymbol{y}-Z\boldsymbol{\delta}]\} \quad (۱۵)$$

تشکیل می‌شود، که با مشتق‌گیری نسبت به $\boldsymbol{\delta}$ و σ^2 و با فرض معلوم بودن ρ ، برآورد $\boldsymbol{\delta}.\sigma^2$. $\hat{\boldsymbol{\delta}} = (Z'Z)^{-1}Z'(I_n-\rho W)\boldsymbol{y}$ و $\hat{\sigma}^2 = n^{-1}\boldsymbol{e}(\rho)\boldsymbol{e}'(\rho)$ خواهد شد، که در آن $\boldsymbol{e}(\rho) = \boldsymbol{e}_0 - \rho\boldsymbol{e}_1$، $\boldsymbol{e}_0$ باقی‌مانده‌ی رگرسیون $\boldsymbol{y}$ روی Z و $\boldsymbol{e}_1$ باقی‌مانده‌ی رگرسیون $W\boldsymbol{y}$ روی Z است، پس $\boldsymbol{e}_0 = \boldsymbol{y} - Z(Z'Z)^{-1}Z'\boldsymbol{y}$، $\boldsymbol{e}_1 = W\boldsymbol{y} - Z(Z'Z)^{-1}Z'W\boldsymbol{y}$ و $\boldsymbol{e}(\rho) = \boldsymbol{y} - \rho W\boldsymbol{y} - Z\hat{\delta}$ هستند. بعد از به‌دست آوردن برآورد ماکسیمم درستنمایی برای $\boldsymbol{\delta}$ و $\hat{\sigma}^2$ در رابطه‌ی درستنمایی، درستنمایی متمرکز شده تشکیل و با روش‌های بهینه‌سازی $\hat{\rho}$ مشابه مدل رگرسیون تأخیر فضایی به‌دست می‌آید و سپس $\hat{\delta}(\hat{\rho})$ و $\hat{\boldsymbol{\gamma}}(\hat{\rho})$ محاسبه می‌شوند.

۴-۱-۳. تحلیل درستنمایی مدل خطای فضایی

برای تحلیل درستنمایی مدل SEM و به دست آوردن برآورد ماکسیمم درستنمایی پارامترها ابتدا با ثابت نگه داشتن λ، $(I_n-\lambda W)(\boldsymbol{y}-X\boldsymbol{\beta}) = \boldsymbol{\varepsilon}$ و بنابراین تابع درستنمایی

$$L(\boldsymbol{\beta}.\lambda.\sigma^2|\boldsymbol{y}) = (\frac{1}{2\Pi\sigma^2})^{\frac{n}{2}}|I_n-\lambda W| \times \exp\left\{-\frac{1}{2\sigma^2}[(\boldsymbol{y}-X\boldsymbol{\beta})'(I_n-\lambda W)'(I_n-\lambda W)(\boldsymbol{y}-X\boldsymbol{\beta})]\right\} \quad (۱۶)$$

و لگاریتم تابع درستنمایی به صورت

$$l(\boldsymbol{\beta}.\lambda.\sigma^2|\boldsymbol{y}) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 + \ln|I_n-\lambda W| - \frac{1}{2\sigma^2}[(\boldsymbol{y}-X\boldsymbol{\beta})'(I-\lambda W)'(I-\lambda W)(\boldsymbol{y}-X\boldsymbol{\beta})] \quad (۱۷)$$

است. اکنون با در نظر گرفتن $\boldsymbol{e} = (I_n-\lambda W)(\boldsymbol{y}-X\boldsymbol{\beta})$ رابطه‌ی (۱۷) به‌صورت

$$l(\boldsymbol{\beta}.\lambda.\sigma^2|\boldsymbol{y}) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 + \ln|I_n - \lambda W| - \frac{1}{2\sigma^2}\boldsymbol{e}'\boldsymbol{e} \qquad (۱۸)$$

می‌شود و با فرض معلوم بودن λ برآورد برار ضرایب رگرسیونی

$$\hat{\beta} = [(X-\lambda WX)'(X-\lambda WX)]^{-1}(X-\lambda WX)'(\boldsymbol{y}-\lambda W\boldsymbol{y}) \\ \times [X'(\lambda)X(\lambda)]^{-1}]X'(\lambda)\boldsymbol{y}(\lambda) \qquad (۱۹)$$

و برآورد پارامتر واریانس به‌صورت

$$\hat{\sigma}^2 = [\boldsymbol{y}(\lambda) - X(\lambda)\boldsymbol{\beta}(\lambda)]'[\boldsymbol{y}(\lambda) - X(\lambda)\boldsymbol{\beta}(\lambda)]n^{-1} \\ = n^{-1}\boldsymbol{e}'(\lambda)\boldsymbol{e}(\lambda) \qquad (۲۰)$$

به دست می‌آیند. پس از جایگذاری $\hat{\beta}$ و $\hat{\sigma}^2$ در لگاریتم تابع درستنمایی و به‌کاربردن یک روش بهینه‌سازی مقدار برآورد ماکسیمم درستنمایی λ به دست می‌آید و سپس در روابط (۱۹) و (۲۰) جایگذاری و برآورد $\hat{\beta}$ و $\hat{\sigma}^2$ محاسبه می‌شود.

برآورد Ω نیز به‌صورت $\widehat{\Omega} = \hat{\sigma}^2[(I_n - \lambda W)'(I_n - \hat{\lambda}W)]^{-1}$ به دست می‌آید.

۴-۲. تحلیل بیزی مدل‌های رگرسیون فضایی

به منظور ارائه یک تحلیل بیزی برای مدل(۱) نیاز است تا پیشین‌هایی برای پارامترهای مدل اختیار شوند. معمولا استخراج اطلاعات در مورد پیشین و فرمول‌بندی آن در غالب یک توزیع، کار مشکلی است. در حالتی‌که درباره‌ی توزیع پیشین پارامترها، اطلاعی در اختیار نباشد، تحلیل بیزی با انتخاب پیشین‌هایی انجام می‌شود که اطلاعات اندکی از پیشین در اختیار می‌گذارند. به عنوان مثال می‌توان از پیشین‌های ناآگاهی بخش استفاده کرد. از جمله پیشین‌های ناآگاهی بخش می‌توان به پیشین‌های تخت، جفریز یا توزیع‌های پیشین مبهم اشاره کرد. در این مقاله پیشین‌های معمولی که در اغلب مقالات به‌کار گرفته شده است استفاده می‌شود. برای پارامارهای رگرسیونی توزیع پیشین ناآگاهی بخش از نوع مبهم، یعنی یک توزیع نرمال با واریانس بزرگ به صورت $N(\boldsymbol{c}.\sigma^2 T)$ و برای σ^2 توزیع پیشین گامای معکوس $IG(a.b)$ فرض می‌شوند و برای پارامتر ρ مطابق لسج و پیس (۲۰۰۹) توزیع یکنواخت به صورت $\pi(\rho) \sim U(\lambda_{min}^{-1}.\lambda_{max}^{-1})$ در نظر گرفته می‌شود، که در آن λ_{min} و λ_{max} کوچک‌ترین و بزرگ‌ترین مقدار ویژه W هستند. بنابراین

$$\pi(\sigma^2) = \frac{b^a}{\Gamma(a)}(\sigma^2)^{-(a+1)}\exp(-\frac{b}{\sigma^2}) \quad \sigma^2 > 0.a.b > 0 \qquad (۲۱)$$

و توزیع پیشین توام

$$\begin{aligned}\pi(\boldsymbol{\beta}.\sigma^2) &= \pi(\boldsymbol{\beta}|\sigma^2)\pi(\sigma^2)\\ &= N(\boldsymbol{c}.\sigma^2 T)IG(a.b)\\ &= \frac{b^a}{(2\pi)^{\frac{k}{2}}|T|^{\frac{1}{2}}\Gamma(a)}(\sigma^2)^{-(a+\frac{k}{2}+1)} \times \exp[\frac{-\{(\boldsymbol{\beta}-\boldsymbol{c})'T^{-1}(\boldsymbol{\beta}-\boldsymbol{c})+2b\}}{2\sigma^2}]\end{aligned} \quad (۲۲)$$

هستند. اکنون فرض کنید تمام مشاهدات به صورت$D = \{\boldsymbol{y}.X.W\}$ و پارامترهای مدل به صورت $\boldsymbol{\theta} = \{\boldsymbol{\beta}.\sigma^2.\rho\}$ در نظر گرفته شوند. تابع درستنمایی برای مدل به صورت

$$\pi(D|\boldsymbol{\beta}.\sigma^2.\rho) = (2\pi\sigma^2)^{\frac{-n}{2}}|A|\exp\{\frac{-1}{2\sigma^2}(A\boldsymbol{y}-X\boldsymbol{\beta})'(A\boldsymbol{y}-X\boldsymbol{\beta})\} \quad (۲۳)$$

است، که در آن $A = (I_n - \rho W)$ است.

از رهیافت بیزی توزیع پسین به صورت $\pi(\boldsymbol{\theta}|D) = \frac{\pi(D|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(D)}$ بیان می‌شود و بنابراین می‌توان نوشت.$\pi(\boldsymbol{\theta}|D) \propto \pi(D|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$ پس

$$\begin{aligned}\pi(\boldsymbol{\beta}.\sigma^2.\rho|D) &\propto \pi(D|\boldsymbol{\beta}.\sigma^2.\rho)\pi(\boldsymbol{\beta}|\sigma^2)\pi(\sigma^2)\pi(\rho)\\ &\propto (\sigma^2)^{a^*+(\frac{k}{2})+1}|A|\\ &\times \exp\{\frac{-1}{2\sigma^2}[2b^* + (\boldsymbol{\beta}-\boldsymbol{c}^*)'(T^*)^{-1}(\boldsymbol{\beta}-\boldsymbol{c}^*)]\}\end{aligned} \quad (۲۴)$$

که در آن

$$\begin{aligned}\boldsymbol{c}^* &= (X'X + T^{-1})^{-1}(X'A\boldsymbol{y} + T^{-1}\boldsymbol{c})\\ T^* &= (X'X + T^{-1})^{-1}\\ a^* &= a + \frac{n}{2}\\ b^* &= b + \frac{(c'T^{-1}c + \boldsymbol{y}'A'A\boldsymbol{y} - (c^*)'(T^*)^{-1}c^*)}{2}\end{aligned}$$

هستند. همان‌طور که مشاهده می‌شود توزیع پسین به صورت حاصل‌ضرب یک توزیع $N(\boldsymbol{c}^*.\sigma^2 T^*)$ و گامای معکوس $IG(a^*.b^*)$ است. اگر در توزیع پسین $\rho = 0$ باشد، بنابراین $A = I_n$ یعنی مدل اقتصادسنجی غیرفضایی در نظر گرفته شود، یا ρ معلوم باشد، آن‌گاه انتخاب توزیع پیشین $N(\boldsymbol{c}.T)$ برای $(\boldsymbol{\beta}|\sigma^2)$ و توزیع گامای معکوس$IG(a.b)$ برای σ^2که به صورت $\pi(\beta.\sigma^2) \approx NIG(\boldsymbol{c}.T.a.b)$ خلاصه می‌شود، منجر به یک توزیع پسین مزدوج $NIG(\boldsymbol{c}^*.T^*.a^*.b^*)$ خواهد شد. اما در

حالت فضایی این نتیجه برقرار نیست و شکل پسین به صورت رابطه (۱۰) می‌باشد. این توزیع پسین شکل مشخصی ندارد و برای به دست آوردن برآوردهای بیزی پارامترها از الگوریتم‌های MCMC استفاده می‌شود.

برای اجرای الگوریتم گیبز ابتدا مقادیر اولیه $(\boldsymbol{\beta}^{(0)}.\sigma^{2^{(0)}}.\rho^{(0)})$ را در نظر بگیرید و $m=0$ قرار دهید و سپس توزیع‌های شرطی کامل پارامترهای رگرسیونی به صورت

$$\pi\left(\boldsymbol{\beta}^{(m+1)}\middle|\sigma^{2^{(m)}}.\rho^{(m)}.D\right)\propto\pi(D|\boldsymbol{\beta}.\sigma^2.\rho)\pi(\boldsymbol{\beta}|\sigma^2)=N(c^*.\sigma^2T^*)\quad (۲۵)$$

و توزیع‌های شرطی کامل پارامتر واریانس

$$\pi\left(\sigma^{2^{(m+1)}}\middle|\boldsymbol{\beta}^{(m+1)}.\rho^{(m)}.D\right)\propto\pi(D|\boldsymbol{\beta}.\sigma^2.\rho)\pi(\sigma^2)=IG(a^*.b^*)\quad (۲۶)$$

تشکیل دهید. توزیع‌های شرطی کامل برای پارامترهای $\boldsymbol{\beta}$ و σ^2 شکل بسته‌ای دارد. اما توزیع شرطی کامل برای ρ به‌صورت

$$\begin{aligned}\pi(\rho^{(m+1)}|\boldsymbol{\beta}^{(m+1)}.\sigma^{2^{(m+1)}}.D)&\propto\pi(D|\boldsymbol{\beta}.\sigma^2.\rho)\pi(\rho)\\&\propto|I_n-\rho W|\\&\times\exp(-\tfrac{1}{2\sigma^2}(A\boldsymbol{y}-X\boldsymbol{\beta})'(A\boldsymbol{y}-X\boldsymbol{\beta})I(\rho)\end{aligned}\quad (۲۷)$$

است، که شکل مشخصی ندارد و از الگوریتم متروپلیس-هاستینگس[1] استفاده می‌شود. برای استفاده از این الگوریتم ابتدا باید یک توزیع پیشنهادی[2] برای $\pi(\rho|\boldsymbol{\beta}.\sigma.D)$ در نظر گرفته شود که مقادیر کاندید ρاز آن تولید شوند. توزیع کاندید به صورت نرمال قدم زدن تصادفی پیشنهاد می‌شود. فرض کنید ρ^c مقداری است که ρ دارد و ρ^* مقدار کاندید است که پذیرفته یا رد می‌شود و احتمال پذیرش به صورت

$$\psi_H(\rho^c.\rho^*)=\min\left[1.\frac{\pi(p^*|\boldsymbol{\beta}.\sigma^2)}{\pi(\rho^c|\boldsymbol{\beta}.\sigma^2)}\right]\quad (۲۸)$$

1. Metropolis Hastings
2. Proposal Distribution

محاسبه می‌شود. در نهایت بعد از اجرای الگوریتم‌های MCMC فوق توزیع پسین مشخص می‌شود و برآورد بیزی پارامترها از روی میانگین توزیع‌های پسین حاشیه‌ای محاسبه می‌شود. برای پیشگویی y^* جدید از توزیع پیشگو به صورت

$$\begin{aligned}\pi(y^*|D) &= \int \pi\,(y^*.\boldsymbol{\theta}|D)d\boldsymbol{\theta} \\ &= \int \pi\,(y^*|D.\boldsymbol{\theta})\pi(\boldsymbol{\theta}|D)d\boldsymbol{\theta}\end{aligned} \qquad (۲۹)$$

استفاده می‌شود، که در آن $\boldsymbol{\theta} = (\boldsymbol{\beta}.\sigma^2.\rho)$ است. به همین ترتیب می‌توان این رهیافت بیزی را برای دو مدل SDM و SEM بیان کرد.

۴-۳. تحلیل بیزی و درستنمایی یک مجموعه داده واقعی

کلمبوس[1] مرکز ایالت اوهایو آمریکا و بزرگ‌ترین شهر این ایالت می‌باشد. برای ۴۹ منطقه ی این شهر متوسط درآمد خانوار منطقه، متوسط قیمت مسکن در منطقه و متوسط تعداد سرقت خانه و خودرو در اختیار می‌باشد. ماتریس متغیرهای کمکی به‌صورت

$$x_i = [\mathbf{1}.Income_i.Housevalue_i]'. \quad \boldsymbol{\beta} = [\beta_0.\beta_{INC}.\beta_{Hoval}]'$$

است. خلاصه آماره‌های میانگین، چارک اول، دوم و سوم این سه متغیر برای ۴۹ منطقه در جدول۱ ارائه شده است.

جدول ۱: خلاصه آماره های متغیرهای مورد بررسی

متغیر	میانگین	چارک اول	میانه	چارک سوم
قیمت مسکن H	۳۸/۴۴	۲۵/۷۰	۳۳/۵۰	۴۳/۳۰
درآمد خانوار I	۱۴/۳۸	۹/۹۶	۱۳/۳۸	۱۸/۳۳
سرقت C	۳۵/۱۳	۲۰/۰۵	۳۴/۰۰	۴۸/۵۹

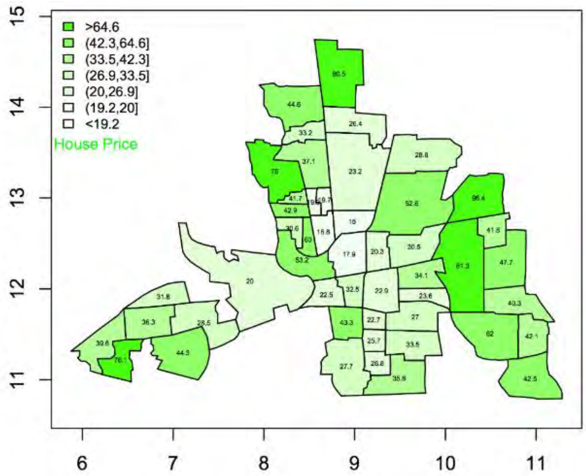
منبع: محاسبه شده توسط مؤلفین با استفاده از نرم‌افزار R

برای توصیف بهتر داده‌ها نمودار تقسیم‌بندی براساس منطقه در شکل‌های ۱، ۲ و ۳ رسم شده است. در شکل ۱ قیمت مسکن، در شکل ۲ درآمد خانواده و در شکل ۳ متوسط سرقت خانه و خودرو در ۴۹ منطقه مورد بررسی ارائه شده‌اند. با مقایسه شکل ۱ و شکل ۲ به‌شور شهودی ارتباط مستقیمی بین قیمت مسکن و درآمد خانواده در این شهر وجود دارد. همچنین با مقایسه شکل ۱ و ۲ با شکل ۳ به نظر می‌رسد ارتباط معکوسی بین

[1]. Columbus

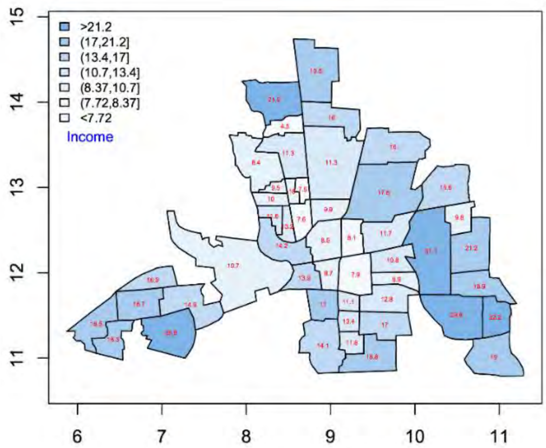
درآمد خانواده، قیمت مسکن و متوسط سرقت خانه و خودرو وجود دارد. برای بررسی بیشتر این مجموعه داده ابتدا مدل رگرسیون معمولی به صورت $M_1: \boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ در نظر گرفته شد و از روش OLS برآورد پارامترها محاسبه شد. نتایج در جدول ۲ ارائه شده است. برآورد ماکسیمم درستنمایی پارامتر σ^2 برای رگرسیون معمولی، ۱۱/۴۳ و مقدار آماره آزمون معنی‌داری رگرسیون ۲۸/۳۹ که ارتباط رگرسیونی بین متغیرهای کمکی و پاسخ پذیرفته می‌شود.

شکل ۱: نقشه شهر کلمبوس و قیمت مسکن در مناطق مورد بررسی



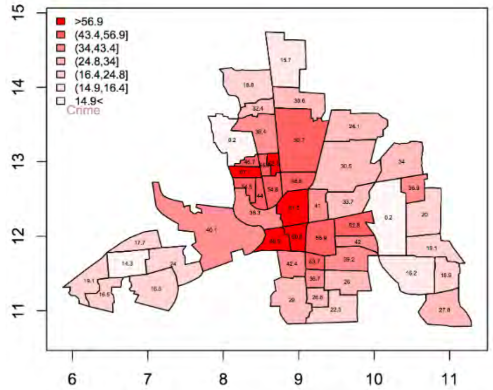
منبع: رسم شده توسط مؤلفین با استفاده از نرم‌افزار R

شکل ۲: نقشه شهر کلمبوس و توزیع درآمد در مناطق مورد بررسی



منبع: رسم شده توسط مؤلفین با استفاده از نرم‌افزار R

شکل ۳: نقشه شهر کلمبوس و توزیع سرقت خانه و ماشین در مناطق مورد بررسی



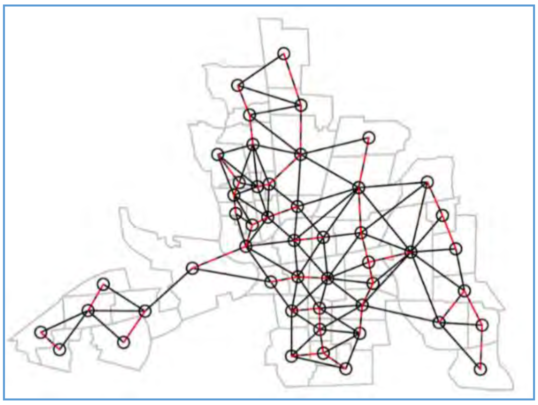
منبع: رسم شده توسط مؤلفین با استفاده از نرم‌افزار R

از طرفی به نظر می‌رسد بین متغیرهای پاسخ همبستگی از نوع مکانی وجود دارد و علاوه بر این متغیرهای کمکی نیز ممکن است دارای همبستگی مکانی باشند. نتایج آزمون I موران برای بررسی وجود همبستگی مکانی مشاهدات و خطاها محاسبه شد که به ترتیب مقدار آماره ۲/۶۸۲ و ۲/۴۷۷ و مقادیر احتمال به ترتیب ۰/۰۰۷۳ و ۰/۰۱۳به دست آمد که همبستگی مکانی بین مشاهدات و خطاها رد نمی‌شود. آزمون هاسمن برای مقایسه مدل خطای فضایی SEM با رگرسیون معمولی انجام شد و مقدار آن ۶/۴۷۳ با مقدار احتمال ۰/۰۹۱ به دست آمد که فرض صفر در سطح ۰/۰۵ یعنی عدم وجود تفاوت معنی دار بین مدل SEM و مدل رگرسیون معمولی رد نمی‌شود. سه مدل رگرسیون فضایی SLM، SDM و SEM نیز بر روی داده اعمال گردید. برای این منظور ماتریس وزنی فضایی W به صورت یک ماتریس سطری تصادفی در نظر گرفته شد. برای این منظور از تابع nb2listw در بسته نرم‌افزاری R-spdep استفاده گردید. حداکثر همسایه‌ی در نظر گرفته شده برای یک منطقه ۱۷ همسایه و حداقل همسایه در نظر گرفته شده برای هفت منطقه با ۲ همسایه می‌باشد. در نهایت W که یک ماتریس49×49 می‌باشد با حدود ده درصد عناصر غیرصفر تشکیل گردید.

متوسط تعداد همسایه‌ی در نظر گرفته شده ۵ می‌باشد. در شکل ۴ شهر کلمبوس و همسایگی‌های انتخاب شده نمایش داده شده است. نتایج تحلیل درستنمایی و برآوردهای ماکسیمم درستنمایی پارامترهای رگرسیونی سه مدل در جدول ۲ ارائه شده است. همچنین برآورد σ^2 برای سه مدل به ترتیب ۹/۹۵۸، ۹/۷۴۹ و ۹/۹۹۹، برآورد ρ برای مدل‌های SLM و SDM به ترتیب ۰/۴۰۴ و ۰/۳۸۴ به دست آمد. برآورد λ برای مدل

SEM، ۰/۵۲۱ حاصل شد. مقدار معیار AIC برای هر سه مدل محاسبه و حدود ۳۷۸ و برای مدل رگرسیون معمولی ۳۸۲/۷۵ به دست آمد. برای تحلیل بیزی مدل‌ها از الگوریتم‌های MCMC با ۵۰۰۰۰ تکرار استفاده شد. نمودار تابع چگالی حاشیه‌ای پسین برای پارامتر خودهمبستگی فضایی برای سه مدل در شکل ۵ رسم شده است. خطوط قرمز برآورد ماکسیمم درستنمایی این پارامتر را نشان می‌دهد.

شکل ۴: شهر کلمبوس، تعیین همسایگی‌ها



منبع: رسم شده توسط مؤلفین با ستفاده از نرم‌افزار R

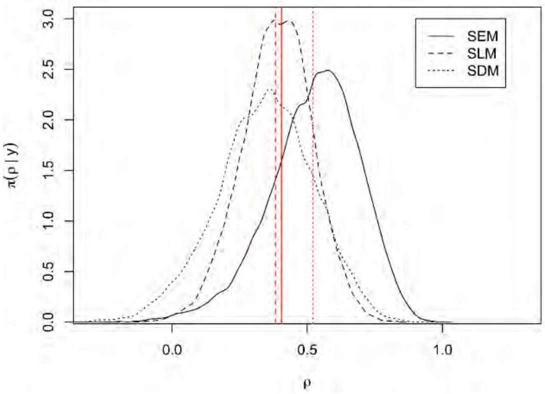
جدول ۲: برآورد درستنمایی پارامترها و خلاصه آماره‌ها

پارامتر	برآورد	انحراف معیار	آماره آزمون	مقدار احتمال
رگرسیون معمولی				
β_0	۶۸/۶۱۹	۴/۷۳۶	۱۴/۹۰	۰/۰۰۰۰
β_H	-۰/۲۷۴	۰/۱۰۳	-۲/۶۵۴	۰/۰۱۰۹
β_I	-۱/۵۹۷	۰/۳۳۴	-۴/۷۸۰	۰/۰۰۰۵
رگرسیون SLM				
β_0	۴۶/۸۵۱	۷/۳۱۵	۶/۴۵۵	۰/۰۰۰۰
β_H	-۰/۲۶۹	۰/۰۹۰	-۲/۹۹۶	۰/۰۰۲۷
β_I	-۱/۰۷۴	۰/۳۱۱	-۳/۴۵۳	۰/۰۰۰۶
رگرسیون SDM				
β_0	۴۵/۵۹۳	۱۳/۱۲۸	۳/۴۷۳	۰/۰۰۰۵
β_H	-۰/۲۹۹	۰/۰۹۰۸	-۳/۲۹۸	۰/۰۰۰۹
β_I	-۰/۹۳۹	۰/۳۳۸	-۲/۷۷۶	۰/۰۰۵۵
γ_H	۰/۲۶۶	۰/۱۸۴	۱/۴۴۹	۰/۱۴۷۳
γ_I	-۰/۶۱۸	۰/۵۷۷	-۱/۰۷۲	۰/۲۸۳۹
رگرسیون SEM				

β_0	۴۸/۰۵۴	۴/۳۱۵	۱۱/۴۸۷	۰/۰۰۰۰
β_H	-۰/۳۰۸	۰/۰۹۲۶	-۳/۳۲۷	۰/۰۰۰۸
β_I	-۰/۹۹۵	۰/۳۳۷	-۲/۹۵۴	۰/۰۰۳۱

منبع: برآوردها توسط مؤلفین با استفاده از نرم‌افزار R انجام شده است.

شکل ۵: تابع چگالی حاشیه‌ای پسین برای پارامتر خودهمبستگی فضایی



منبع: رسم شده توسط مؤلفین با ستفاده از نرم‌افزار R

پس از همگرایی الگوریتم‌های MCMC و تولید نمونه به اندازه کافی از توزیع‌های شرطی کامل میانگین این توزیع‌ها مشخص و در نهایت برآورد بیزی پارامترها مشخص می‌شود. نتایج برآورد بیزی پارامترها و انحراف معیار برآوردها و چندک‌های ۰/۰۲۵، ۰/۵ و ۰/۹۷۵ برای پارامترها در جدول ۳ مشاهده می‌شود. با توجه به نتایج تحلیل درستنمایی و بیزی به نظر می‌رسد که متوسط درآمد افراد و متوسط قیمت مسکن خانه‌ها در مناطق مختلف شهر کلمبوس روی متوسط سرقت خودرو و دزدی از منازل تاثیرگذار بوده و تاثیر آن منفی است یعنی در مناطق پردرآمد متوسط سرقت کم و در مناطق با قیمت بالای مسکن سرقت کمتری انجام شده است.

جدول۳: برآورد بیزی پارامترها و خلاصه آماره‌ها برای سه مدل رگرسیون فضایی

پارامتر	برآورد	انحراف معیار	چندک ۰/۰۲۵	چندک ۰/۵	چندک ۰/۹۷۵
رگرسیون SLM					
ρ	۰/۳۸۴	۰/۱۳۳	۰/۱۰۸	۰/۳۸۸	۰/۶۳۷
β_0	۴۸/۸۳۲	۸/۴۱۰	۳۱/۶۹۲	۴۷/۶۲۲	۶۴/۸۳۲
β_H	-۰/۲۷۱	۰/۰۹۴	-۰/۴۵۶	-۰/۲۷۱	-۰/۰۸۵
β_I	-۱/۰۹۵	۰/۳۵۳	-۱/۷۹۸	-۱/۰۹۳	-۰/۳۹۸
رگرسیون SDM					
ρ	۰/۳۴۰	۰/۱۷۷	-۰/۰۲۳	۰/۳۴۹	۰/۶۶۶

β_0	۴۸/۶۵۳	۱۴/۶۹۳	۲۱/۱۶۰	۴۸/۱۴۷	۷۸/۲۸۸
β_H	-۰/۳۰۰	۰/۰۹۹	-۰/۴۹۴	-۰/۲۹۸	-۰/۱۰۷
β_I	-۰/۹۵۷	۰/۳۷۷	-۱/۷۰۰	-۰/۹۵۹	-۰/۲۲۷
γ_H	۰/۲۶۲	۰/۱۹۸	-۰/۱۱۵	۰/۲۶۱	۰/۶۵۴
γ_I	-۰/۶۹۵	۰/۶۳۹	-۱/۹۹۱	-۰/۶۷۹	۰/۵۵۷
رگرسیون SEM					
$\rho = \lambda$	۰/۵۲۹	۰/۱۷۱	۰/۱۵۱	۰/۵۴۶	۰/۸۱۷
β_0	۶۰/۸۹۲	۶/۶۳۹	۴۷/۱۰۶	۶۱/۰۹۹	۷۳/۳۴۱
β_H	-۰/۳۰۸	۰/۰۹۷	-۰/۴۹۸	-۰/۳۰۹	-۰/۱۱۷
β_I	-۰/۹۹۵	۰/۴۰۰	-۱/۸۰۵	-۰/۹۸۳	-۰/۲۳۰

منبع: برآوردها توسط مؤلفین با ستفاده از نرم‌افزار R انجام شده است.

برای مقایسه بین سه مدل برازش و مقادیر مجذور میانگین مربع خطای نسبی یعنی

$$RRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{(y_i-\hat{y}_i)^2}{\hat{y}_i^2}} \qquad (۳۰)$$

استفاده شد. برای سه مدل این معیار محاسبه و به‌ترتیب ۰/۳۹۸۰، ۰/۳۸۵۳ و ۰/۳۵۰۲ برای روش بیزی و به ترتیب ۰/۴۰۱۵، ۰/۴۲۹۸ و ۰/۳۹۵۹۶ برای رهیافت درستنمایی به‌دست آمدند، که بیان‌گر برتری مدل SEM نسبت به دو مدل SLM و SDM و دقت بالاتر رهیافت بیزی نسبت به درستنمایی می‌باشد. همچنین برای مقایسه مدل‌ها از معیار عامل بیزی[1] که توسط گودمن (۱۹۹۹) به‌صورت $\pi(M_j|y) = \frac{\pi(y|M_j)}{\sum_{j=1}^{J}\pi(y|M_j)}$تعریف می شود، استفاده شد. با در نظر گرفتن $M_1 = SLM$، $M_2 = SDM$ و $M_3 = SEM$ مقادیر لگاریتم معیار مذکور برای سه مدل به ترتیب به‌صورت

$$\log\pi(M_1|y) = -8.2112$$
$$\log\pi(M_2|y) = -19.8553$$
$$\log\pi(M_3|y) = -0.0003$$

به دست آمد، که مدل M_3 یعنی SEM با این معیار به عنوان مدل بهتر انتخاب می‌شود.

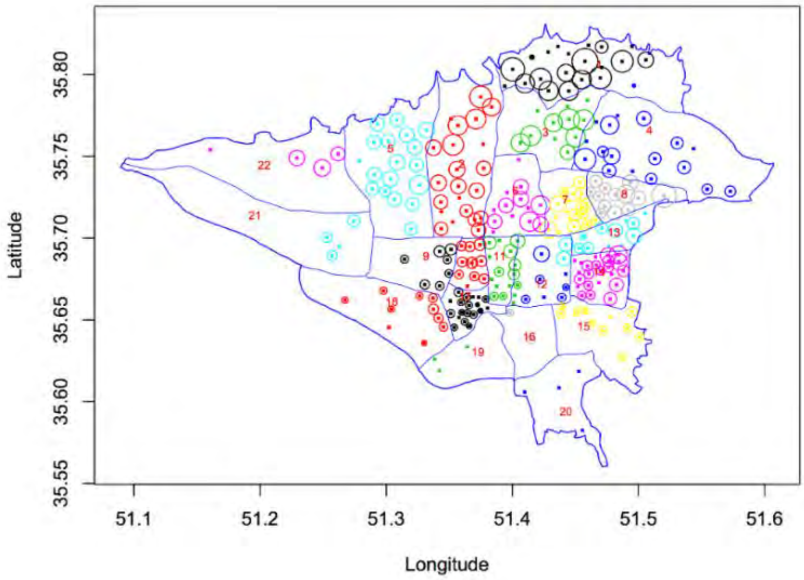
۴-۴. تحلیل داده‌های قیمت مسکن

قیمت مسکن نسبت به موقعیت قرارگیری مسکن در محله‌های مختلف همبسته‌است بنابراین داده‌های قیمت مسکن دارای همبستگی از نوع فضایی (مکانی) است. در این

[1]. Bayes Factor

قسمت داده‌های مربوط به متوسط قیمت مسکن در مناطق مختلف تهران با استفاده از سه مدل رگرسیون فضایی بیان شده و دو رهیافت بیزی و درستنمایی تحلیل می‌شود. داده‌های قیمت مسکن برای شهر تهران به‌صورت روزانه در سامانه اطلاعات بازار املاک ایران ثبت می‌شود که این داده‌ها را می‌توان در سایت www.hmi.mrud.ir مشاهده کرد. قیمت مسکن نوساز (تا یک سال ساخت) برای محله‌های ۲۲منطقه در شهریور ماه سال ۱۳۹۴ استخراج شد. موقعیت‌مناطق و داده‌ای در دسترس روی نقشه تهران در شکل ۶ مشخص شده است. دایره بزرگتر نشان‌دهنده قیمت مسکن بیشتر و اعداد منطقه را نشان می‌دهند. خلاصه آماره‌های مربوط به این مجموعه داده در جدول ۴ ارائه شده است. میانگین قیمت مسکن حدود چهار میلیون، ماکسیمم قیمت مسکن ۱۲ میلیون و مینیمم قیمت مسکن حدود یک و نیم میلیون است.

شکل ۶: موقعیت داده‌های متوسط قیمت مسکن تهران (شهریور ۱۳۹۴)



منبع: رسم شده توسط مؤلفین با استفاده از نرم‌افزار R

جدول ۴: خلاصه آماره های مربوط به قیمت مسکن شهریور ۹۴ تهران (میلیون تومان)

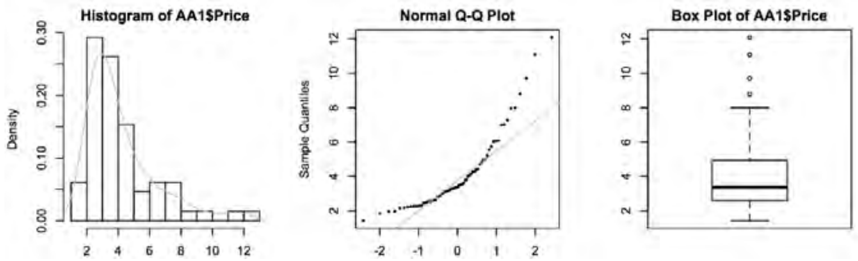
متغیر	میانگین	چارک اول	میانه	چارک سوم	مینیمم	ماکسیمم
قیمت مسکن	۴/۱۷۷	۲/۵۸۴	۳/۳۷۶	۴/۹۳۰	۱/۴۲۳	۱۲/۰۸۰

منبع: معیارها توسط مؤلفین با استفاده از نرم‌افزار R به‌دست آمده است.

نمودار هیستوگرام، جعبه‌ای و نرمال چندک-چندک داده‌ها در شکل ۷ رسم شده است که عدم نرمال بودن داده‌ها را نشان می‌دهد. از آماره آزمون شاپیرو ویلک و کلموگروف اسمیرنف برای اطمینان از عدم نرمال بودن استفاده شد. مقدار احتمال این دو آزمون به ترتیب ۰/۰۰۰ و ۰/۰۴ به دست آمد و نرمال بودن داده‌ها رد شد.

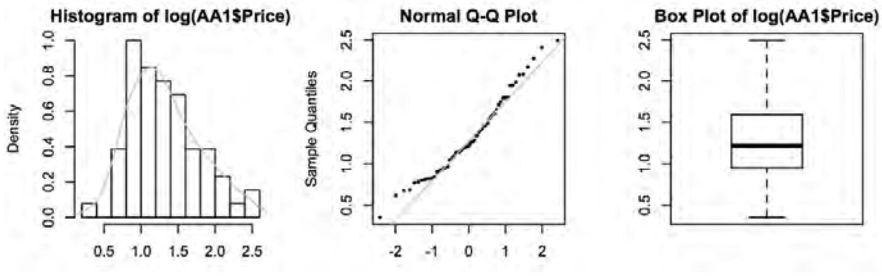
از تبدیل لگاریتمی برروی داده‌ها استفاده شد. برای تبدیل لگاریتمی داده‌ها نمودار هیستوگرام، جعبه‌ای و نرمال چندک-چندک در شکل ۸ رسم شد،که نرمال بودن داده‌ها را نشان می‌دهد. مقدار آماره شاپیرو ویلک و کلموگروف اسمیرنف به ترتیب ۰/۰۹۶۳ و ۰/۵۸۴۸ به دست آمد که فرض نرمال بودن پذیرفته می‌شود. مساحت ملک به عنوان متغیر کمکی در نظر گرفته شد. داده‌ها با سه مدل رگرسیون فضایی SLM ،SEM و SDM برازش و نتایج تحلیل ماکسیمم درستنمایی و بیزی به ترتیب در جداول ۵ و ۶ آورده شده است.

شکل ۷: نمودار جعبه‌ای، چندک-چندک و هیستوگرام داده‌های مسکن تهران



منبع: رسم شده توسط مؤلفین با ستفاده از نرم‌افزار R

شکل ۸: نمودار جعبه‌ای، چندک-چندک و هیستوگرام لگاریتم داده‌های مسکن تهران



منبع: رسم شده توسط مؤلفین با ستفاده از نرم‌افزار R

جدول ۵: برآورد درستنمایی پارامترها و خلاصه آماره‌ها (داده‌های مسکن)

پارامتر	برآورد	انحراف معیار	آماره آزمون	مقدار احتمال
رگرسیون SLM				
β_0	-۰/۴۲۴۹	۰/۱۳۳۴	-۳/۲۸۸۴	۰/۰۰۱۰
β_{Area}	۰/۰۰۶۸	۰/۰۰۱۱	۶/۳۷۱۱	۰/۰۰۰۰
رگرسیون SDM				
β_0	-۰/۳۱۳۶	۰/۲۷۴۵	-۱/۲۷۰۳	۰/۲۰۴۰
β_{Area}	۰/۰۰۶۹	۰/۰۰۱۱	۶/۴۱۴۲	۰/۰۰۰۰
γ_{Area}	-۰/۰۰۱۶	۰/۰۰۲۸	-۰/۵۸۷۲	۰/۵۵۷۱
رگرسیون SEM				
β_0	۰/۶۶۵۸	۰/۳۷۸۷	۱/۷۵۸۲	۰/۰۷۸۷
β_{Area}	۰/۰۰۶۵	۰/۰۰۱۱	۶/۱۵۹۶	۰/۰۰۰۰

منبع: برآوردها توسط مؤلفین با ستفاده از نرم‌افزار R به‌دست آمده است.

برای تحلیل بیزی از الگوریتم MCMC با ۵۰۰۰۰ تکرار استفاده شد. برازش و مقادیر مجذور میانگین مربع خطای نسبی (RRMSE) برای سه مدلSLM، SDM و SEM محاسبه و به‌ترتیب ۰/۲۳۶۷، ۰/۲۴۰۲ و ۰/۲۳۸۹ برای رهیافت بیزی و به ترتیب ۰/۲۴۱۸، ۰/۲۴۴۷ و ۰/۲۴۳۵ برای رهیافت درستنمایی به‌دست آمدند، که بیان‌گر برتری مدل SLM نسبت به دو مدل SEM و SDM و دقت تقریبا بهتر رهیافت بیزی نسبت به درستنمایی برای این مجموعه داده می‌باشد. برآورد σ^2 برای سه مدل رگرسیون فضایی SLM، SDM و SEM به ترتیب ۰/۰۶۱۱، ۰/۰۶۰۱۲ و ۰/۰۶۰۸ و مقدار معیار AIC برای سه مدل به ترتیب ۱۷/۲۲۴، ۱۸/۹۵۹ و ۱۹/۲۲۲ حاصل شد. مقدار AIC برای مدل رگرسیون معمولی برای این داده‌ها ۶۲/۱۲۱ بدست آمد که نشان می‌دهد هر سه مدل رگرسیون فضایی از رگرسیون معمولی خیلی بهتر عمل کرده‌اند و مدل SLM از AIC کمتری برخوردار است و مدل مناسب برای برازش این داده‌ها است.

جدول ۶: برآورد بیزی پارامترها و خلاصه آماره‌ها (برای داده‌های مسکن)

پارامتر	برآورد	انحراف معیار	چندک ۰/۰۲۵	چندک ۰/۵	چندک ۰/۹۷۵
رگرسیون SLM					
ρ	۰/۸۱۶۸	۰/۰۸۲۲	۰/۶۳۹۷	۰/۸۲۷۲	۰/۹۵۶۳
β_0	-۰/۴۰۱۳	۰/۱۲۹۲	-۰/۶۵۳۹	-۰/۴۰۳۱	-۰/۱۲۸۹
β_{Area}	۰/۰۰۶۹	۰/۰۰۱۱	۰/۰۰۴۷	۰/۰۰۶۹	۰/۰۰۹۳
رگرسیون SDM					
ρ	۰/۸۲۵۹	۰/۰۹۵۳	۰/۶۰۵۷	۰/۸۴۰۳	۰/۹۶۸۴

β_0	۰/۳۴۴۳-	۰/۲۴۶۹	۰/۸۹۳۴-	۰/۳۴۱۹-	۰/۱۹۸۱
β_{Area}	۰/۰۰۶۹	۰/۰۰۱۰	۰/۰۰۴۷	۰/۰۰۶۹	۰/۰۰۹۳
γ_{Area}	۰/۰۰۰۷-	۰/۰۰۲۶	۰/۰۰۷۲-	۰/۰۰۰۸-	۰/۰۰۶۲
رگرسیون SEM					
λ	۰/۹۱۹۸	۰/۰۵۸۸	۰/۷۷۱۴	۰/۹۳۴۳	۰/۹۹۰۹
β_0	۰/۵۶۱۸	۰/۳۲۸۶	۱/۲۷۹۰-	۰/۶۱۳۹	۲/۰۲۰۶
β_{Area}	۰/۰۰۶۵	۰/۰۰۱۰	۰/۰۰۴۳	۰/۰۰۶۵	۰/۰۰۸۶

منبع: برآوردها توسط مؤلفین با ستفاده از نرم‌افزار R به‌دست آمده است.

برآورد ρ برای مدل‌های SLM و SDM به ترتیب ۰/۸۴۳۳ و ۰/۸۷۱۸ به دست آمد که با استفاده از آماره آزمون والد صفر بودن این پارامتر در دومدل رد شد. برآورد λ برای مدل SEM، ۰/۹۱۶۲ حاصل شد و با استفاده از آماره آزمون والد صفر بودن آن رد شد. با توجه به نتایج جدول هر سه مدل ارتباط بین مساحت و قیمت را تایید می‌کنند و صفر بودن اثر ثابت عرض از مبدا فقط در مدل SLM رد نمی‌شود.

۵. نتیجه‌گیری و پیشنهادات

در این مطالعه بر استفاده از اقتصادسنجی فضایی برای مشاهداتی که دارای وابستگی مکانی هستند تاکید شد و به بررسی مفصل سه مدل رگرسیونی فضایی معروف که برای این نوع از داده‌ها استفاده شود پرداخته شد. نحوه برآورد پارامترها با استفاده از روش ماکسیمم درستنمایی و رهیافت بیزی ارائه شد. در یک مجموعه داده واقعی مربوط به بررسی ارتباط بین متوسط قیمت مسکن، درآمد و سرقت خانه و خودرو در ۴۹ منطقه از شهر کلمبوس مدل‌ا و روش‌ها پیاده‌سازی شد.با توجه به نتایج حاصل برای این مجموعه داده متوسط درآمد افراد و متوسط قیمت مسکن خانه‌ها در مناطق مختلف شهر کلمبوس روی متوسط سرقت خودرو و دزدی از منازل تاثیرگذار بوده و تاثیر آن منفی است یعنی در مناطق پردرآمد متوسط سرقت کم و در مناطق با قیمت بالای مسکن سرقت کمتری انجام شده است. برای مقایسه بین سه مدل SDM، SLM و SEM از معیار مجذور میانگین مربع خطای نسبی[1] استفاده شد که برای این سه مدل فضایی با استفاده از رهیافت درستنمایی به‌ترتیب ۰/۳۹۸۰، ۰/۳۸۵۳ و ۰/۳۵۰۲ و برای رهیافت بیزی به ترتیب ۰/۴۰۱۵، ۰/۴۲۹۸ و ۰/۳۹۵۹۶ برآورد شدند که بیان‌گر برتری مدل SEM نسبت به دو مدل SLM و SDM و دقت بالاتر رهیافت بیزی نسبت به درستنمایی می‌باشد. در

[1]. Relative Root Mean Square Error

یک مجموعه داده دیگر به بررسی ارتباط قیمت و مساحت مسکن در شهریور ماه ۱۳۹۴ شهر تهران پرداخته شد. با توجه به عدم نرمال بودن داده‌ها از تبدیل لگاریتمی داده‌ها استفاده شد. پس از برازش مدل‌ها و مقایسات انجام شده مدل مناسب مدل SLM انتخاب و در هر سه مدل و با هر دو رهیافت بیزی و درستنمایی وجود اثر فضایی و ارتباط مستقیم بین قیمت مسکن و مساحت پذیرفته می‌شود. با استفاده از معیار مجذور میانگین مربع خطای نسبی و براساس این دو مجموعه داده مشاهده شد که رهیافت بیزی برای این نوع از مدل‌های اقتصادسنجی فضایی نسبت به رهیافت درستنمایی از عملکرد تقریبا بهتری برخوردار است و بستگی به نوع داده‌ها و با استفاده از معیارهای موجود می‌توان یکی از سه مدل SLM،SDM و SEM را به عنوان مدل مناسب انتخاب نمود، که برای داده‌های کلمبوس مدل SEM و برای داده‌های قیمت مسکن تهران مدل SLM از دقت بیشتری برخوردار بود، که ممکن است با تغییر داده‌ها و بستگی به نوع و ماهیت داده‌ها هر کدام از سه مدل به عنوان مدل مناسب انتخاب شوند.

در عمل داده‌های اقتصادی معمولا حجیم هستند و به اصطلاح با داده‌های بزرگ سروکار داریم که در اینصورت به دست آوردن برآوردهای ماکسیمم درستنمایی ناممکن و رهیافت بیزی با الگوریتم‌های MCMC معمول ارائه شده بسیار زمان‌بر و گاهی اوقات غیرممکن است و به راحتی همگرایی‌ها حاصل نمی‌شود. به عنوان پیشنهاد می‌توان از روش‌های درستنمایی و بیزی تقریبی استفاده نمود و زمان و دقت محاسبات را با روش‌های معمول مقایسه کرد. در مدل‌های معرفی شده برای راحتی محاسبات معمولا توزیع نرمال برای خطاها منظور می‌شود، به طوری‌که حسینی و همکاران در مدل‌های آمیخته خطی تعمیم-یافته فضایی نشان دادند که استفاده از توزیع چوله نرمال برای اثرات تصادفی فضایی باعث افزایش دقت برآورد پارامترها و پیشگویی در مدل می‌شود. به عنوان پیشنهاد می‌توان در مدل‌های اقتصادسنجی فضایی از توزیع‌های انعطاف‌پذیرتر مثل چوله نرمال، چوله نرمال بسته و چوله تی استفاده کرد که کلاس بزرگتری نسبت به توزیع نرمال هستند.

فهرست منابع:

اکبری، نعمت‌الله و ناهید، توسلی (۱۳۸۷)، تحلیل تاثیر عوارض شهرداری‌ها بر قیمت مسکن: مطالعه‌ی موردی شهر اصفهان (یک رهیافت اقتصادسنجی فضایی)، فصلنامه بررسی‌های اقتصادی، ۵ (۱): ۴۷-۶۴.

اکبری، نعمت‌الله و شکوفه، فرهمند (۱۳۸۴)، همگرایی اقتصادی کشورهای اسلامی و بررسی سرریزهای منطقه‌ای با تاکید بر نقش منتخبی از کشورهای حوزه خلیج فارس: مطالعه‌ای بر مبنای اقتصادسنجی فضایی، پژوهشنامه بازرگانی، ۳۴: ۳۲-۱.

براتی، جواد، کریمی موغاری، زهرا و نادر، مهرگان (۱۳۹۷)، بررسی محرک‌های توسعه منطقه‌ای در ایران: رویکرد اقتصادسنجی فضایی، فصلنامه علمی پژوهشی اقتصاد مقداری، ۱۵ (۱): ۲۰۱-۲۲۴.

رحمانی، تیمور، عباسی نژاد، حسین و میثم، امیری (۱۳۸۶)، بررسی تاثیر سرمایه اجتماعی بر رشد اقتصادی ایران؛ مطالعه موردی استان‌های کشور با روش اقتصادسنجی فضایی، فصلنامه پژوهش‌های اقتصادی. ۶ (۲): ۲۴-۳۳.

رحمانی، زانیار و محمود، حاجی رحیمی (۱۳۹۴)، مقایسه الگوی اقتصادسنجی ساده و اقتصادسنجی فضایی جهت ارزش‌گذاری هدانیک زمین کشاورزی، (مطالعه موردی: بخش مرکزی شهرستان سنندج)، تحقیقات اقتصاد کشاورزی، ۷ (۲): ۱۴۵-۱۶۲.

شهیادی، سعید (۱۳۸۴)، تحلیل فضایی و تقاضای نیروی کار در ایران، فصلنامه پژوهش‌های اقتصادی، ۶ (۲۵): ۷۲-۸۵.

صادقی، سید کمال، پورعباداللهان کویج، محسن، محمدزاده، پرویز، کریمی، زهرا و پروین، علی‌مرادی افشار (۱۳۹۶)، بررسی عوامل اقتصادی موثر بر دموکراسی در کشورهای در حال توسعه با استفاده از رهیافت اقتصادسنجی فضایی، فصلنامه علمی پژوهشی اقتصاد مقداری، ۱۴ (۱)۱: ۱۱۹-۱۴۲.

طالبلو، رضا، محمدی، تیمور و هادی، پیردایه (۱۳۹۵)، تحلیل انتشار فضایی تغییرات قیمت مسکن در استان‌های ایران؛ رهیافت اقتصادسنجی فضایی، فصلنامه پژوهش‌های اقتصادی، ۱۷ (۶۶): ۵۵-۹۵.

عسگری، حشمت‌الله، عالی، رضا و عزیز، مراسلی (۱۳۹۶)، تاثیر صادرات نفت بر همگرایی یا واگرایی GDP سرانه کشورهای عضو اوپک با رویکرد اقتصادسنجی ترکیبی فضایی، فصلنامه علمی پژوهشی اقتصاد مقداری، ۱۴(۱): ۳۱-۶۶.

کسرایی، اسرافیل (1385)، نظریه هم‌گرایی، وابستگی فضایی و رشد منطقه‌ای، مجله تحقیقات اقتصادی، شماره ۷۷، ۲۵-۱۶.

Anselin, L. (1988), Spatial Econometrics: Methods and Models, Studies in Operational Regional Science, Springer, Texas, United States.

Anselin, L. (2010), Thirty years of Spatial Econometrics, Papers in Regional Science, 89: 3-25.

Anselin, L. & Florax, R., & Rey, S. (2004), Advances in Spatial Econometrics, Methodology, Tools and Applications. Springer, Berlin, 10.1007/978-3-662-05617-2.

Arbia, G. (2016), Spatial Econometrics: A Broad View, Foundations and Trends in Econometrics, (8:3-4): 145-265.

Barry, R. P. & Pace, R. K. (1997), Kriging with Large Data Sets Using Sparse Matrix Techniques, Communications in Statistics Simulation and Computation,
26(2): 619-629.

Bivand R. S. & Portnov, B. (2004), Exploring Spatial Data Analysis Techniques Using R: the Case of Observations with No -Neighbors. In: Anselin L., Florax R.J.G.M., Rey S.J. (eds) Advances in spatial Econometrics. Advances in Spatial Science. Springer, Berlin, Heidelbergv.

Goodman, S. (1999), Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. Annals of Internal Medicine. 130(12): 1005-13.

Holly, S., Pesaran, H. M., & Yamagata, T. (2011), The Spatial and Temporal Diffusion of House Prices in the UK, Journal of Urban Economics, 69(1): 2-23.

Hosseini, F., Eidsvik, J., & Mohammadzadeh M. (2011), Approximate Bayesian Inference in Spatial GLMM with Skew Normal Latent Variables, Computational Statistics and Data Analysis, 55: 1791–1806.

Kelejian, H., & Piras, G. (2017), Spatial Econometrics. London, England: Academic Press.

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

Lesage, J. P. & Ficher, M. (2008), Spatial Growth Regression: Model Specification, Estimation and Interpretation, Spatial Economic Analysis, 3: 275–304.

Lesage, J. P. & Pace, R. K. (2009), Introduction to Spatial Econometrics, Chapman and Hall/CRC, Boca Raton, FL.

Liao, W. C. & Wang, X. (2012), Hedonic House Prices and Spatial Quantile Regression, Journal of Housing Economics, 21: 16-27.

Monkkonen, P., Wong, K. & Begley, J. (2012), Economic Restructuring, Urban Growth, and Short-Term Trading: The Spatial Dynamics of the Hong Kong housing Market (1992–2008), Regional Science and Urban Economics, 42: 396-406.

Teye, A. L. (2018), Diffusion and Risks of House Prices in the Netherlands. DOI:10.7480/abe.2018.3.