مقارنة بين طريقتي الامكان الاعظم والبيزية في تقدير انموذج الانحدار الذاتي المكاني مع تطبيق عملي على مرض السرطان في العراق

أ.م.د هيفاء طه عبد
قسم الاحصاء / كلية الادراة والاقتصاد
الجامعة المستنصرية

الباحثة ساره اسامة سعد
قسم الاحصاء / كلية الادراة والاقتصاد
الجامعة المستنصرية

تاريخ استلام البحث: 17 / 04 /2022

تاريخ قبول البحوث: 19/ 05 / 2022

نشر البحث في العدد السابع عشر: ديسمبر / كانون اول 2022

رمز التصنيف ديوي / النسخة الالكترونية (Online): 2522-64X/512.943

رمز التصنيف ديوي / النسخة الورقية (Print): 2519-948X/512.943

مقارنة بين طريقتي الامكان الاعظم والبيزية في تقدير انموذج الانحدار الذاتي المكاني مع تطبيق عملي على مرض السرطان في العراق

أ.م.د هيفاء طه عبد
قسم الاحصاء / كلية الادراة والاقتصاد
الجامعة المستنصرية

الباحثة ساره اسامة سعد
قسم الاحصاء / كلية الادراة والاقتصاد
الجامعة المستنصرية

المستخلص

لقد جذب تحليل البيانات المكانية في الاونة الاخير اهتمام الباحثيين الاحصائيين وخاصة لان اهمال البعد المكاني في التحليل يؤثر على النتائج ويؤدي الى ضياع معلومات مهمة، ولذلك تم اللجوء الى نماذج الانحدار المكانية التي يمكن من خلالها دراسة مدى تاثر المتغير المعتمد بالمتغيرات التوضيحية في ظل وجود الاعتماد المكاني لمفردات الظاهرة المدروسة، ولقد تم استعمال طريقة الامكان الاعظم MLE وطريقة بيز المعلمية في تقدير انموذج الانحدار الذاتي المكاني (SAR) في ظل وجود مصفوفة الاوزان المكانية المعدلة التي تم بناءها بلاعتماد على معيار روك Rook للتجاور، ولقد تم الاعتماد على بيانات مرضى السرطان المتمثلة بعدد مرضى السرطان في كل محافظة من محافظات العراق كمتغير معتمد اما المتغيرات التوضيحية فقد تمثلت بمعدل العمر ومعدل حجم الورم لمرضى السرطان وعدد المناطق الملوثه باليوارنيوم في المحافظة، ومن خلال نتائج معيار متوسط الخطأ النسبي المطلق MAPE الذي استعمل للمقارنة بين طرائق التقدير تم التوصل الى ان طريقة الامكان الاعظم MLE افضل من طريقة بيز في تقدير انموذج الانحدار الذاتي المكاني في ظل وجود مصفوفة الاوزان المكانية المعدلة التي تم بناءها بلاعتماد على معيار روك Rook للتجاور.

الكلمات المفتاحية: انموذج الانحدار الذاتي المكاني، طريقة الامكان الاعظ ML ، طريقة بيز، معيار متوسط الخطأ المطلق النسبي MAPE ، معيار روك Rook للتجاور، معامل موران

A Comparison between the Maximum Likelihood and Bayesian Methods to estimate the Spatial Auto Regressive Model with a practical application on Cancer in Iraq

Abstract

Recently, spatial data analysis has attracted the attention of statistical researchers, especially because neglecting the spatial dimension in the analysis affects the results and leads to the loss of important information. Therefore, spatial regression models have been resorted to, through which it is possible to study the extent to which the dependent variable is affected by the explanatory variables in light of the presence of spatial dependence for the studied vocabulary. The Bayes method and the MLE method were used to estimate the spatial autoregressive model (SAR) in the presence of the modified spatial weights matrix that was built based on Rook juxtaposition criterion. Cancer patient data on the number of cancer patients in each province of Iraq have been relied upon as an approved variable, but the illustrative variables were the age rate, the rate of tumor size of cancer patients and the number of areas contaminated with uranium in the province, and through the results of the Mean Absolute Percentage Error Criteria (MAPE) that used to compare the methods of estimation it was found that the Maximum Likelihood method (MLE) is better than Bayes method to Estimate the Spatial Auto Regressive Model in the absence of a modified spatial weight matrix built up based on the Rook juxtaposition standard.

Key words: Spatial Auto Regressive Model, Moran Coefficient, Bayes method, MLE method, Mean Absolute Percentage Error Criteria, Moran Coefficient.

1- المقدمة

يمثل الاقتصاد القياسي المكاني أحد أهم الفروع في الاقتصاد القياسي، وتاتي أهميتة من كونة يتعامل مع البيانات المكانية التي تتميز بأحتواءها على صفة التبعية المكانية (الارتباط المكاني) وصفة عدم التجانس المكاني، وقد قام الباحث Anselin بمناقشة هذا الأمر في ابحاثه عام (1988-2001)، وإن هاتان الصفتين تجعل من غير المناسب التعامل مع البيانات المكانية بتقنيات الاقتصاد القياسي التقليدي التي تهتم النماذج الخاصة به بالاعتمادية بين المشاهدات خلال فترة زمنية معينة دون الاخذ بنظر الاعتبار الاعتمادية المكانية مما يؤدي الى الحصول على تقديرات غير كفؤه نتيجة لعدم تحقق فرضيات التحليل بالاضافة الى إهمال الكثير من المعلومات المكانية الخاصة بالبيانات نتيجة عدم القدرة على توظيفها والاستفادة منها عند استعمال نماذج القياس الاقتصادي الغير مكانية(2). ان مشكلة البحث تتلخص في وجود الاعتماد المكاني في البيانات للظاهرة المدروسة والتي لا تؤخذ بنظر الاعتبار عند دراسة هذا النوع من الظواهر مما يؤدي الى ضياع معلومات مهمة تنعكس على دقة تقدير معلمات الانموذج وعلى النتائج التي تم الحصول عليها، الامر الذي يستدعي الى البحث عن انموذج قياسي يأخذ بنظر الاعتبار الارتباط المكاني لبيانات الظاهرة، ويهدف هذا البحث الى تقدير معلمات إنموذج الانحدار الانحدار الذاتي المكاني (SAR) باستعمال طريقة الامكان الاعظم MLE وطريقة بيز ومن ثم المقارنة بين هاتين الطريقتين من اجل التوصل الى الطريقة الافضل في تقدير انموذج الانحدار المكاني .

2- انموذج الانحدار الذاتي المكاني

ويسمى ايضا انموذج الانحدار الذاتي المكاني المختلط Mixed Spatial Autoregressive Model ويمثل هذا الانموذج حالة خاصة من انموذج الانحدار الذاتي المكاني العام (SAC) الذي تم اقتراحة من قبل (Anselin) ، ويمكن التعبير عن انموذج الانحدار الذاتي المكاني رياضيا من خلال الصيغة التالية: (5)

$$\underline{y} = \lambda W \underline{y} + X\underline{B} + \underline{u} \dots\dots\dots\dots\dots\dots\dots (1)$$

$$\underline{\mathrm{u}} \sim \mathrm{N}(0, \sigma^2 \mathrm{I}_n)$$

حيث ان :

$\underline{y}$: متجة ابعاده (nx1) ويمثل ألمتغير المعتمد.

λ : معلمة الاعتماد المكاني.

X : مصفوفة ابعادها (n * k+1) للمتغيرات التوضيحية .

B : متجة ابعاده (k+1 * 1) للمعلمات المتعلقة بالمتغيرات التوضيحية x.

$\underline{u}$: متجة الاخطاء ابعاده (nx1) والذي يتوزع توزيع طبيعي بمتوسط صفر وتباين $\sigma^2 \mathrm{I}_n$.

W : مصفوفة الاوزان المكانية ابعادها (nxn) وهي ثابتة وتكون محددة مسبقاً.

3- مصفوفة التجاور المكانية

وهي احدى مصفوفات الاوزان المكانية ويتم بناءها بالاعتماد على معيار التجاور بين الوحدات المكانية ، فاذا كانت الوحدتين المكانيتين متجاورة فيعطى لها القيمة (1) اما اذا كانت الوحدتين المكانيتين غير متجاوره فيعطى لها القيمة (0) وكذلك فإن الوحدة المكانية لا تجاور نفسها فيعطى لها القيمة (0) ايضا في مصفوفة الاوزان المكانية عند بناءها ، وهنالك عدة معايير لبناء لمصفوفة التجاور المكانية وقد تم الاعتماد على معيار روك للتجاور في هذا البحث، حيث يكون التجاور بالاعتماد على اذا كان الوحدتين المكانية مشتركتين بحد طولي غير صفري.[3]

4- مصفوفة الاوزان المكانية المعدلة

لقد سميت هذه المصفوفة بالمصفوفة المعدلة لانها تتكون من مصفوفة التجاور المكانية ولكن بعد اجراء بعض العمليات عليها من أجل ان يصبح مجموع كل صف فيها يساوي الواحد الصحيح اي ان: [3]

$$\sum_{j=1}^{n} W_{ij}^{Adj} = 1$$

ويتم بناء مصفوفة الاوزان المعدلة من خلال تطبيق الصيغة التالية على قيم اوزان مصفوفة التجاور:

$$W_{ij}^{Adj} = \begin{cases} \frac{W_{ij}}{\sum W_{ij}} & if \quad W_{ij} = 1 \\ 0 & if \quad W_{ij} = 0 \end{cases} \qquad \dots (2)$$

5- طرائق التقدير المعلمية

ولتقدير انموذج الانحدار الذاتي المكاني سيتم استعمال الطريقتين التالية :

5-1 طريقة الامكان الاعظم:

ان طريقة الامكان الاعظم تم استعمالها في تقدير انموذج الانحدار الذاتي المكاني اول مرة من قبل Ord (1975)، وهي احدى الطرائق المهمة في التقدير لكونها تعطي افضل تقدير لمعالم الانموذج من بين عدة تقديرات،(8) حيث ان استعمال طريقة المربعات الصغرى في تقدير نماذج الانحدار المكاني يؤدي الى الحصول على تقديرات متحيزه وغير متسقة وبالتالي يكون الانموذج غير كفوء ولهذا تم اللجوء الى طريقة الامكان الاعظم في التقدير،(1) ومن اجل تقدير معلمات الانموذج نقوم باعادة كتابة الانموذج في الصيغة (1) كما يلي :

$$\underline{u} = \underline{y} - \lambda W\underline{y} - X\underline{B} \qquad \dots\dots \qquad (3)$$

$$\underline{u} = (I - \lambda W)\underline{y} - X\underline{B} \qquad \dots\dots \qquad (4)$$

وان صيغة دالة الامكان الاعظم لانموذج الانحدار الذاتي المكاني كما يأتي: [8]

$$L(\underline{B}, \sigma^2, \lambda) = (2\pi\sigma^2)^{\frac{-n}{2}} + |I - \lambda W| \exp((1/2\sigma^2)\underline{u}'\underline{u}) \dots (5)$$

حيث ان :

$$\underline{u}'\underline{u} = \left((I - \lambda W)\underline{y} - X\underline{B}\right)' \left((I - \lambda W)\underline{y} - X\underline{B}\right)$$
$$\underline{u}'\underline{u} = \left(\underline{y}'(I - \lambda W)' - \underline{B}'X'\right) \left((I - \lambda W)\underline{y} - X\underline{B}\right)$$
$$\underline{u}'\underline{u} = \underline{y}'(I - \lambda W)'(I - \lambda W)\underline{y} - \underline{y}'(I - \lambda W)'X\underline{B} - \underline{B}'X'(I - \lambda W)\underline{y} + \underline{B}'X'X\underline{B}$$
$$\underline{u}'\underline{u} = \underline{y}'(I - \lambda W)'(I - \lambda W)\underline{y} - 2\underline{B}'X'(I - \lambda W)\underline{y} + \underline{B}'X'X\underline{B}$$

ومن اجل تبسيط دالة الامكان الاعظم سنأخذ اللوغاريتم الطبيعي للصيغة (5) لنحصل على الصيغة الاتيه: [3]

$$LnL(\underline{B}, \sigma^2, \lambda) = \frac{-n}{2} Ln2\pi - \frac{n}{2} Ln\,\sigma^2 + Ln\,|I - \lambda W| - \left(\frac{1}{2\sigma^2}\right) \underline{y}'(I - \lambda W)'(I - \lambda W)\underline{y} - 2\underline{B}'X'(I - \lambda W)\underline{y} + \underline{B}'X'X\underline{B} \quad \dots (6)$$

ولتقدير معلمات الانموذج نقوم باشتقاق دالة الامكان الاعظم في الصيغة (6) كما يأتي :

$$\frac{\partial LnL(\underline{B}, \sigma^2, \lambda)}{\partial \underline{B}} = -\left(\frac{1}{2\sigma^2}\right) 2X'(I - \lambda W)\underline{y} + 2X'X\,\widehat{\underline{B}}_{MLE} \quad \dots\dots(7)$$

وبجعل الصيغة (7) مساوية الى الصفر وبتبسيطها يتم الحصول على:

$$\widehat{\underline{B}}_{MLE} = (X'X)^{-1} X'(I - \lambda W)\underline{y} \quad \dots\dots \quad (8)$$

$$\widehat{\sigma^2}_{MLE} = \frac{\underline{u}'\underline{u}}{n} \quad \dots\dots \quad (9)$$

$$\widehat{\sigma^2}_{MLE} = (\underline{y} - \lambda W\underline{y} - X\,\widehat{\underline{B}}_{MLE})'(\underline{y} - \lambda W\underline{y} - X\,\widehat{\underline{B}}_{MLE})/n \quad \dots (10)$$

حيث يلاحظ في الصيغ اعلاة انه لا يمكن ايجاد كل من معلمات الانموذج $\widehat{\underline{B}}_{MLE}$، $\widehat{\sigma^2}_{MLE}$ اذا كانت معلمة الاعتماد المكاني λ مجهولة ولهذا اقترح Ord في عام (1975م) صيغة من اجل حساب المحدد $|I - \lambda W|$ كما يأتي: [3]

$$|I - \lambda W| = \prod_{i=1}^{n}(1 - \lambda\,\omega_i) \quad \dots\dots \quad (11)$$

$$Ln\,|I - \lambda W| = \sum_{i=1}^{n} Ln\,(1 - \lambda\omega_i) \quad \dots\dots \quad (12)$$

حيث ان :

ω_i : قيم Eigen values لمصفوفة الاوزان المكانية W.

وبتعويض قيمة كل من الصيغ التالية (8)،(9) ،(12) في صيغة دالة الامكان الاعظم (6) نحصل على دالة الامكان المركزة (Concentrated likelihood function) وهي دالة غير خطية، ومن خلال استعمال الطرق التكرارية لدالة الامكان المركزه في الصيغة (13) ادناة سيتم الحصول على معلمة الاعتماد المكاني λ لانموذج الانحدار الذاتي المكاني (SAR):

$$Lf = \frac{-n}{2} Ln \left[\frac{(e_o - \lambda e_L)'((e_o - \lambda e_L)}{n} \right] + \sum_{i=1}^{n} Ln\,(1 - \lambda \omega_i) \ldots \ldots (13)$$

$$\underline{e_o} = \underline{y} - X\underline{\widehat{B}_o}$$

$$\underline{e_L} = W\underline{y} - X\underline{\widehat{B}_L}$$

$$\underline{\widehat{B}_o} = (X'X)^{-1} X'y$$

$$\underline{\widehat{B}_L} = (X'X)^{-1} X'Wy$$

حيث ان:

$\underline{\widehat{B}_o}$: موجة معلمات انموذج انحدار y على X.

$\underline{\widehat{B}_L}$:موجة معلمات انموذج االنحدار Wy على X.

$\underline{e_o}$:موجة البواقي لانموذج أنحدار y على X.

$\underline{e_L}$: موجة البواقي لانموذج انحدار Wy على X.

5 -2 طريقة بيز :Bayes Method

إن طريقة بيز في التقدير قد نشأت من قبل العالم البريطاني توماس بيز (1702-1761) وهو صاحب مبرهنة بيز التي تستند على الاحتمالات الشرطية، وتختلف طريقة بيز عن الطرق الاخرى في التقدير بكونها تتعامل مع المعلمات المراد تقديرها على إنها متغيرات عشوائية، بمعنى ان المعلمات تمتلك توزيع احتمالي بفتره معينة بعكس الطرائق التقليدية التي تتعامل مع معلمات الانموذج كثوابت غير معلومة القيمة، ويتألف الانموذج البيزي من ثلاثة اجزاء ويمكن توضيحها من خلال الصيغة الاتية : [4]

$$p(Z|D) = \frac{P(D|Z)\,P(Z)}{P(D)} \quad \ldots \quad (14)$$

حيث ان :

$P(Z)$: التوزيع المسبق للمعلمة Z، ويمثل أعتقاد الباحث السابق لتوزيع المعلمة (Z) قبل ملاحظة البيانات وهذه الاعتقادات تكون إما معلوماتية او غير معلوماتية .

$P(D|Z)$: دالة الامكان الاعظم للإنموذج، وهي التي توضح العلاقة بين البيانات ومعلمة الإنموذج Z.

$p(Z|D)$:التوزيع اللاحق للمعلمة Z، ويحتوي على البيانات المحدثة الخاصة بتوزيع المعلمة Z بعد أخذ بيانات العينة الحالية (D) بنظر الاعتبار.

ويمثل التوزيع اللاحق قاعدة أرتكاز الاستدلال البيزي حيث تم إشتقاقة من قاعدة بيز، وإذا تم تجاهل $P(D)$ في الصيغة (14) بسبب كونها لا تتضمن المعلمة (Z) فيمكن كتابة التوزيع اللاحق بالشكل الأتي : [6]

$$p(Z|D) \;\alpha\; P(D|Z)\,P(Z) \quad \ldots \quad (15)$$

وهذا يعني إن التوزيع اللاحق ناتج من حاصل ضرب دالة الإمكان الأعظم $P(D|Z)$ بالمقدار $P(Z)$ والذي يمثل التوزيع المسبق للمعلمة (Z)، وفي حالة إحتواء الإنموذج على أكثر من معلمة

كما هو الحال في إنموذج SAR المستعمل في البحث فإن التوزيع اللاحق سيمثل التوزيع المشترك لكل المعالم مشروطاً على بيانات المشاهدة y، حيث يتم الحصول علية عن طريق ضرب دالة الامكان الاعظم $L(\underline{B}, \sigma^2, \lambda)$ بالتوزيعات المسبقة للمعلمات $p(\underline{B}), p(\lambda), p(\sigma^2)$، ومن خلال قاعدة بيز يمكن كتابة التوزيع اللاحق لانموذج الانحدار الذاتي المكاني SAR بالشكل الاتي:[6]

$$p(\underline{B}, \sigma^2, \lambda|D) = \frac{p(D|\underline{B}, \sigma^2, \lambda)}{p(D)}\ \pi(\underline{B}, \sigma^2)\ \pi(\lambda) \ldots . (16)$$

حيث ان $p(D)$ يتضمن اي جزء من التوزيع اللاحق الذي لا يحتوي على معلمات الانموذج ويمكن كتابة التوزيع اللاحق في الصيغة (16) بالشكل الاتي :

$$p(\underline{B}, \sigma^2, \lambda|D)\ \alpha\ \ p(D|\underline{B}, \sigma^2, \lambda)\ \pi(\underline{B}, \sigma^2)\ \pi(\lambda) \quad \ldots (17)$$

$$\pi(\underline{B}, \sigma^2) \sim NIG(c, T, a, b)$$

$$\pi(\underline{B}, \sigma^2) = \pi(\underline{B}|\sigma^2)\ \pi(\sigma^2)$$

$$= N(c, \sigma^2\ T)IG\ (a, b)$$

$$\pi(\underline{B}, \sigma^2) = \frac{b^n}{(2\pi)^{k/n}\,|T|^{1/2}\Gamma(a)}(\sigma^2)^{-(a+(k/2)+1)} \times \exp\left[-\{(B-c)'T^{-1}(B-c)+2b\}/(2\sigma^2)\right] \ldots . (18)$$

$$\pi(\sigma^2) = \frac{b^a}{\Gamma(a)}\ (\sigma^2)^{-(a+1)} \exp(-b/\sigma^2) \qquad \ldots (19)$$

$$\Gamma(a) = \int_0^\infty t^{a-1}\ e^{-t}\ dt$$

$$\sigma^2 > 0, \qquad a, b > 0$$

$$\pi(\lambda) \sim U(\omega_{min}^{-1}, \omega_{max}^{-1}) \qquad \ldots (20)$$

حيث ان :

$\pi(\underline{B}, \sigma^2)$: التوزيع الحدي (الهامشي) السابق للمعلمتين $\underline{B}, \sigma^2$.

$\pi(\sigma^2)$: التوزيع الحدي السابق (الاولي) للمعلمة σ^2.

$\pi(\lambda)$: التوزيع الحدي السابق للمعلمة λ .

ومن خلال الصيغة (18)،(19)،(20) يمكن اعادة كتابة التوزيع اللاحق لانموذج الانحدار الذاتي SAR في الصيغة (16) بالشكل الاتي:

$$p(\underline{B}, \sigma^2, \lambda|D)\ \alpha\ (\sigma^2)^{a^*+(k/2)+1}|A| \times \exp\left\{-\frac{1}{2\sigma^2}[2b^* + (B-c^*)'(T^*)^{-1}\ (B-c^*)]\right\} \ldots \qquad (17)$$

حيث ان :

$$(B-c^*)'^{(T^*)^{-1}}\ (B-c^*) \equiv (Ay-XB)'(Ay-XB) + (B-c)'T^{-1}(B-c) + 2b^*$$

$c^* = (X'X + T^{-1})^{-1}(X'Ay + T^{-1}c)$
$c = zero\ matrix$
$T^* = (X'X + T^{-1})^{-1}$
$T = I \cdot 10^3$
$a^* = a + n/2$
$b^* = b + (c'T^{-1}c + y'A'Ay - (c^*)'(T^*)^{-1}c^*)/2$
$A = I - \lambda W$

ويعد التوزيع الاحق في الصيغة (17) العنصر الرئيسي للتقدير بطريقة بيز حيث يتم من خلالة اشتقاق التوزيعات الحدية التي تحتوي على كل المعلومات الخاصة بمعلمات الانموذج$(\underline{B}, \sigma^2, \lambda)$مشروطة على بيانات المشاهدة y، وان عملية الحصول على التوزيعات الحدية الشرطية$p(\underline{B}, \sigma^2, \lambda|D)$ من التوزيع الاحق المشترك $p(\sigma^2|y), p(\underline{B}|y), p(\lambda|y)$ عملية صعبة وذلك لكون عملية الاشتقاق معقدة وشبة مستحيلة باستعمال الورقة والقلم، وهذه المشكلة سببت صعوبة للطريقة البيزية في ايجاد التوزيع الخاص بكل معلمة من معلمات الانموذج على حدة ومن اجل حل هذه المشكلة تم اللجوء الى استعمال سلاسل ماركوف مونت كارلو (MCMC)، وهي عبارة عن طرائق توليد مرنة ومن اشهرها طريقة كابز Gibbs وطريقة Metropolis-Hastings، ويتلخص مبدأ عملها بتحليل التوزيع اللاحق لمجموعة من التوزيعات الشرطية الكاملة لكل معلمة من معلمات انموذج SAR وهي:

$$p(\sigma^2|\underline{B}, \lambda), p(\underline{B}|\sigma^2, \lambda), (\lambda|\underline{B}, \sigma^2) \quad \dots (18)$$

ثم القيام بعملية سحب عينات عشوائية من هذه التوزيعات الشرطية الكاملة مما يمكن الباحث من استعمال هذه العينة من السحبات في الحصول على التقديرات البيزية الصحيحة لمعلمات الانموذج $(\underline{B}, \sigma^2, \lambda)$، حيث تستعمل طريقة كابز Gibbs مع التوزيعات الشرطية التي تمتلك صيغة معروفة فقط بينما طريقة Metropolis-Hastings يمكن استعمالها مع التوزيعات الشرطية الكاملة التي لا تمتلك صيغة معروفة.[6]

1-2-5 التوزيعات الشرطية الكاملة

ان عملية الحصول على التوزيعات الشرطية الكاملة هي اولى متطلبات معاينة كابز حيث تم اشتقاق التوزيعات الشرطية الكاملة لانموذج SAR من التوزيع اللاحق في الصيغة (17)، ومن الجدير بالذكر ان هنالك ثلاثة افكار رياضية تم تطبيقها والالتزام بها عند الاشتقاق وهي : [4]

اولا: اهمال كل الثوابت الغير ضرورية.

ثانيا: جمع الحدود الخاصة فقط بالتوزيع الشرطي عند القيام بعملية اشتقاق التوزيع الشرطي المطلوب .

ثالثا: إكمال المربع.

ان التوزيعات الشرطية الكاملة لانموذج SAR التي تم اشتقاقها كما يلي : [6]

- **التوزيع الشرطي الكامل للمعلمة σ^2**، $p(\sigma^2|\underline{B}, \lambda)$: يملك توزيع كاما المعكوس (IG) المبين كما يلي:

$$p(\sigma^2 \mid \underline{B}, \lambda) \sim IG(a^*, b^*)$$

$$p(\sigma^2 \mid \underline{B}, \lambda) = \frac{b^{*a^*}}{\Gamma(a^*)} \ \sigma^{2^{-(a^*+1)}} . e^{-\ b^* \backslash \sigma^2} \quad \ldots\ldots (19)$$

$$a^* = a + n/2$$

$$b^* = b + (A\underline{y} - X\underline{B})'(A\underline{y} - X\underline{B})/2$$

$$\pi\sigma^2 \sim IG(a, b)$$

$$a = \frac{n(\bar{x}^2 - 2\sum x_i^2)}{n\bar{x}^2 - \sum x_i^2}$$

$$b = (a-1)\bar{x}$$

- **التوزيع الشرطي الكامل للمعلمة $\underline{\mathbf{B}}$**، $p(\underline{B}|\sigma^2, \lambda)$: يملك التوزيع الطبيعي المتعدد المتغيرات (N) كما يلي:

$$p(\underline{B}|\sigma^2, \lambda) \sim N(C^*, \sigma^2 T^*)$$

$$p(\underline{B}|\sigma^2, \lambda) = \frac{1}{(2\pi)^{-p/2}|\sigma^2 T^*|^{1/2}} \exp\left\{-\frac{1}{2}(\underline{B} - C^*)' \ (\sigma^2 T^*)^{-1} \ (\underline{B} - C^*)\right\} \ldots (20)$$

$$c^* = (X'X + T^{-1})^{-1}(X'Ay + T^{-1}c)$$

$$c = \text{zero matrix}$$

$$T^* = (X'X + T^{-1})^{-1}$$

$$T = I \,.\, 10^3$$

- **التوزيع الشرطي الكامل للمعلمة λ**، $p(\lambda|\underline{B}, \sigma^2)$: يملك توزيع بصيغة غير معروفة يتم الحصول علية من الصيغة التالية:

$$p(\lambda|\underline{B}, \sigma^2) \alpha \frac{p(\lambda, \underline{B}, \sigma^2|D)}{p(\underline{B}, \sigma^2|D)} \quad \ldots (21)$$

$$p(\lambda|\underline{B}, \sigma^2) \alpha \ |I - \lambda W| \exp(-\tfrac{1}{2\sigma^2}(Ay - X\underline{B})'(Ay - X\underline{B})) \quad \ldots (22)$$

$$p(\underline{B}, \sigma^2|D) \sim NIG(c, T, a, b)$$

2-2-5 معاينة Metropolis-Hastings (M-H)

ان معاينة (*M-H*) عباره عن خوارزمية رفض وقبول حيث يتم فيها اقتراح توزيع معين نولد منه قيمة مرشحة للمعلمة □ والتي سنسميها *□ ثم يتم المقارنة او تقييم القيمة المرشحة مع القيمة الحالية للمعلمة التي نحصل عليها من الصيغة (*22*) والتي سنسميها ^{Q}□، ومن اجل حساب احتمال القبول للقيم المرشحة بعد المقارنة نستعمل الصيغة التالية: (6)

$$\psi H\left(\lambda^{Q},\lambda^{*}\right)=\min\left[1.\frac{p(\lambda^{*}|B,\sigma^{2})}{p\left(\lambda^{Q}\middle|B,\sigma^{2}\right)}\right] \qquad \dots(23)$$

وغالبا ما يتم ضبط احتمالية القبول ضمن المدى (40-60) وذلك لكي يتم سحب نصف العينات المرشحة ، ولقد تم اختيار التوزيع الطبيعي كتوزيع مقترح جنبا الى جنب مع اجراءات السير العشوائي التي اقترحها كل من *Holloway, Shankara* و *Rahman* عام 2002 من اجل الحصول على القيمة المرشحة للمعلمة □، وان عملية المعاينة ومعلمة الضبط *Q* يمكن توضيحها من خلال الصيغة التالية:

$$\lambda^{*}=\lambda^{Q}+\mathrm{Q.N}(0,1) \qquad \dots. \quad (24)$$

حيث ان الصيغة(*24*) توضح سبب تسمية اجراء توليد البيانات بالسير العشوائي، وان الهدف من ضبط وتقييم العينات التي تسحب من التوزيع الطبيعي المقترح هو للتاكد من ان عملية سحب العينات العشوائية تتم خلال التوزيع الشرطي باكملة.

3-2-5 معاينة كابز

تمثل معاينة كابز *Gibbs* احدى طرائق توليد البيانات وهي حالة خاصة من معاينة *(M_H)* وذلك لكونها تتعامل مع التوزيعات الشرطية الكاملة ذات الصيغ المعروفة فقط ، حيث يتم سحب عينات عشوائية من التوزيعات الشرطية الكاملة الخاصة بمعالم انموذج *SAR* ويتم اولا اختيار قيم افتراضية لمعلمات الانموذج $\sigma^{2}{}_{(0)}$, $B_{(0)},\lambda_{(0)}$ وبعدها يبدأ اخذ عينات بشكل متسلسل من التوزيعات الشرطية الكاملة الثلاث وكالاتي : [6]

1- تسحب عينة من التوزيع الشرطي الكامل للمعلمة *B*، حيث تمتلك $p\left(\underline{B}\mid\ \sigma^{2}{}_{(0)},\lambda_{(0)}\right)$ التوزيع الطبيعي المتعدد $N(C^{*},\sigma^{2}T^{*})$ المبين في الصيغة (*20*) ونرمز للمعلمة المسحوبة بالرمز $\underline{B}_{(1)}$، ويتم استعمال هذة المعلمة بدل من المعلمة $\underline{B}_{(0)}$ بالخطوة التالية.

2- تسحب عينة من التوزيع الشرطي الكامل للمعلمة σ^{2}، حيث تمتلك $p\left(\sigma^{2}\mid\ \underline{B}_{(1)},\lambda_{(0)}\right)$ توزيع كاما المعكوس $IG(a^{*},b^{*})$ المبين في الصيغة (*19*) ونرمز للمعلمة المسحوبة بالرمز $\sigma^{2}{}_{(1)}$ ويتم استعمال هذة المعلمة بدل من المعلمة $\sigma^{2}{}_{(0)}$ بالخطوة التالية.

3- تسحب عينة من التوزيع الشرطي الكامل للمعلمة λ، حيث تمتلك $p\left(\square\middle|\underline{B}_{(1)},\sigma^{2}{}_{(1)}\right)$ توزيع غير معروفة الصيغة و المبين في الصيغة (*22*) ولهذا تم سحب العينة باستعمال طريقة معاينة (*M-H*) ونرمز للعينة الجديدة بالرمز $\lambda_{(1)}$ ونعوضها بالخطوه رقم (1)

حيث ان الخطوات المتسلسلة من 1 الى 3 تمثل مسار واحد او دورة واحدة وفي كل دوره نقوم بسحب عينات جديدة للمعلمات، وكمثال توضيحي من الممكن ان نسحب 7500 مرة ونتجاهل اول 2500 سحبة من العينات ونستعمل الباقي للتوصل للتوزيع الاحق والاستدلالات الخاصة به.

6- معيار متوسط الخطأ المطلق النسبي

يمثل معيار (*MAPE*) الوسيلة التي سيتم بها المقارنة بين طرائق تقدير معلمات الانموذج المكاني ويتم حسابة من خلال الصيغة الرياضية التالية : [9]

$$\text{MAPE} = \frac{1}{\text{n}}\sum_{\text{i}=1}^{\text{n}} \left|\frac{\text{y}_\text{i} - \hat{\text{y}}_\text{i}}{\text{yi}}\right| \qquad \dots (25)$$

حيث ان طريقة التقدير التي تحصل على القيمة الاصغر ستكون هي الطريقة الافضل في تقدير معلمات الانموذج .

7- أختبار معامل موران

ان معامل موران عبارة عن اداة لقياس الاعتماد المكاني في بيانات الظاهرة المراد دراستها ومدى التشابه فيما بينها ويمكن ان نرمز لة بالرمز I_{MC} ويكون مناضر لاختبار ديربن واتسون (*Durbin Watson*) في بيانات السلاسل الزمنية، وتتراوح قيمة معامل موران بين (1+, 1-) حيث نلاحظ انه اذا اقتربت قيمة معامل موران من (1+) يكون نمط الانتشار للبيانات متقارب اما اذا كانت قيمة معامل موران قريبة من (1-) يكون انتشار البيانات متباعد فيما بينها اما اذا كانت قيمة معامل موران تقترب من (0) نلاحظ ان البيانات عشوائية الانتشار المكاني،(7) وان صيغة معامل موران لمصفوفة الاوزان المستعملة من النوع التي يكون فيها مجموع الصف يساوي الواحد الصحيح كما يلي: [3]

$$\text{I}_{\text{MC}} = \frac{(u'w\,u)}{(u'u)} \dots\dots\dots\dots\dots\dots\dots\dots (26)$$

حيث ان :

W: مصفوفة الاوزان(التجاورات) ذو الأبعاد $n \times n$.

u: متجة الاخطاء (البواقي) ذو البعد $n \times 1$.

ولقد تم التوصل للتوزيع التقاربي الخاص باحصاءة موران من قبل الباحثان *Ord* و *Cliff* في عام (1972) حيث انه يتوافق مع التوزيع الطبيعي القياسي ويمكن اجراء اختبار موران (*Z*) من خلال الصيغة التالية : (3)

$$Z_{I_{MC}} = \frac{I_{MC} - E(I_{MC})}{\sqrt{V(I)}} \dots\dots\dots\dots\dots (27)$$

$$E(I_{MC}) = E(I_{MC}) = \frac{tr\ (M\ \ W)}{n-k}$$

$$V(I_{MC}) = \frac{tr(MWMW') + tr(MW)^2 + \left(tr(MW)\right)^2}{(n-k)(n-k+1)} - \left(E(I_{MC})\right)^2$$

$$M = (I_n - X(X'X)^{-1}X')$$

حيث ان :

tr: مجموع العناصر للقطر الرئيسي.

K: يمثل عدد المتغيرات التفسيرية.

M: تمثل مصفوفة صماء وتكون مربعة و متماثلة.

ومن اجل اختبار وجود الاعتماد المكاني من عدمة نستخدم الفرضية التالية :

$H_o: \lambda = 0$ لا يوجد أعتماد مكاني

$H_1: At\ least\ one\ of\ \lambda \neq 0$ يوجد أعتماد مكاني

ومن خلال الصيغة *(27)* اذا كانت قيمة $Z_{I_{MC}}$ المحسوبة اكبر من قيمة Z الجدولية عند مستوى دلالة معينة نقبل الفرضية البديلة وهذا يعني وجود اعتماد مكاني بين مفردات الظاهرة المدروسة، اما ذا كانت $Z_{I_{MC}}$ المحسوبة اصغر من قيمة Z الجدولية فإن ذلك يؤدي الى قبول الفرضية العدم بمعنى انه لا يوجد اعتماد مكاني بين مفردات الظاهرة المدروسه.

8- الجانب التطبيقي

لقد تم تطبيق طريقتي التقدير لانموذج الانحدار الذاتي المكاني التي تم ذكرها على بيانات حقيقية متمثلة بمرض السرطان حيث ان المتغير المعتمدy يمثل عدد مرضى السرطان اذ تم اخذ عدة انواع من السرطان التي تمثل النوع الاكثر انتشارا في المحافظة في كل محافظة من محافظات العراق، بينما تمثلت المتغيرات التوضيحية كما يلي:

X_1 : معدل العمر

X_2 : معدل حجم الورم

X_3 *:* عدد المناطق الملوثة باليورانيوم في المحافظة

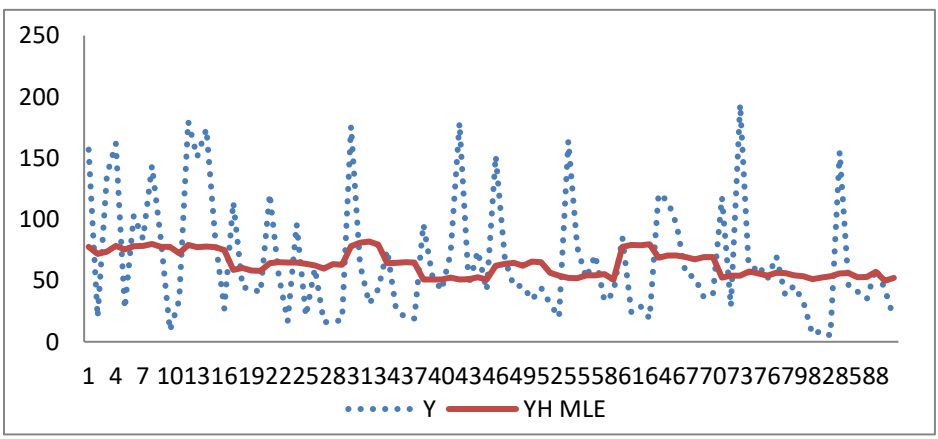
اما مصفوفة الاوزان المكانية التي تم اعتمادها في البحث فقد تم بناءها بالاعتماد على معيار روك للتجاور بين محافظات العراق ومن ثم تم تحويلها الى مصفوفة اوزان مكانية معدلة بالاعتماد على الصيغة في الرقم (2)، وفيما يلي نتائج تطبيق كل من طريقة الامكان الاعظم *MLE* وطريقة بيز في تقدير انموذج الانحدار الذاتي المكاني:

8-1 تقدير انموذج الانحدار الذاتي المكاني SAR باستعمال طريقة الامكان الاعظم MLE

باستعمال برنامج (*Matlab*) تم تقدير قيمة معلمة الاعتماد المكاني (λ) للانموذج من خلال الصيغة (13) و المساوية الى (0.287) ومن خلالها تم تقدير انموذج *SAR* باستعمال طريقة *MLE* والحصول على قيم المتغير المعتمد التقديرية $\hat{y}$ المبينة في الجدول (1) الذي يضم القيم الحقيقية والتقديرية للمتغير المعتمد نفسة y الذي يمثل عدد مرضى السرطان، وكذلك تم تمثيل القيم الحقيقية والتقديرية للمتغير y في الجدول (1) بالرسم لبياني في الشكل (1)، وبعد ان تم الحصول على القيم التقديرية $\hat{y}$ للمتغير المعتمد تم حساب قيمة معيار متوسط الخطأ النسبي المطلق(*MAPE*) لانموذج الانحدار الذاتي المكاني *SAR* والتي تساوي (*1.0472*). اما نتائج اختبار موران Z_M عند استعمال مصفوفة الاوزان المكانية المعدلة كانت *(* 4.2832) وعند المقارنة بين القيمة المحسوبة وبين القيمة الجدولية التي تساوي(*1.96*) عند مستوى معنوية (*0.05*) نلاحظ ان القيمة الجدولية اصغر من القيمة المحسوبة وهذا دليل على وجود الاعتماد المكاني في البيانات .

جدول (1) القيم الحقيقية والتقديرية للمتغير المعتمد y باستعمال طريقة الامكان الاعظم MLE

T	y	$\hat{y}$	T	y	$\hat{y}$	T	y	$\hat{y}$	T	y	$\hat{y}$
1	157	77	24	96	65	47	67	64	70	39	69
2	20	72	25	21	64	48	45	64	71	118	52
3	136	74	26	60	62	49	46	62	72	29	54
4	162	78	27	16	60	50	34	65	73	194	54
5	27	75	28	16	63	51	44	65	74	54	57
6	103	78	29	18	63	52	31	56	75	63	56
7	84	78	30	176	78	53	20	54	76	51	54
8	144	80	31	63	81	54	164	52	77	70	56
9	83	77	32	31	82	55	77	52	78	38	56
10	9	77	33	43	79	56	52	54	79	46	54
11	34	72	34	75	64	57	71	54	80	32	53
12	180	79	35	26	64	58	33	55	81	6	51
13	151	77	36	20	65	59	40	51	82	9	53
14	174	78	37	19	65	60	85	77	83	5	54
15	83	77	38	96	51	61	24	79	84	155	56
16	27	75	39	54	51	62	29	79	85	45	56
17	114	59	40	42	51	63	19	80	86	41	53
18	44	60	41	72	52	64	121	69	87	35	53
19	44	58	42	179	51	65	115	70	88	60	57
20	40	58	43	46	51	66	95	71	89	43	50
21	121	64	44	74	53	67	55	69	90	19	52
22	54	65	45	42	51	68	53	67			
23	17	65	46	152	62	69	36	69			



شكل (1) الرسم البياني للقيم الحقيقية للمتغير المعتمد y والقيم التقديرية $\hat{y}$ باستعمال طريقة MLE

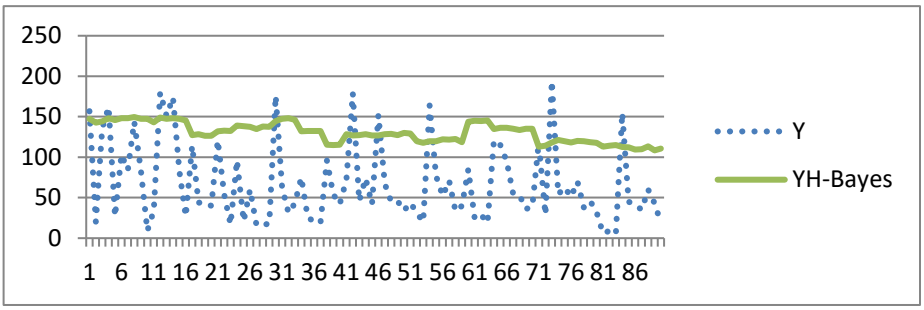
2-8 تقدير انموذج الانحدار الذاتي المكاني SAR باستعمال طريقة بيز Bayes

باستعمال برنامج (*Matlab*) تم تقدير انموذج *SAR* باستعمال طريقة بيز والحصول على قيم المتغير المعتمد التقديرية $\hat{y}$ المبينة في الجدول (2) الذي يضم القيم الحقيقية والتقديرية للمتغير المعتمد نفسة y الذي يمثل عدد مرضى السرطان، وكذلك تم تمثيل القيم الحقيقية والتقديرية للمتغير y في الجدول (2) بالرسم لبياني في الشكل (2)، وبعد ان تم الحصول على القيم التقديرية $\hat{y}$ للمتغير المعتمد تم حساب قيمة معيار متوسط الخطأ النسبي المطلق(*MAPE*) لانموذج الانحدار الذاتي المكاني *SAR* والتي تساوي (*2.6829*).

اما نتائج اختبار موران Z_M عند استعمال مصفوفة الاوزان المكانية المعدلة كانت (*4.0731*) وعند المقارنة بين القيمة المحسوبة وبين القيمة الجدولية التي تساوي(*1.96*) عند مستوى معنوية (*0.05*) نلاحظ ان القيمة الجدولية اصغر من القيمة المحسوبة وهذا دليل على وجود الاعتماد المكاني في البيانات .

جدول (2) القيم الحقيقية والتقديرية للمتغير المعتمد y باستعمال طريقة بيز Bayes

T	y	$\hat{y}$	T	y	$\hat{y}$	T	y	$\hat{y}$	T	y	$\hat{y}$
1	157	148	24	96	139	47	67	128	70	39	135
2	20	143	25	21	138	48	45	129	71	118	113
3	136	144	26	60	137	49	46	127	72	29	114
4	162	148	27	16	135	50	34	130	73	194	118
5	27	146	28	16	138	51	44	129	74	54	121
6	103	148	29	18	137	52	31	120	75	63	120
7	84	148	30	176	145	53	20	118	76	51	118
8	144	150	31	63	147	54	164	120	77	70	120
9	83	148	32	31	148	55	77	120	78	38	120
10	9	148	33	43	146	56	52	122	79	46	118
11	34	143	34	75	132	57	71	122	80	32	118
12	180	149	35	26	133	58	33	122	81	6	113
13	151	147	36	20	132	59	40	119	82	9	114
14	174	148	37	19	132	60	85	144	83	5	115
15	83	147	38	96	115	61	24	145	84	155	113
16	27	145	39	54	115	62	29	145	85	45	112
17	114	127	40	42	116	63	19	145	86	41	110
18	44	129	41	72	128	64	121	135	87	35	110
19	44	127	42	179	127	65	115	136	88	60	114
20	40	127	43	46	127	66	95	136	89	43	108
21	121	132	44	74	129	67	55	135	90	19	110
22	54	133	45	42	127	68	53	134			
23	17	132	46	152	127	69	36	135			



شكل (2) الرسم البياني للقيم الحقيقية للمتغير المعتمد y والقيم التقديرية $\hat{y}$ باستعمال طريقة بيز Bayes

3-8 تحليل نتائج التطبيقي

نلاحظ من خلال الجدول رقم (1) والجدول رقم (2) ان القيم التقديرية للمتغير المعتمد باستعمال طريقة *MLE* اقرب الى القيم الحقيقية للمتغير المعتمد من القيم التقديرية للمتغير المعتمد باستعمال طريقة بيز حيث يبين هذا ايضا من خلال الشكل التوضيحي رقم (1) ورقم (2) حيث نلاحظ ان تقديرات بيز بعيدة عن القيم الحقيقية ويوجد مسافة بينهما كبيرة بينما هذة المسافة تبدو اقل عند مقدرات طريقة الامكان الاعظم، وهذا ما جعل قيمة معيار متوسط الخطأ المطلق النسبي لطريقة الامكان الاعظم اصغر من طريقة بيز في التقدير.

9- الاستنتاجات والتوصيات

بعد ان تم تحليل نتائج الجانب التطبيقي سيتم ذكر اهم الاستنتاجات والتوصيات التي تم التوصل اليها فيما يلي :

1-9 الاستنتاجات

1. اظهر اختبار موران ان بيانات مرض السرطان تعاني من الاعتمادية المكانية عند استعمال مصفوفة الاوزان المكانية المعدلة.
2. بالاعتماد على نتائج معيار المقارنة (MAPE) وعند المقارنة بين طريقة الامكان الاعظم MLE وطريقة بيز Bayes في تقدير انموذج الانحدار الذاتي المكاني SAR باستعمال مصفوفة الاوزان المكانية المعدلة تم التوصل الى ان طريقة الامكان الاعظم افضل من طريقة بيز Bayes في التقدير كونها حصلت على اقل قيمة لمعيار (MAPE).
3. ان طريقة بيز في التقدير تعطي نتائج تقديرية بعيدة عن القيم الحقيقية عند استعمالها على بيانات مرضى السرطان.

2-9 التوصيات

1- استعمال طريقة الامكان الاعظم في تقدير انموذج الانحدار الذاتي المكاني.
2- استعمال معيار روك للتجاور كونه اظهر كفاءته.

3- كدراسات مستقبليه نوصى بالمقارنة بين مصفوفة التجاور المكانية المعدلة ومصفوفة المسافة التي تعتمد على احداثيات مواقع الاشخاص وليس مركز الوحدات المكانية .

4- كدراسات مستقبلية نوصى بدراسة مكانية للعوامل التي تؤدي الى الاصابة بمرض السرطان وتاثيرها على عدد الاصابات في كل منطقة .

المصادر

1- Anselin, L. (1988). "Spatial Econometrics: Methods and Models". Kluwer Academic Publishers Dordrecht, the Netherlands.

2- Anselin, L. (1992), "Spatial Data Analysis with GIS: An Introduction to Appalachian the Social Sciences", National Center for Geographic Information and Analysis, University of California, Santa Barbara.

3- Anselin, L. and Bera, A. (1998). "Spatial dependence in linear regression models with an introduction to spatial econometrics ", Handbook of Applied Economic Statistics. Marcel Dekker, New York.

4- Lacombe, Donald J. (2008). " An Introduction to Bayesian Inference in Spatial Econometrics". Department of Economics Ohio University.

5- Lesage, J. P., (1999), "The Theory and Practice of Spatial Econometrics", Department of Economics University of Toledo.

6- Lesage, J. P, (2009). "Introduction to Spatial Econometrics". Statistics, a series of textbooks and monographs.

7- Michiel, D., Pooter, D., Segers, R. Herman, K. Dijk, v., (2006). " Gibbs Sampling in Econometric Practice", Econometric Institute and Tinbergen Institute Erasmus University Rotterdam, The Netherlands.

8- Omar, Dalia Badeea. (2020), "Estimating Parameters of Some Spatial Regression Models with Experimental and Applied Study", College of Administration and Economics at Kirkuk University.

9- Rahim, Sh. A., Hussein, M. F.,(2020), "A study on atmospheric pressure in krg using spatial regression (SAR and SEM)" A Scientific Quarterly Refereed Journal Issued by Lebanese French University.