استخدام طريقة M الحصينة في تقدير انموذج انحدار بواسون لأعداد العاطلين عن العمل في العراق

محمد غازي جواد / باحث mohamed2722471@gmail.com
أ.م اسيل عبد الرزاق رشيد aseelstat@uomustansiriyah.edu.iq

P: ISSN : 1813-6729
E : ISSN : 2707-1359
https://doi.org/10.31272/jae.i139.1091
تأريخ أستلام البحث : 2023/2/27 مقبول للنشر بتأريخ :2023/3/13

المستلخص :
يعد انموذج انحدار بواسون احد نماذج العد المهمة حيث استخدم لنمذجة العديد من الظواهر السكانية والأقتصادية والصحية . وبسبب وجود القيم الشاذة في البيانات تم استعمال طريقة M الحصينة لتقدير معالم هذا الانموذج وذلك بالأعتماد على بيانات للعاطلين عن العمل في العراق كمتغير استجابة للفترة (2004-2021) ومعدل النمو السكاني السنوي كمتغير توضيحي اول ومعدل التضخم السنوي كمتغير توضيحي ثاني وضمن نفس الفترة , حيث افرزت النتائج ارتباط هذه المتغيرات بعلاقة معنوية ذات اثر عكسي مع متغير الاستجابة .

الكلمات المفتاحية : انحدار بواسون , طريقة مقدرات M , اعداد العاطلين

مجلة الادارة والاقتصاد
مجلد 48 العدد 139 / أيلول / 2023
الصفحات : 213 - 220

*** بحث مستل من رسالة ماجستير .**

1- المقدمة :

تحليل الانحدار هو اداة احصائية تقوم ببناء انموذج احصائي وذلك لتقدير العلاقة بين متغير واحد يدعى المتغير المعتمد ومتغير اخر او عدة متغيرات اخرى تدعى المتغيرات التوضيحية (التفسيرية) بحيث ينتج معادلة احصائية توضح العلاقة بين المتغيرات . يعد انموذج انحدار بواسون احد نماذج الانحدار غير الخطية , وكذلك يعد احد اهم نماذج الانحدار اللوغارتمية الخطية وهو الاداة التي يتم من خلالها نمذجة المتغير المعتمد عندما تكون قيم ذلك المتغير على شكل قيم قابلة للعد , وفي حالة وجود القيم الشاذة في البيانات يتعذر علينا تحقيق الفروض الواجب توفرها لتقدير معالم انموذج انحدار بواسون بالطرائق الاعتيادية وبالتالي نتجه للطرائق الحصينة ومن ابرزها طريقة M الحصينة لتقدير معالم هذا الانموذج بأستخدام بيانات العاطلين عن العمل في العراق والعوامل المؤثرة عليها .

2- انموذج انحدار بواسون Poisson Regression Model [4][5]

يعرف انموذج انحدار بواسون بأنه احد انواع نماذج الانحدار اللاخطية والمستخدم في نمذجة البيانات المعدودة (Countable Data) ويستعمل في اغلب الاحيان في تحليل الاحداث النادرة, ويمكن كتابة انموذج انحدار بواسون على النحو التالي:

$$Y_i = \mathrm{E}(Y_i) + u_i = \lambda_i + u_i \quad \ldots\ldots\ldots (1) \qquad i = 1, 2, \ldots n$$ حيث ان

وان متغير الاستجابة **Yi** يتبع توزيع بواسون بمتوسط وتباين يساوي **(λi)**

$$\lambda_i = E(Y_i|X_i) = e^{\underline{x}_i \underline{B}} \quad \ldots\ldots\ldots(2)$$

وعلى وفق هذه الصيغة فان المتغيرات التوضيحية تقترح اي قيمة حقيقية لمتوسط متغير الاستجابة (بمعنى ان المتوسط اما يكون موجب او سالب) , وهذا يتعارض مع خاصية معلمة توزيع بواسون وهي ان قيمة **λ** موجبة فقط , لذلك تم اقتراح نمذجة علاقة خطية لوغارتم المتوسط **(Log λ)** والمتغيرات التوضيحية **(Xi)** اي اعتماد النموذج الخطي اللوغارتمي للتخلص من القيم السالبة لمتوسط متغير الاستجابة **(Y)**

$$\log (E(Yi)) = \text{Log } (\lambda i) = \underline{X}_i' \underline{B} \quad \ldots\ldots\ldots(3)$$

$\underline{B}$: يمثل متجه المعلمات ذو الدرجة **(K+1)*1** المطلوب تقديرها

$\underline{X}_i'$**:** يمثل متجه المتغيرات التوضيحية ذو الدرجة **1*(k+1)**

Yi: يمثل متغير الاستجابة حيث **(i=1,2,….,n)** و **n** هي عدد مشاهدات العينة

وبأخذ **exp** لطرفي العلاقة رقم (3) نحصل

$$\lambda i = \exp (\underline{X}_i' \underline{B}) \quad \ldots\ldots\ldots(4)$$

وبالتالي سوف نحصل على الدالة الاحتمالية الشرطية التالية :

$$f(yi,xi) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} = \frac{e^{-e^{(\underline{x}_i'\underline{B})}} e^{yi\underline{x}_i'\underline{B}}}{y_i!} \qquad y=0,1,2,\ldots\ldots \quad \ldots.(5)$$

ان انموذج انحدار بواسون يحقق دالة الوسط الحسابي الشرطية الاسية (الخطية اللوغارتمية) لان الانموذج يتحول بأخذ اللوغارتم الطبيعي الى انموذج خطي .

وعليه فأن الصيغة الخطية ل λ_i تصبح $\text{Log } \lambda_i = \underline{x}_i' \underline{B}$

3- طرائق التقدير الحصينة Robust Estimation Methods

اول من اطلق مصطلح المقدرات الحصينة هو الباحث **Box** في عام 1953 ثم توالت الدراسات بخصوصها من قبل عدة باحثين , ان كلمة الحصانة تطلق على المقدرات التي لاتتأثر بوجود مخالفة في احدى فرضيات التوزيع الطبيعي , وان الهدف الاساسي من ايجاد طرائق التقدير الحصين هو لتقليل تأثير القيم الشاذة على المقدر .[1]

ان اتباع الطرق الكلاسيكية في تقدير المعالم سواء كان الانموذج خطيا او غير خطي ليس اكثر امانا من تطبيق الطرق الحصينة وذلك بسبب ان الفروض الواجب توفرها لتطبيق الطرق الكلاسيكية ليست بالسهلة مثل عدم وجود القيم الشاذة او اتباع الخطا العشوائي توزيعا غير التوزيع الذي يناسب الطريقة المعتمدة في التقدير .[3]

1-3 طريقة M-estimator [2]

هي مقدر انحدار حصين امتداد لطريقة (**MLE**) وتستعمل على نطاق واسع وتسمى احيانا تقديرات (**Huber**) وتعتبر تقديرات قوية ضد القيم الشاذة في المتجه المعتمد (**Y**) ولكنها ليست قوية ضد القيم الشاذة في المتجه (**X**) وتكون مقدرات هذه الطريقة غير متحيزة وتمتلك اقل تباين ويكون هدف تقدير **M** هو تقليل البواقي للاخطاء .

$$\widetilde{B}_{M} = \min p(y_i - \exp(\acute{X}B) \quad \ldots\ldots\ldots (6)$$

وبتقليل دالة البواقي الى اقل ما يمكن

$$\widehat{B}_M = \min \sum_{i=1}^{n} P(u_i) = \min \sum_{i=1}^{n} P\left(\frac{e_i}{\widehat{\sigma}}\right) = min \sum_{i=1}^{n} P\left[\frac{y_i - \exp(\acute{X}B)}{\widehat{\sigma}}\right] \quad \ldots\ldots (7)$$

إذ ان e_i : تمثل البواقي

حيث يقدر التباين وفق الصيغة الاتية :

$$\widehat{\sigma} = \frac{MAD}{0.6745} \quad \ldots\ldots\ldots (8)$$

$$\text{MAD} = \text{median } |e_i - median(e_i)| \quad \ldots\ldots\ldots (9)$$

إذ ان **MAD** : تمثل الوسيط لمطلق الفروقات بين البواقي ووسيطها

عند اشتقاق المعادلة (7) بالنسبة الى **B** نحصل :

$$\sum_{i=1}^{n} x_i \varphi[\frac{y_i - \exp(XB)}{\widehat{\sigma}_M}] = 0 \quad \ldots\ldots\ldots (10)$$

إذ ان $\varphi(.)$: تمثل مشتقة دالة **P(.)** وتسمى ايضا بدالة التأثير **Influence function**

ويتم حساب الاوزان للقيم

$$W_i = \frac{\varphi[\frac{y_i - \exp(XB)}{\widehat{\sigma}_M}]}{\frac{y_i - \exp(XB)}{\widehat{\sigma}_M}} \quad \ldots\ldots\ldots (11)$$

وبالتالي يمكن صياغة المعادلة المشار إليها بالاعتماد على دالة توكي وكما يلي :

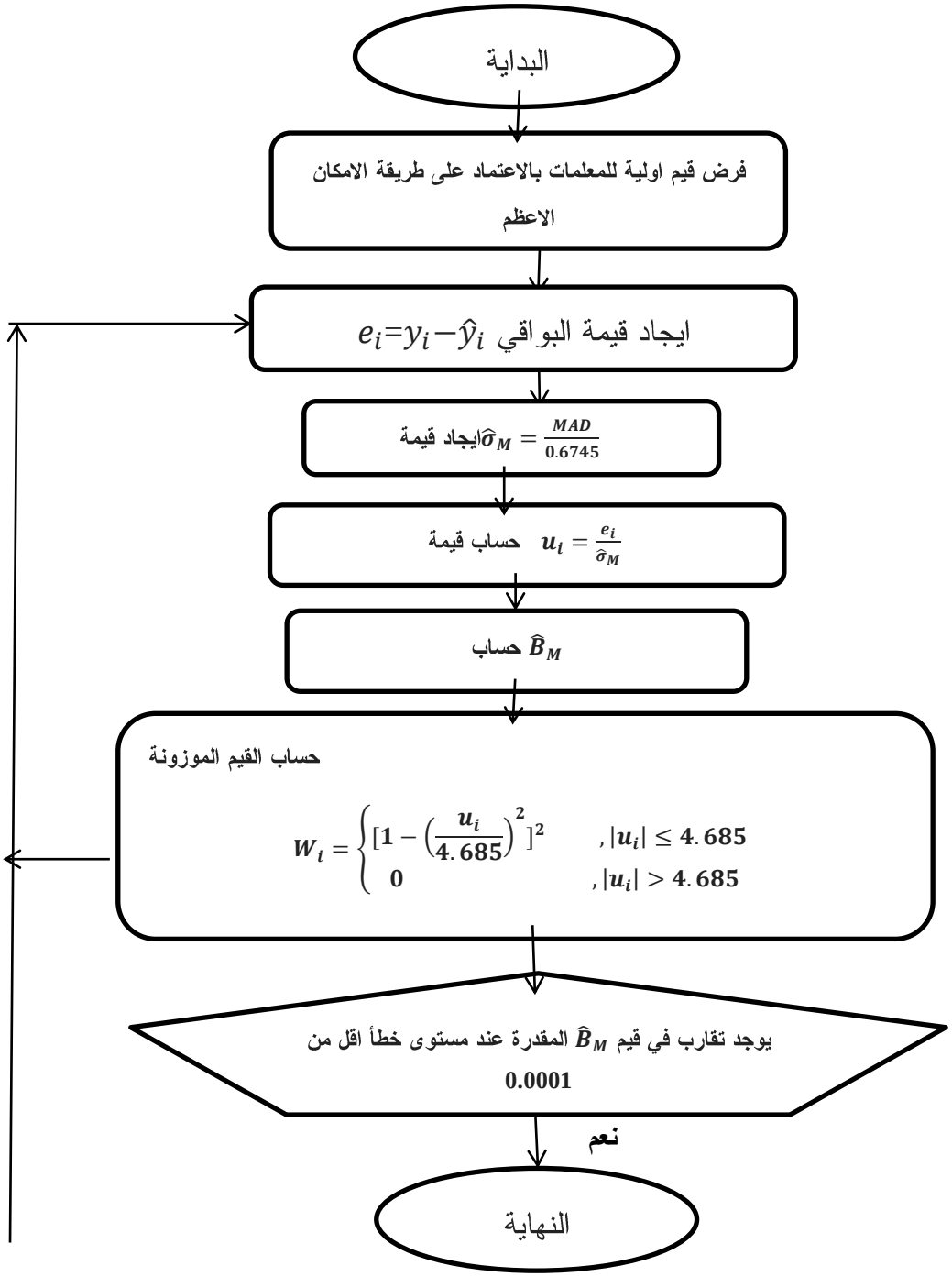
$$W_i = \begin{cases} \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 & , |u_i| \le c \\ 0 & , |u_i| > c \end{cases} \quad \ldots\ldots\ldots (12)$$

حيث ان قيمة **(c=4.685) حسب دالة Tukey**

$$\text{Where } u_i = \frac{e_i}{\widehat{\sigma}}$$

$$\sum_{i=1}^{n} x_i W_i(y_i - \exp(XB)) = 0 \quad \ldots\ldots\ldots (13)$$

خوارزمية طريقة M تكون كالتالي :



4- وصف بيانات الدراسة

يعتبر عدد العاطلين عن العمل الشغل الشاغل لكثير من البلدان سواء النامية او المتقدمة التي تستهدف خفض هذا العدد الى ادنى حد ممكن من دون التأثير على عجلة النمو الأقتصادي .

ان اهمال الجانب الزراعي والصناعي في بعض الدول الفقيرة والدول غير المنتجة والدول ذات الدخل الريعي يُعجل من زيادة عدد العاطلين فيها بشكل كبير جدا ويصعب على حكومات هذه الدول خفضه لاحقا .

قامت دائرة العمل والتدريب المهني لوزارة العمل والشؤون الأجتماعية في العراق بعمل قاعدة بيانات للعاطلين عن العمل تتضمن مختلف الأختصاصات والشهادات في محاولة منها لرفد القطاع الحكومي والقطاع الخاص بالعمالة المناسبة لكن ظهر لها بعض المعوقات مثل زيادة عدد الموظفين في الدوائر الحكومية لدرجة الفيض فيها , وازدياد اعداد العمالة الوافدة من الخارج للعمل في القطاع الخاص ومنافستها للعمالة الوطنية .

في هذا البحث تم اختيار عدد العاطلين عن العمل في العراق كمتغير معتمد , حيث اخذت بيانات سنوية لمدة 18 سنة (منذ عام 2004 ولغاية عام 2021) من وزارة العمل والشؤون الأجتماعية / دائرة العمل والتدريب المهني . يعتمد عدد العاطلين على متغيرات كثيرة بينها اساسي وبينها ثانوي بينها متوفر وبينها غير متوفر , **حيث اعتمدنا في هذه الدراسة على متغيرين توضيحين هما :-**

1- **المتغير التوضيحي الأول :** يمثل معدل النمو السكاني السنوي في العراق بالأستناد على بيانات وزارة التخطيط / الجهاز المركزي للأحصاء ولنفس الفترة الزمنية , حيث يرتبط معدل النمو السكاني بعلاقة عكسية مع عدد العاطلين عن العمل , ويحسب كما يلي:

$$\text{معدل النمو السكاني السنوي في العراق} = \left[1 - \frac{\text{عدد سكان العراق في السنة المحددة}}{\text{عدد سكان العراق في السنة السابقة لها}}\right] * 100$$

2- **المتغير التوضيحي الثاني :** يمثل معدل التضخم السنوي في العراق بالأستناد الى بيانات تقرير البنك المركزي ولنفس الفترة الزمنية , حيث يرتبط معدل التضخم بعلاقة عكسية مع عدد العاطلين عن العمل،ويحسب كما يلي :

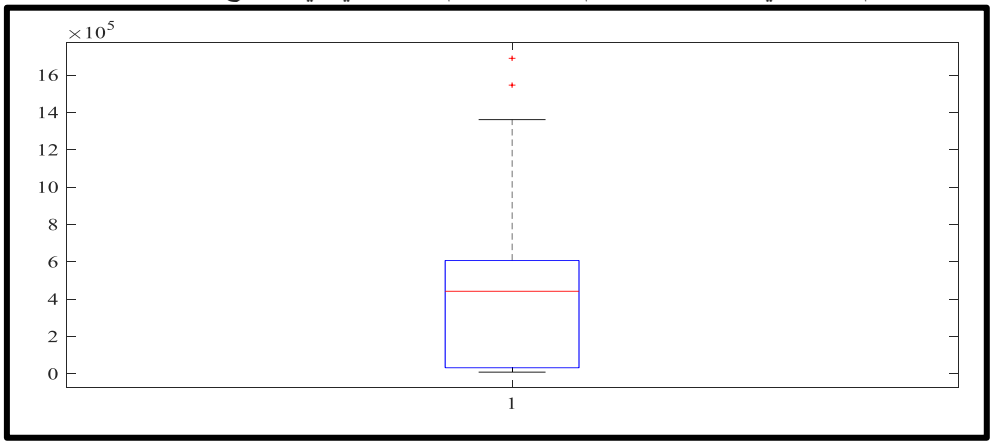
$$\text{معدل التضخم السنوي في العراق} = \left[1 - \frac{\text{معدل اسعار السلع والخدمات في السنة المحددة}}{\text{معدل اسعار السلع والخدمات في السنة السابقة لها}}\right] * 100$$

الجدول رقم (1) بيانات الدراسة

السنة	عدد العاطلين عن العمل(y_i)	معدل النمو السكاني(x_{i1})	معدل التضخم(x_{i2})
2004	22642	3.0347	26.96
2005	8532	3.0338	36.62
2006	55815	3.0306	53.61
2007	31899	3.0254	30.83
2008	38256	3.0177	2.67
2009	22295	3.5537	-2.8
2010	10910	2.607	2.46
2011	281169	2.6124	5.95
2012	542326	2.605	6.06
2013	606169	2.5974	1.86
2014	404734	2.5894	2.2
2015	585105	-2.1995	1.4
2016	478926	2.7164	0.5
2017	508295	2.6829	0.2
2018	1074182	2.6512	0.4
2019	1361090	2.6327	-0.2
2020	1689831	2.6126	0.6
2021	1546069	2.5914	6

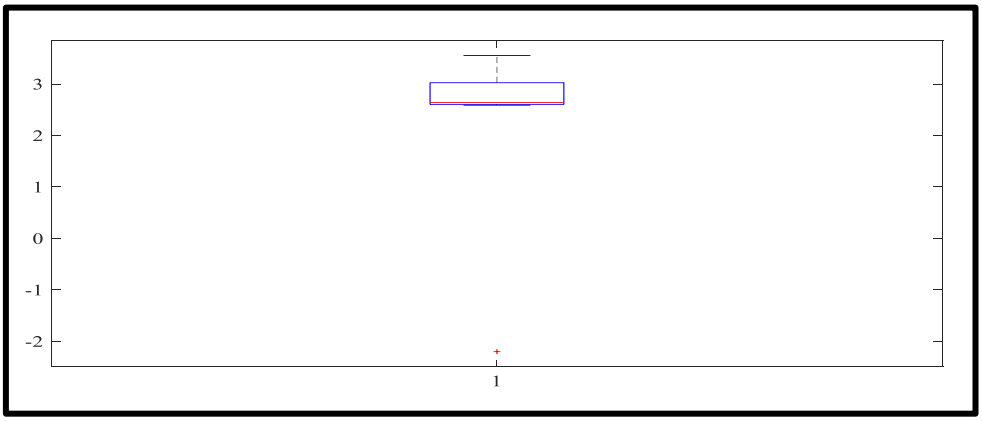
5- الكشف عن القيم الشاذة بأستعمال طريقة الرسم الصندوقي

تم الكشف عن القيم الشاذة في البيانات بأستخدام طريقة الرسم الصندوقي في برنامج SPSS



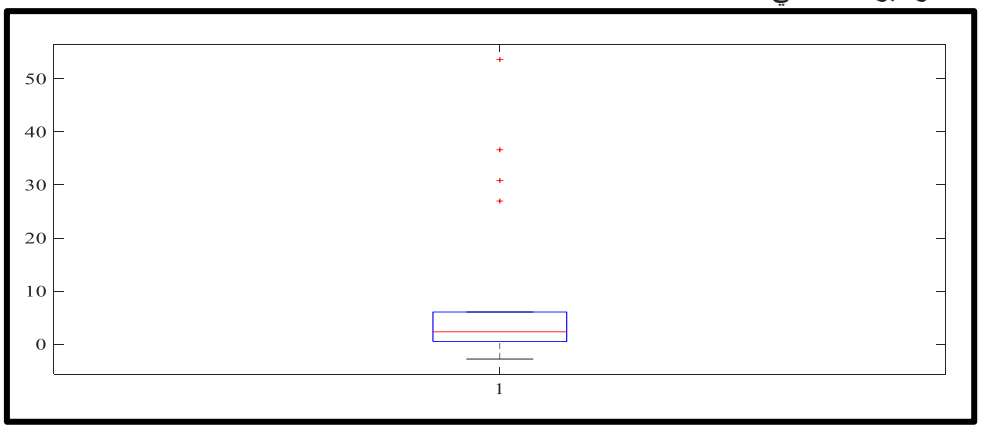
الشكل رقم (1)

حيث يمثل شكل رقم (1) الرسم الصندوقي للمتغير المعتمد (عدد العاطلين عن العمل في العراق) مبيناً القيم الشاذة وغير الشاذة في هذه الاعداد .



الشكل رقم (2)

حيث يمثل شكل رقم (2) الرسم الصندوقي للمتغير التوضيحي الاول (معدل النمو السكاني في العراق) مبيناً القيم الشاذة وغير الشاذة في هذه المعدلات .



الشكل رقم (3)

حيث يمثل شكل رقم (3) الرسم الصندوقي للمتغير التوضيحي الثاني (معدل التضخم في العراق) مبيناً القيم الشاذة وغير الشاذة في هذه المعدلات .

6- تقدير المعلمات واعتماد الأنموذج المقدر

نطبق طريقة **M** الحصينة في تقدير معالم هذا الأنموذج وذلك باستخدام بيانات الدراسة :

Criteria	M
β_0	13.2416
β_1	-0.1175
β_2	-0.0610

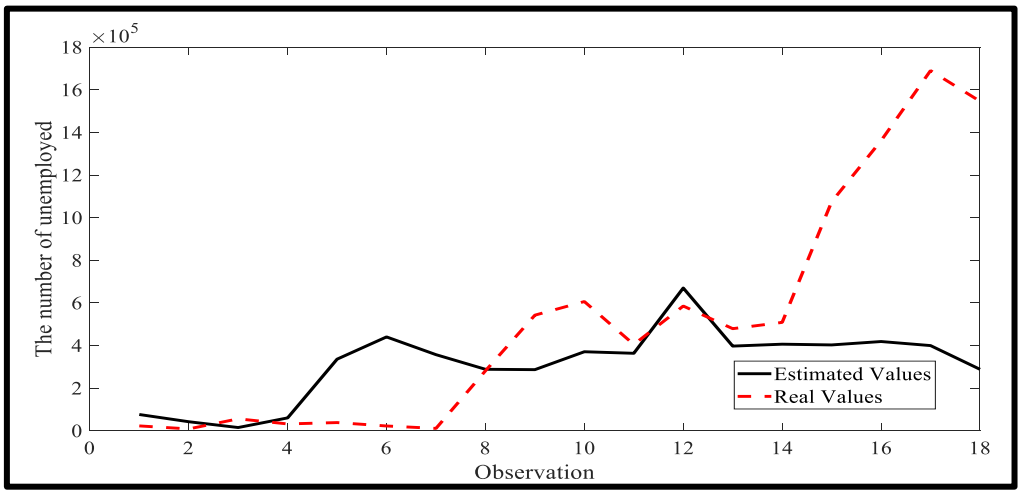
فيكون النموذج المقدر كالآتي :

$$\hat{Y} = exp(13.2416 - 0.1175X_1 - 0.061X_2)$$

بالتالي نستخدم الأنموذج المقدر لتقدير قيم y_i ومقارنتها بالقيم الحقيقة وحسب الجدول الآتي :

الجدول رقم (2) يمثل القيم الحقيقة والمقدرة للمتغير التابع

القيم المقدرة	القيم الحقيقية	القيم المقدرة	القيم الحقيقية
370655	606169	76164	22642
363389	404734	42256.8	8532
669644	585105	14996.3	55815
397125	478926	60215.1	31899
406054	508295	335795	38256
402627	1074182	440186	22295
418543	1361090	356933	10910
399552	1689831	288309	281169
288141	1546069	286630	542326



الشكل رقم (4)

حيث يمثل شكل رقم (4) رسم القيمة الحقيقية والمقدرة للمتغير التابع حيث تتقارب القيم تارة وتتباعد تارة اخرى .

7- الاستنتاجات

من خلال تطبيق طريقة M الحصينة لتقدير انموذج انحدار بواسون في حالة وجود القيم الشاذة في بيانات اعداد العاطلين عن العمل والمتغيرات المؤثرة عليها تبين مايلي :

1- ارتباط المتغير المعتمد (عدد العاطلين عن العمل) بعلاقة معنوية عكسية مع المتغير التوضيحي الأول (معدل النمو السكاني) , حيث نلاحظ قيمة سالبة للمعلمة B_1 , اي كلما زاد معدل النمو السكاني يدل على حياة مرفهه وازدهار سوق العمل وهذا يعني زيادة في عدد الوظائف (سواء في القطاع العام او الخاص) وبالتالي انخفاض عدد العاطلين عن العمل .

2- ارتباط المتغير المعتمد (عدد العاطلين عن العمل) بعلاقة معنوية عكسية مع المتغير التوضيحي الثاني (معدل التضخم) , حيث نلاحظ قيمة سالبة للمعلمة B_2 , اي كلما زاد معدل التضخم يدل على زيادة الطلب التي بدورها تعني زيادة في عدد الأجور التي بدورها تعني زيادة اعداد الموظفين (سواء في القطاع العام او الخاص) وبالتالي انخفاض عدد العاطلين عن العمل .

3- قُرب القيمة التقديرية من القيمة الحقيقية للمتغير المعتمد تارة وبُعدها تارة اخرى يعني هذا ان الأنموذج يحتاج الى متغيرات توضيحية اضافية لزيادة قوتهِ التنبؤية .

8- التوصيات :

1- اضافة متغير توضيحي اخر لتقوية الأنموذج مثل (حجم الاستيرادات) الذي يرتبط بعلاقة طردية مع عدد العاطلين عن العمل حيث كلما زاد استيراد البضائع من الخارج زاد عدد العاطلين عن العمل داخل البلد , خاصة وقد استجد في الوقت الحاضر اتباع نظام الكتروني للتحويل المالي يحدد حجم الاستيرادات بدقة .

2- اضافة متغير توضيحي اخر لتقوية الأنموذج مثل (عدد العمالة الوافدة من الخارج) الذي يرتبط بعلاقة طردية مع عدد العاطلين عن العمل حيث كلما زاد عدد العمالة الوافدة من الخارج زاد عدد العاطلين عن العمل , شرط توفر البيانات الدقيقة عن العمالة الوافدة .

المصادر :

1- هادي , خديجة عبد الكريم , 2021 " مقارنة بعض الطرائق الحصينة لتقدير انموذج الانحدار للبيانات ذات التوزيع الطبيعي الملتوي " رسالة ماجستير , كلية الادارة والاقتصاد , الجامعة المستنصرية .

2- باهض , زهراء خالد , 2022 " التقدير الحصين لأنموذج الانحدار شبه المعلمي مع تطبيق عملي " رسالة ماجستير , كلية الادارة والاقتصاد , الجامعة المستنصرية .

3- الجشعمي , حسين علي , 2007 " مقارنة لبعض المقدرات الحصينة لمعالم النماذج اللاخطية " , اطروحه دكتوراه , كلية الادارة والاقتصاد , الجامعة المستنصرية .

4- Winklmann Rainer , 2008 , " Econometric Analysis of Count Data " Fifth edition , springer-verlag Berlin Heidelberg .

5- Philip , N. and Sebastian , N. , 2015 , " Application of Poisson Regression on Traffic Safety "

The robust M method was used to estimate the Poisson regression model for the number of unemployed persons in Iraq

Muhammad Ghazi Jawad (researcher) mohamed2722471@gmail.com
Aseel Abd Al-Razzaq Rashid /aseelstat@uomustansiriyah.edu.iq

Abstract:

The Poisson regression model is one of the important enumeration models, as it was used to model much population, economic and health phenomena. Because of the presence of outliers in the data, the robust M method was used to estimate the parameters of this model, based on data for the unemployed in Iraq as a variable in response to the period (2004-2021) and the annual population growth rate as the first explanatory variable and the annual inflation rate as a second explanatory variable and within the same period, where The results revealed the association of these variables with a significant relationship with an inverse effect with the response variable.

Keywords: Poisson regression, M estimator method, number of unemployed .

****
****
****