المفاضلة بين طرائق تحليل الانحدار الحصين

دراسة إحصائية تطبيقية لتأثير أهم العوامل الديموغرافية على معدل الخصوبة الكلي في اليمن للفترة (1990-2013)

الاستـــلام: 26/ أغسطس /2023
التحكيم: 2/ سبتمبر /2023
القبـــــول: 10 / سبتمبر /2023

د. أحمد عبد الرزاق محمد الأنصاري [1،*]
د. عبد الرزاق أحمد الرازحي [2]

© 2023 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2023 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] أستاذ مساعد، قسم علوم البيانات وتكنولوجيا المعلومات- كلية العلوم الإدارية – جامعة تعز.
[2] أستاذ مشارك - قسم الإحصاء والمعلوماتية - كلية العلوم الإدارية – جامعة تعز.
* عنوان المراسلة: a733050161@gmail.com

https://doi.org/10.20428/jss.v29i3.2160

المفاضلة بين طرائق تحليل الانحدار الحصين

دراسة إحصائية تطبيقية لتأثير أهم العوامل الديموغرافية على معدل الخصوبة الكلي في اليمن للفترة (1990-2013)

الملخص:

هدف البحث إلى المقارنة بين مقدرات طرائق نموذج الانحدار الحصين المتمثلة في (طريقة M الحصينة- طريقة MM الحصينة -طريقة S الحصينة)، والكشف عن طبيعة توزيع البيانات، واكتشاف البيانات الشاذة والمتطرفة لمتغيرات الدراسة، بالإضافة إلى الكشف عن استقرارية السلاسل الزمنية، واختبار وجود علاقة التكامل المشترك، عبر تحليل العلاقة بين المتغيرات الديموغرافية المدروسة المتمثلة في (معدل وَفيَات الرضع-عدد وَفيَات الأمهات -معدل الوَفيَات الخام- معدل انتشار سوء التغذية - معدل وَفيَات الإناث- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي - نسبة وَفيَات الأمهات لكل 1000 مولود حي) كمتغيرات مستقلة، ومعدل الخصوبة الكلي كمتغير تابع في الجمهورية اليمنية للفترة (1990 – 2015).

وخلص البحث إلى العديد من النتائج أهمها: إن أفضل طريقة لتقدير معالم الانحدار الحصين هي طريقة M الحصينة، حيث حققت أفضل قيم لمعايير المفاضلة المتمثلة في (RMSE- MAE- MAPETheil Coefficient)، كما بينت اختبارات القيم الشاذة المتمثلة في (DFFITS- Covratio - Hat-Matrix (H)) أن بيانات الدراسة تعاني من قيم شاذة، وأوضحت الدراسة أن المتغيرات الديموغرافية المدروسة تؤثر على معدل الخصوبة الكلي بنسبة (84٪) بينما (14٪) تعود لعوامل أخرى غير مدرجة في النموذج المقدر.

الكلمات المفتاحية:نموذج الانحدار الحصين، طرائق الانحدار الحصين، القيم الشاذة اختبار جذر الوحدة، اختبار جوهانسون.

A comparison between Robust regression analysis methods

Applied statistical study of the most important demographic factors on the total fertility rate in Yemen for the period (1990-2013).

Abstract

the research aims to compare the estimators of robust regression methods (the robustM method - the robust MM method - the robustS method), the detection of the nature of the distribution of data, the discovery of abnormal and extreme data of the study variables, in addition to revealing the stability of time chains, and testing a relationship Joint integration by analyzing the relationship between the studied demographic variables represented in (infant mortality - the number of maternal deaths - raw death rate - the rate of widespread malnutrition - female death rate - child mortality rate under the age For every 1,000 neighborhood births) as independent variables, and the total fertility rate as a variable in the Republic of Yemen for the period (1990-2015).

The research concluded with many results, the most important of which is that the best way to estimate the features of the stoned slope is the m fort- as it achieved the best values of the comparison standards of (RMSE-Mae-MAPE Theil Coefficient), and tests of abnormal values represented in (H) HAT showed -Matrix- Covratio - DFFITS) The study data suffers from abnormal values, and the study showed that the studied demographic variables affect the total fertility rate by 84%, while 14% belong to other factors that are not included in the estimated form.

Keywords: a robust regression model, a robust regression method, outlier values, unit root test, Johansson test

https://doi.org/10.20428/jss.v29i3.2160

المقدمة:

يعد تحليل الانحدار من أهم النماذج الإحصائية الذي يهتم بتحليل العلاقات الرياضية بين متغير الاستجابة والمتغيرات التفسيرية، وتتمثل هذه العلاقة كتركيبة خطية تدعى معادلة الانحدار، إذ دقتها تعتمد على صحة تقدير معلماتها، والتي تشترط توفر فروض التحليل، المتمثلة في (توزيع البواقي طبيعيا- عدم وجود ارتباط ذاتي – تجانس تباين حد الخطأ العشوائي- التعدد الخطي منعدم) أما عند عدم تحقق فروض تحليل الانحدار فإن مقدرات تحليل الانحدار الخطي تصبح مقدرات غير متسقة ومتحيزة. إن الفرض الأساسي في نموذج تحليل الانحدار الخطي هو أن المتغيرات يجب أن تتبع التوزيع الطبيعي، ولكن من الصعوبة أن يتحقق هذا الفرض في كثير من الدراسات والبحوث، كما أن وجود قيم شاذة ينتج عنها مقدرات لا تحقق فروض الانحدار الخطي، لذا كان من الضروري البحث عن طرق بديلة يمكن منها الوصول على مقدرات متسقة، وأكثر دقة، وأقل تحيزًا في حالة أن بيانات الظواهر المدروسة لا ينطبق عليها افتراضات الانحدار الخطي التقليدي، وكان أحد أهم هذه الأساليب هو أسلوب الانحدار الحصين .ويتناول هذا البحث قياس وتحليل أثر بعض المتغيرات المتمثلة الديموغرافية في (معدل وَفيَات الرضع– عدد وَفيَات الأمهات– معدل الوَفيَات الخام- معدل انتشار سوء التغذية- معدل وَفيَات الإناث- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي - نسبة وَفيَات الأمهات لكل 1000 مولود حي) كمتغيرات مستقلة على معدل الخصوبة الكلي كمتغير تابع في الجمهورية اليمنية للفترة (1990-2015) عبر تطبيق نموذج الانحدار الحصين باستخدام أهم ثلاث طرائق للتقدير، والمتمثلة في (طريقة MM الحصينة - طريقة S الحصينة- طريقة M الحصينة) واختيار أفضل نموذج، وفقًا لمعايير المفاضلة المتمثلة في مقياس تشتت بواقي النموذجDeviance). ولغرض تحقيق أهداف البحث قام الباحثان بتقسيم البحث إلى ثلاثة محاور رئيسة: المحور الأول يشمل منهجية الدراسة، والمحور الثاني يمثل الإطار النظري لنموذج الانحدار الحصين، والمحور الثالث يشمل الجانب التطبيقي المتمثل في التحليل القياسي لأثر المتغيرات الديموغرافية على معدل الخصوبة الكلي باستخدام أسلوب الانحدار الحصين.

الإطار النظري للانحدار الحصين والدراسات السابقة:

إن الطريقة الإحصائية تسمى حصينة إذا كان الاستدلال الإحصائي لا يتأثر بشكل ملحوظ نتيجة لاختراق أي من شروطها الأساسية وإن مفهوم الحصانة بالنسبة للاختبار الإحصائي هي إهمال القيم الشاذة أو تقليل تأثيرها على البيانات، وإن التقدير الحصين هو التقدير الذي يكون قليل التأثير تجاه القيم الشاذة، وذا كفاءة تعادل كفاءة مقدرات المربعات الصغرى في حالة عدم وجود القيم الشاذة (خطار وآخرون، 2016: 4).

وتعد طريقة الانحدار الحصين بديلة لتقدير معلمات نموذج الانحدار، وتستخدم في حالة وجود القيم الشاذة في البيانات، حيث إن تقدير معلمات نموذج الانحدار الكلاسيكي في ظل وجود القيم الشاذة تكون غير كفؤة، وتسبب حالة عدم تطابق بين بيانات موضوع الدراسة والفروض الأساسية الواجب توافرها في النموذج. وبذلك تفقد الطرائق التقليدية المتمثلة في المربعات الصغرى خصائصها الجيدة لتقدير معلمات النموذج المدروس. (حسن ورضا،2011: 199).

إن التقدير الحصين جوهريًا يحاول أن يساوي ما بين الحالات التي تحاول تعديل الشواذ وعدم انتهاك الفروض لتقدير انحدار (OLS) أي أن صيغته من المربعات الصغرى الموزونة تعامل بشكل تكراري، وبكل خطوة يكون لدينا مجموعة أوزان تحدد على ضوء البواقي، بشكل عام تتناسب البواقي عكسيًا مع الأوزان، فكلما كبرت البواقي قلت الأوزان، لذلك الأوزان تعتمد على البواقي، والبواقي تعتمد على النموذج، والنموذج يعتمد على الأوزان (صالح، 2010).

يستخدم الانحدار الحصين لتقدير معلمات نموذج الانحدار الخطي في ظل عدم توزيع أخطاء نموذج الانحدار بشكل غير مطابق للتوزيع الطبيعي (البدري وصالح، 2010: 180). وفي حالة عدم تطابق البيانات مع التوزيع الطبيعي تستخدم الطرق بالطرق الحصينة، أما المقدرات الناتجة عن هذه الطرق فتسمى بالمقدرات الحصينة، وتكون غير حساسة تجاه القيم الشاذة، والقيم غير الموزعة توزيعًا طبيعيًا، وتعطي مقدرات كفؤة مقارنة بطريقة المربعات الصغرى. (حسن ورضا، 2011: 200).

أهم مقاييس اكتشاف القيم الشاذة:

يوجد العديد من المقاييس الإحصائية التي يمكن منها اكتشاف وجود القيم الشاذة، نستعرض أهمها في الآتي:

1- فحص عناصر قطر المصفوفة Hat-Matrix(H) (شاكر، 2009: 22).

تسمى هذه المصفوفة مصفوفة قيم قوة الرفع (Leveragevalues) وعادة ما تستخدم للكشف عن القيم الشاذة في المتغيرات التفسيرية، وتأخذ الشكل الآتي:

$$H_{n\times n} = X(X'X)^{-1}X' = \begin{bmatrix} h_{11}h_{12}\ldots..h_{1n} \\ h_{21}h_{22.}\ldots\ h_{2n} \\ . \qquad\qquad . \\ h_{n1}h_{n2}\ldots.h_{nn} \end{bmatrix} \qquad (1)$$

فإذا كانت

$h_{ii} > \frac{3m}{n}$ or $h_{ii} > \frac{2m}{n}$ فهذا يدل أن المشاهدة(i) مشاهدة متطرفة أو شاذة

2- إحصاءة (DFFITS)

حيث تعرف بأنها إحصائية تقيس مدى تأثير المعلمات المقدرة للتغيرات، وتحتسب كالآتي:

إذا كان $|DFFITS| > CUT\ OFF$ فإن هذه المشاهدة تعد مؤثرة.

حيث $CUT\ OFF$ تختلف من مصدر إلى آخر، وغالبًا ما تحدد بالقيمة 2. (حسن ورضا، 2011: 200).

وتحسب هذه الإحصاءة وفقًا للصيغة الرياضية الآتية: (شاكر، 2009: 223).

$$DF = \frac{e_i(h_{ii})^{\frac{1}{2}}}{s_{[i]}(1-h_{ii})} \qquad (2)$$

h_{ii}: قيم القطر الرئيس في مصفوفة قيم قوة الرفع المقدرة باستخدام جميع المشاهدات لتقدير دالة الانحدار

$s_{[i]}$: الخطأ القياسي بعد حذف المشاهدة [i]

https://doi.org/10.20428/jss.v29i3.2160

ولتحديد القيم الشاذة والمتطرفة وفقاً لهذا المقياس هناك عدة طرق أهمها:(Belsley,1980)

1- طريقة (Belsley)

اقترح (Belsley) (1980) أن المشاهدة الشاذة المؤثرة على نموذج الانحدار إذا كانت $DF>2\sqrt{\frac{m}{n}}$ فإن هذا يدل أن المشاهدة [i] مشاهدة شاذة أو متطرفة.

حيث m: عدد المتغيرات المستقلة

n: حجم العينة

2- طريقة (Neter)

تستخدم هذه الطريقة في حالة العينات الكبيرة والصغيرة والمتوسطة، وتعد المشاهدة الشاذة مؤثرة إذا كانت:

$$|DFFITS| > 1$$

3- طريقة (Chatterjee&Hadi)

وفقًا لهذه الطريقة المقترحة من قبل (Chatterjee&Hadi) فإن المشاهدة الشاذة إذا كانت:

$$|DFFITS| > 2\sqrt{\frac{m}{n-m}} \qquad (3)$$

4- مقياس(COVRATIO)

يحسب هذا المقياس وفقًا للصيغة الرياضية الآتية :(Brown ,2014: 202).

$$\mathrm{COVRATIO} = \left(\frac{S_i^2}{MSE}\right) * \frac{1}{1 - \mathrm{h_i}} \qquad (4)$$

حيث:

$$S_i^2 = \frac{[(N-P) * MSE] - e_i / 1 - \mathrm{h_i}}{N-P-1}$$

P: عدد المتغيرات

N: حجم العينة

h_i : قيم العنصر القطري رقم i في المصفوفة المقدرة H باستخدام جميع المشاهدات لتقدير دالة الانحدار

MSE : متوسط مربعات الخطأ

ووفقًا لهذا المقياس تعد المشاهدة شاذة إذا كانت القيمة المطلقة للمقياس (COVRATIO) أكبر من $\frac{1+3p}{n}$

طرائق التقديرات الحصينة (Robust Estimation Methods):

بما أن تفسير أي بيانات مبني على أساس قواعد وطرائق خاصة، ولكن في بعض الأحيان يواجه الباحث مشكلات في البيانات، مثل ابتعاد بيانات العينة عن التوزيع الطبيعي، وذلك بسبب وجود القيم الشاذة فيها، أواختلاف توزيع المجتمع قيد الدراسة عن التوزيع الطبيعي، وهذا يسبب انحرافات عن الافتراضات التي وضعت للطرائق التقليدية

منها: طريقة المربعات الصغرى الاعتيادية لتقدير معلمات النموذج الخطي، لهذا السبب لابد من وضع طرائق بديلة وإضافةً إلى ذلك توصّل كثير من الباحثين إلى نتيجة بأن هذه الطرائق تكون غير كفؤة في حالة عدم تحقيق أحد الافتراضات اللازمة، أو الشروط التي تعتمدها هذه الطرائق، لذا عمل كثير من الباحثين من أجل إيجاد طرائق أكثر كفاءة، ولا تتأثر كثيراً بالانحرافات عن الافتراضات المحددة، فكانت هذه الطرائق المسماة بالطرائق الحصينة، حيث أقل تأثراً في حالة اختلال شروط الطرائق التقليدية، وأهم الطرائق الحصينة لتقدير نموذج الانحدار الحصين الآتي:

(حسن ورضا، 2011: 201).

1- طريقة M الحصينة M -EstimationRobust

2- طريقة S الحصينة S –EstimationRobust

3- طريقة MM الحصينة MM -EstimationRobust

وسيتم في هذا البحث المقارنة بين هذه الثلاث الطرائق، وسنستعرض فيما يلي الإطار النظري لها.

1- طريقة MM الحصينة (MM -Estimation Robust)

ستستخدم طريقة MM الحصينة MM–Estimation Robust التي اقترحها (Yohai) حيث يضيف تحسينًا على كفاءة مقدرات معلمات نموذج الانحدار الخطي في ظل عدم توزيع أخطاء نموذج الانحدار بشكل غير متطابق مع التوزيع الطبيعي، وتتلخص هذه الطريقة في تحجيم تأثير قيم البواقي الكبير كالآتي: (جواد، واخرون 2009: 156).

وأن مقدر MM الحصين يُعرَّف بأنه حل للمعادلة الآتية:

$$\sum_{i=i}^{n}\psi_i(\frac{r_i}{s_i})x_i = 0 \qquad (5)$$

حيث r_i: البواقي، s_i: تقدير طريقة M لأخطاء القياس، ψ_i: تأثير الدالة M الحصينة والتي تحقق الشرط التالي

$$s(T_1) \le s(T_0) \qquad (6)$$

حيث T_0 ؛ T_1 مقدرات الانحدار

حيث إن

$$S(\theta) = \sum_{i=1}^{n}\rho(\frac{r_i}{s_n}) \qquad (7)$$

2- طريقة M الحصينة (M -EstimationRobust)

تستخدم طريقة M -EstimationRobust لتقدير معلمات نموذج الانحدار الخطي في ظل عدم توزيع أخطاء نموذج الانحدار بشكل غير متطابق مع التوزيع الطبيعي، وتتلخص هذه الطريقة في تحجيم تأثير قيم البواقي الكبير باستخدام معادلة (Huber,1964) الآتية: (شاكر، 2009: 223).

$$\phi^*_{e_i} = \frac{\psi_{(e_i)}}{e_i} \qquad (8)$$

https://doi.org/10.20428/jss.v29i3.2160

إذ أن:

$$C>0, \psi_{(e_i)} = \max\{-c, \min(e_i, c)\}$$

حيث c=1.7orc=1.5 ، $e_i = y_i - \hat{y}_i$

وبعد الحصول على قيم $\psi_{(e_i)}$ تتم موازنة قيمها كما في المعادلة رقم (1)

من المعادلة رقم (8) يمكننا ملاحظة أنه عندما تكون $|e_i| \to \infty$ فإن قيمة $\phi^*_{e_i}$ تكون صغيرة، وتقترب من الصفر $[\phi^*_{e_i} \to 0]$ وباستخدام مبدأ طريقة المربعات الصغرى العادية والقاضي بتصغير مجموع مربعات الخطأ إلى أقل ما يمكن للحصول مقدرات Mالحصينة (R.M) وتحسب وفقًا للصيغة الرياضية الآتية:

$$\hat{B}_M = (X'\phi X)^{-1} X'\phi Y \qquad (9)$$

حيث إن:

$$\phi = \begin{bmatrix} \phi^*_{e_i} & 0 & 0 & 0 \\ 0 & \phi^*_{e_i} & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & \phi^*_{e_i} \end{bmatrix}$$

3- طريقة Sالحصينة (S -EstimationRobust)

اقترحت طريقة S الحصينة كطريقة لتقدير معالم الانحدار الحصين من قبل (Yohai Rousseeuw and) (1984) وتهدف إلى تقليل تشتت البواقي، ويعبر عنها رياضيًا كالآتي: **(Alma,2011:415)**

$$MIN\ S = (e_1(\beta),...e_n(\beta)) \qquad (10)$$

حيث:

$e_i(\beta)$: البواقي للمعلمة β_i ، $i = 1,2,..n$

فإن دالة تقليل تشتت البواقي تعطى بالصيغة الآتية:

$$\frac{1}{n-p}\sum_{i=1}^{n} x\left(\frac{Y_i - \hat{Y}_i}{S}\right) = K \qquad (11)$$

حيث:

K: ثابت يمثل القيمة المتوقعة لعدد المتغيرات $E_\phi[P]$ بانحراف معياري ϕ

واقترح: (Rousseeuw and Yohai) (1984) دالة استدلال الانحدار الآتية: **(Dehnel,2015:140)**

$$p(x) = \begin{cases} \frac{x^2}{2} - \frac{x^4}{2c^2} + \frac{x^6}{6c^4} & \text{for}|x| \le c \qquad (12) \\ \frac{c^2}{6} & \text{for}|r > k| \end{cases}$$

حيث:

c : معلمة ثابتة ضابطة عادة تحدد بـ ($c = 1.48$)، ($k = 0.199$)

معايير قياس أخطاء التنبؤ: (Criteria for measuring forecasting errors)

يوجد العديد من المعايير لقياس حجم الأخطاء والتي من خلالها المفاضلة بين مقدرات النماذج أو الطرق المختلفة للتقدير، وتعد أفضل النماذج تقديراً تلك النماذج التي تحقق أدنى القيم لهذه المعايير، وأهم هذه المعايير ما يلي:(شعراوي،2005، 14).

1- مجموع الأخطاء (Sum of Errors)ويرمز له عادة بالرمز (SE) ويعرف بالصورة الآتية:

$$SE = \sum_{t=1}^{n} e_t = \sum_{t=1}^{n} (y_t - \hat{y}_t) \qquad (13)$$

وهذا المقياس لا يفيد كثيراً إذا كانت الأخطاء عشوائية، فإن هذا مجموع فروق الأخطاء عادة ما يكون قريباً جداً من الصفر، بغض النظر عن حجم هذه الأخطاء.

2-متوسط الانحرافات المطلقة (Mean Absolute Deviation)

يعبر عن متوسط الفرق المطلق لخطأ البواقي، ويعطى بالصيغة الرياضية الآتية:

$$MAD = \frac{1}{n} \sum_{t=1}^{n} \left| e_t \right| = \frac{1}{n} \sum_{t=1}^{n} \left| y_t - \hat{y}_t \right| \qquad (14)$$

3- متوسط مربعات الأخطاء (Mean Squared Error)

يعد أهم المقاييس لقياس الأخطاء، ويعطى بالصيغة الرياضية الآتية:

$$MSE = \frac{1}{n} \sum_{t=1}^{n} e_t^2 = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \qquad (15)$$

4- متوسط الأخطاء النسبية المطلقة (Mean absolute percentage error)

يعطى هذا المقياس وفقًا للصيغة الرياضية الآتية:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{e_t}{y_t} \right| \qquad (16)$$

الإطار النظري لاختبارات الكشف عن استقرارية السلاسل الزمنية ومشكلات النموذج القياسي:

https://doi.org/10.20428/jss.v29i3.2160

يتناول هذا الجزء الإطار النظري لاختبارات الكشف عن استقرارية السلاسل الزمنية المتمثلة في اختبار جذر الوحدة، واختبار التكامل المشترك (لجوهانسون)، واختبار السببية لجرانجر، بالإضافة إلى الكشف عن مشكلات النموذج القياسي المتمثلة في الارتباط الذاتي، وعدم ثبات تباين حد الخطأ والتعدد الخطي.

1- اختبار جذر الوحدة (ADF) للكشف عن استقرارية السلاسل الزمنية

يقوم اختبار (DF) على افتراض أنه لا يوجد ارتباط تسلسلي بين الأخطاء، وبناءً على ذلك فإذا تبين عن طريق اختبار (دربن واتسون) رفض هذ الفرض فإن تطبيق هذا الاختبار لا يكون مناسبًا، ويعطي نتائج غير دقيقة بشأن سكون أو عدم سكون السلسلة. ولهذه الأسباب فقد اقترح (ديكي فوللر) اختبار آخر يسمى (ADF) يستخدم عندما تظهر النتائج وجود ارتباط تسلسلي بين الأخطاء.

وهناك ثلاث صيغ للنموذج الذي يمكن استخدامه في هذه الحالة (عناني، 2009: 663).

الصيغة الأولى:

$$\Delta Y_T = \lambda Y_{t-1} + \sum_{j=1}^{m} \alpha_j \Delta Y_{t-1} + u_t \quad (16)$$

وبناءً على هذه الصيغة يجري اختبار الفروض الآتية:

فرض العدم: $H_0 : \lambda = 0$ ، الفرض البديل: $H_1 : \lambda < 0$

ويلاحظ في هذه الصيغة أنها لا تحتوي على مقدار ثابت أو اتجاه زمني، كما يلاحظ أيضًا أنه يجري إدراج عدد من الفروق ذات الفجوة (m) حتى تختفي مشكلة الارتباط التسلسلي بين الأخطاء. وبعد تقدير النموذج (16) باستخدام طريقة المربعات الصغرى وحساب قيمة اختبار τ^* للمعلمة λ معامل (Y_{t-1}) ومقارنتها بقيمة τ المستخرجة من جدول (ADF) فإذا كانت τ^* أكبر من τ نرفض فرض العدم $H_0 : \lambda = 0$ ونقبل الفرض $H_1 : \lambda < 0$ أي أن السلسة لا تحتوي على جذر الوحدة، أي إن السلسلة الزمنية مستقرة.

الصيغة الثانية:

$$\Delta Y_T = B_1 + \lambda Y_{t-1} + \sum_{j=1}^{m} \alpha_j \Delta Y_{t-1} + u_t \quad (17)$$

ويلاحظ في هذه الصيغة أنها تحتوي على مقدار

ثابت B_1 ولا تحتوي على اتجاه زمني كما

وبناءً على هذه الصيغة يجري اختبار الفروض الآتية:

فرض العدم: $H_0 : \lambda = 0, B_1 = 0$ ، الفرض البديل: $H_1 : \lambda < 0, B_1 = 0$

وبعد تقدير النموذج باستخدام طريقة المربعات الصغرى، واختبار معنوية λ، B_1 كما يلي:

$$\tau^*_{\lambda} = \frac{\hat{\lambda}}{S_{\hat{\lambda}}} \quad \tau^*_{B_1} = \frac{\hat{B}}{S_{\hat{B}_1}}$$

وبمقارنة قيم $\tau^*_{B_1}$ ، τ^*_{λ} بقيم τ المستخرجة من جدول (ADF) فإذا كانت $\tau^*_{B_1}$ ، τ^*_{λ} أكبر من نظيراتها الجدولية نرفض فرض العدم $H_0 : \lambda = 0, B_1 = 0$ ونقبل الفرض $H_1 : \lambda < 0, B_1 = 0$ أي إن السلسة لا تحتوي على جذر الوحدة، أي إن السلسلة الزمنية مستقرة.

الصيغة الثالثة

$$\Delta Y_T = B_1 + B_2 T + \lambda Y_{t-1} + \sum_{j=1}^{m} \alpha_j \Delta Y_{t-1} + u_t \quad (18)$$

وتحتوي هذه الصيغة على مقدار ثابت، وكذلك على متغير الزمن، وبناءً على هذه الصيغة يجري اختبار الفروض الآتية:

$$H_1 = \lambda < 0 \;\; H_0 = \lambda = 0$$

$$H_1 = B_1 \neq 0 \;\; H_0 = B_1 = 0$$

$$H_1 = B_2 \neq 0 \;\; H_0 = B_2 = 0$$

ثم يجري حساب القيم المحسوبة لاختبار τ^* المختلفة وكما يلي:

$$\tau^*_{\lambda} = \frac{\hat{\lambda}}{S_{\hat{\lambda}}} \quad \tau^*_{B_2} = \frac{\hat{B}_2}{S_{\hat{B}_2}} \quad \tau^*_{B_1} = \frac{\hat{B}_1}{S_{\hat{B}_1}}$$

وتقارن هذه القيم المحسوبة بالقيم الحرجة المستخرجة من جدول (ADF) واتخاذ قرار بقبول أو رفض الفروض السابقة.

2- اختبار التكامل المشترك لجوهانسون

ولتحديد عدد متجهات التكامل المشترك اقترح (Johansen 1992,juselius 1990) اختبار إحصائيتين: الأول: اختبار الأثر (Trace test-λtrace)، والثاني: اختبار القيمة الكامنة العظمى Maximal (eigenvalue) **وفيما يلي استعراض الاختبارين: (العبدلي، 2007: 50).**

1- **اختبار الأثر** (Trace test-λ trace)

يختبر هذا الاختبار الفرضيات الآتية:

فرضية العدم H_0: إن عدد متجهات التكامل المشترك يقل عن أو يساوي العدد (q)

الفرضية البديلة H_1: عدد متجهات التكامل المشترك (q=r)

ويحسب وفقًا للصيغة الرياضية الآتية:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{p} In(1 - \hat{\lambda}_i) \quad (19)$$

حيث $(\lambda_{r+1}, \ldots . \lambda_n)$ تمثل أقل المتجهات الكامنة، وتشير فرضية العدم أن عدد متجهات التكامل المشترك الكامنة يساوي أو يقل عن (r).

2- **اختبار القيمة الكامنة العظمى** (Maximal eigenvalue)

يختبر هذا الاختبار الفرضيات الآتية:

https://doi.org/10.20428/jss.v29i3.2160

فرضية العدم H_0: إن هناك (r) متجه للتكامل المشترك

الفرضية البديلة H_1 إن هناك (r+1) متجه للتكامل المشترك

ويحسب هذا الاختبار وفقًا للصيغة الآتية:

$$\lambda_{max}(r, r+1) = -Tln(1-\hat{\lambda}_{r+1}) \quad (20)$$

الدراسات السابقة

نستعرض في هذا الجزء أهم ثلاث دراسات سابقة في طرائق الانحدار الحصين:

1- دراسة شاكر (2009)

هدف البحث إلى إيجاد تركيبة حصينة موزونة لمقدرات معلمات نموذج الانحدار الخطي بدمج أسلوبM الحصين وطريقة المربعات الصغرى الموزونة، لمعالجة أثر القيم الشاذة والتي لا تنسجم ونمط البيانات العام، وبالتالي الحصول على مقدرات جديدة، وبالاعتماد على معيار متوسط مربعات الخطأ جرى فحص كفاءة المقدرات الجديدة والتي تبين أن نتائجها ذات دقة عالية، مقارنة مع طريقة المربعات الصغرى الاعتيادية والمربعات الصغرى الموزونة، وأسلوبMالحصين.

2- دراسة Alma (2011)

هدفت الدراسة على التعرُّف على سلوك القيم الشاذة والمتطرفة في نموذج تحليل الانحدار، عبر تطبيق أهم طرائق الانحدار الحصين، حيث جرى استخدام أسلوب المحاكاة في مقارنة مقدرات طرائق الانحدار الحصينة المتمثلة في (طريقة المربعات الصغرى المشذبة - طريقة M الحصينة - طريقة MM الحصينة - طريقة S الحصينة)، وتوصلت الدراسة إلى أن مقدرات طريقتي (MM-S) الحصينة مقارنة مع أكثر كفاءة، مقارنة بطريقة المربعات الصغرى الاعتيادية.

3- دراسة شاكر (2009)

هدفت الدراسة إلى إيجاد تركيبة حصينة موزونة لمقدرات معلمات نموذج الانحدار الخطي بدمج أسلوبMالحصين وطريقة المربعات الصغرى الموزونة لمعالجة أثر القيم الشاذة والتي لا تنسجم ونمط البيانات العام، وتوصلت الدراسة إلى أن أسلوبM الحصين الموزون (R.M.W) المقترح قد أثبت كفاءة في تقدير معلمات نموذج الانحدار وبدقة عالية عند وجود التطرف في قيم المتغير(المتغيرات) التوضيحية أو في المتغير المعتمد أو في كليهما.

مشكلة الدراسة:

إن استخدام طريقة المربعات الصغرى تؤدي الى مقدرات غير متسقة وأقل كفاءة، حيث تتأثر طريقة المربعات الصغرى (OLS) بالقيم المتطرفة، فتكون المشاهدات منحرفة عن العلاقة الخطية بين المتغير التابع والمستقل. (Wen,2013) وبالتالي تكون مقدرات المربعات الصغرى غير فعالة (Cetin, M., &Toka, 2011) وللتغلب على هذه المشكلة جرى تطوير طرق تقدير الانحدار جديدة، سميت بطرق الانحدار الحصينة (Schumacker, & Mount., 2002). وتعد هي الطرق الأساسية لتحليل البيانات التي تعاني من القيم المتطرفة، وتعطي نتائج جيدة، وتتمثل هذه الطرق في (طريقة M الحصينة - طريقة M الحصينة - طريقة S الحصينة) (Abd-Almonem,2015). وتكمن مشكلة الدراسة في أن أغلب البيانات لكثير من الظواهر الاقتصادية والاجتماعية وغيرها لا تتوزع توزيعًا غير طبيعي، كما أنها تعاني من القيم الشاذة والمتطرفة، وبالتالي فإن استخدام طرائق الانحدار الحصين تعطي مقدرات

أكثر دقة، وأقل تحيز، وعليه سيستخدم الباحثان في هذا البحث أهم طرائق أسلوب الانحدار الحصين عبر دراسة أهم العوامل الديموغرافية المتمثلة في (معدل وَفيَات الرضع– عدد وَفيَات الأمهات– معدل الوَفيَات الخام - معدل انتشار سوء التغذية - معدل وَفيَات الإناث- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي - نسبة وَفيَات الأمهات لكل 1000 مولود حي) على معدل الخصوبة الكلي في الجمهورية اليمنية.

أهداف الدراسة:

تهدف الدراسة إلى:

1- المقارنة بين مقدرات طرائق نموذج الانحدار الحصين المتمثلة في (طريقة M الحصينة - طريقة M الحصينة - طريقة S الحصينة).

2- الكشف عن طبيعة توزيع البيانات واكتشاف البيانات الشاذة والمتطرفة لمتغيرات الدراسة.

3-الكشف عن استقرارية السلاسل الزمنية، واختبار وجود علاقة التكامل المشترك بين المتغيرات الديموغرافية المدروسة ومعدل الخصوبة الكلي.

فرضيات الدراسة:

1- لا يوجد فرق في مقدرات نموذج تحليل الانحدار الحصين باختلاف طريقة التقدير المستخدمة.

2- لا تعاني السلاسل الزمنية لمتغيرات الدراسة المتمثلة في (معدل وَفيَات الرضع– عدد وَفيَات الأمهات– معدل الوَفيَات الخام - معدل انتشار سوء التغذية- معدل وَفيَات الإناث- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي - نسبة وَفيَات الأمهات لكل 1000 مولود حي- معدل الخصوبة الكلي) من جذر الوحدة عند مستوى معنوية (0.05).

3- عدم وجود أي متجه للتكامل المشترك بين سلاسل المتغيرات (معدل وَفيَات الرضع– عدد وَفيَات الأمهات– معدل الوَفيَات الخام - معدل انتشار سوء التغذية- معدل وَفيَات الإناث- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي - نسبة وَفيَات الأمهات لكل 1000 مولود حي) كمتغيرات مستقلة، ومعدل الخصوبة الكلي كمتغير تابع عند مستوى معنوية (0.05).

4- لا يوجد تأثير ذات دلالة إحصائية للمتغيرات الديموغرافية المدروسة في معدل الخصوبة الكلي عند مستوى معنوية (0.05).

أهمية الدراسة:

تأتي هذه الدراسة كدراسة تحليلية إحصائية قياسية لتوضيح أهمية استخدام نموذج الانحدار الحصين في حالة وجود قيم شاذة وعدم تطابق البيانات للتوزيع الطبيعي، وذلك من تحليل العلاقة بين المتغيرات الديموغرافية المتمثلة في (معدل وَفيَات الرضع– عدد وَفيَات الأمهات– معدل الوَفيَات الخام - معدل انتشار سوء التغذية- معدل وَفيَات الإناث- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي - نسبة وَفيَات الأمهات لكل 1000 مولود حي) ومعدل الخصوبة الكلي، والتوصل إلى بناء نموذجٍ قياسيٍّ إحصائيٍّ ملائمٍ للبيانات الشاذة والمتطرفة.

https://doi.org/10.20428/jss.v29i3.2160

الجانب التطبيقي للدراسة

في هذا الجزء يقوم الباحثان باختبار التوزيع الطبيعي لمتغيرات الدراسة، واختبار استقرارية السلاسل الزمنية، واختبارات القيم الشاذة والتكامل المشترك بين المتغير التابع (معدل الخصوبة الكليY) والمتغيرات المستقلة (معدل وَفيَات الرضعx1– عدد وَفيَات الأمهاتx2– معدل الوَفيَات الخامx3- معدل انتشار سوء التغذيةx4- معدل وَفيَات الإناثx5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات لكل 1000 مولود حي X7) خلال الفترة (1990-2014) ومن ثم بتطبيق طرائق الانحدار الحصين المتمثلة في (طريقة MM الحصينة طريقة M الحصينة- طريقة S الحصينة)، بهدف المقارنة بين مقدراتها المختلفة.

منهج الدراسة:

اعتمدت الدراسة على المنهج التحليلي القياسي، ممثلًا باستخدام أسلوب الانحدار الحصين لتحليل لعلاقة بين أهم المتغيرات الديموغرافية المدروسة ومعدل الخصوبة الكلي في الجمهورية اليمنية.

توصيف متغيرات البحث

أولاً: مفهوم الاقتراع

المتغير التابع

- معدل الخصوبة الكليy

المتغيرات المستقلة

- معدل وَفيَات الرضع لكل1000 مولود حيx1
- عدد وَفيَات الأمهاتx2
- معدل الوَفيَات الخام لكل (1000) من السكانx3
- معدل انتشار سوء التغذية (٪ من تعداد السكان) x4
- معدل وَفيَات الإناث لكل (1000) من الإناثx5
- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود)
- نسبة وَفيَات الأمهات لكل (1000) مولود حي.

قاعدة بيانات الدراسة

البيانات الآتية تمثل بيانات متغيرات الدراسة المتمثلة في (معدل الخصوبة الكلي - معدل وَفيَات الرضعx1– عدد وَفيَات الأمهاتx2– معدل الوَفيَات الخامx3- معدل انتشار سوء التغذيةx4- معدل وَفيَات الإناثx5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات لكل 1000 مولود حي X7) والجدول الذي يوضح بيانات متغيرات الدراسة في الفترة (1990-2014):

https://doi.org/10.20428/jss.v29i3.2160

جدول رقم (1) بيانات متغيرات الدراسة للفترة (1991-2017)

السنوات	Y	X1	X2	X3	X4	X5	X6	X7
1990	8.607	88.70	3500.00	11.57	NA	246.69	116.80	547.00
1991	8.459	86.40	3500.00	11.34	28.90	244.08	114.40	535.00
1992	8.27	84.40	3500.00	11.13	28.60	241.48	112.20	523.00
1993	8.045	82.70	3500.00	10.92	27.90	239.82	109.80	517.00
1994	7.792	81.10	3500.00	10.70	28.30	238.16	107.20	507.00
1995	7.521	79.70	3400.00	10.45	29.30	236.50	103.80	498.00
1996	7.248	78.20	3400.00	10.19	30.30	234.84	99.90	488.00
1997	6.986	76.50	3300.00	9.92	30.80	233.18	95.40	476.00
1998	6.744	74.40	3200.00	9.66	30.70	231.88	90.70	462.00
1999	6.527	71.90	3100.00	9.40	30.10	230.58	85.90	451.00
2000	6.335	69.00	3100.00	9.15	29.60	229.28	81.30	440.00
2001	6.164	66.10	3100.00	8.93	29.40	227.98	76.80	435.00
2002	6.002	63.00	3100.00	8.72	30.00	226.67	72.60	431.00
2003	5.842	60.10	3100.00	8.52	30.50	225.60	68.50	429.00
2004	5.68	57.20	3200.00	8.35	30.90	224.52	64.70	429.00
2005	5.515	54.50	3200.00	8.19	30.60	223.45	61.00	428.00
2006	5.347	51.90	3300.00	8.03	29.70	222.37	57.60	428.00
2007	5.18	49.40	3300.00	7.88	28.50	221.29	54.20	424.00
2008	5.016	47.00	3300.00	7.73	27.70	218.20	51.10	417.00
2009	4.856	44.70	3400.00	7.58	27.30	215.12	48.40	417.00
2010	4.703	42.40	3400.00	7.43	27.00	212.03	46.00	416.00
2011	4.556	40.30	3400.00	7.28	26.30	208.94	43.80	413.00
2012	4.416	38.40	3400.00	7.15	25.80	205.85	41.90	410.00
2013	4.284	36.70	3400.00	7.02	25.70	203.67	NA	406.00
	4.16							

المصدر: المؤشرات الاقتصادية لليمن، البنك الدولي، http://databank.albankaldawli.org/data/home.asp

اختبار التوزيع الطبيعي للبيانات والكشف عن القيم الشاذة

اختبارات الكشف عن القيم الشاذة : في هذا الجزء قام الباحثان باستخدام اختباري(DIFET) و(COVRATIO) **ومصفوفة** (Hat-Matrix) (H) **وكانت النتائج كما في الجدول الآتي:**

جدول رقم (2) نتائج اختبارات (COVRATIO - DFFITS)- (H) Hat-Matrix

Influence Statistics

Date: 10/28/18 Time: 23:27

Sample: 1990 2014

Included observations: 22

Obs.	Resid.	RStudent	DFFITS	COVRATIO	Hat Matrix
1991	0.004268	0.331862	-0.591561	**7.064524**	0.760622
1992	-0.007827	-0.432514	0.453040	**3.384315**	0.523166
1993	0.029600	1.458144	-0.981613	0.782550	**0.311859**
1994	0.006295	0.313853	-0.266138	**2.927793**	0.418285
1995	-0.031915	-1.631927	1.167694	0.615929	**0.338617**
1996	0.000996	0.046884	-0.034580	**2.789599**	0.352338
1997	-0.001456	-0.068601	0.050696	**2.789087**	0.353218
1998	-0.021645	-0.982272	0.569141	1.362847	0.251340
1999	-0.020269	-0.980512	0.705878	1.552170	0.341354
2000	-0.007349	-0.337190	0.227260	**2.453929**	0.312360
2001	0.036245	1.831774	-1.179341	0.407701	**0.293042**
2002	0.019033	0.828313	-0.417073	1.502871	0.202255
2003	0.004650	0.212226	-0.141747	**2.544827**	0.308483
2004	0.020842	1.013955	-0.736890	1.503841	**0.345620**
2005	-0.008234	-0.384139	0.271528	**2.478833**	0.333170
2006	0.023323	1.111561	-0.741670	1.264600	**0.308054**
2007	-0.013011	-0.691138	0.654289	**2.570349**	0.472633
2008	-0.055380	**-3.284973**	**1.914208**	0.019261	**0.253485**
2009	0.005065	0.257281	-0.228491	**3.107236**	0.440941
2010	0.004353	0.185335	-0.094404	**2.230963**	0.206006
2011	0.007910	0.359738	-0.235074	**2.384999**	0.299232
2012	0.004507	0.262252	-0.304368	**4.070560**	0.573920

المصدر: من إعداد الباحثيْن بالاعتماد على نتائج البرنامج الإحصائي (Eviews9)

يتضح من الجدول السابق رقم (2) وجود قيم شاذة في بيانات متغيرات الدراسة المتمثلة في (معدل الخصوبة الكلي - معدل وَفيَات الرضع x1– عدد وَفيَات الأمهات x2– معدل الوَفيَات الخام x3- معدل انتشار سوء التغذية x4- معدل وَفيَات الإناث x5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات لكل 1000 مولود حي X7) وتركزت القيم الشاذة (المتطرفة) في بيانات السنة (2008) وفقًا لاختبار (DFFITS)، أما فيما يخص اختبار مصفوفة (Hat-Matrix) (H) فقد تركزت القيم الشاذة في السنوات (2008، 2006، 2004، 2001، 1995، 1993)

https://doi.org/10.20428/jss.v29i3.2160

وبالنسبة لاختبار (COVRATIO) تركزت القيم الشاذة في السنوات (1996،1994،1992،1991، 2000،1997، 2003،2005 ،2010،2009،2007، 2011 ،2012) ولوجود بيانات شاذة في بيانات السلاسل الزمنية فإن الأسلوب الملائم لتحليل العلاقة بين بعض المتغيرات الاقتصادية المتمثلة في - معدل وَفيَات الرضعx1– عدد وَفيَات الأمهاتx2– معدل الوَفيَات الخامx3- معدل انتشار سوء التغذيةx4- معدل وَفيَات الإناثx5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات (لكل 1000 مولود) حي X7 كمتغيرات مستقلة ومعدل الخصوبة الكلي كمتغير تابع هو أسلوب الانحدار الحصين.

اختبار التوزيع الطبيعي لبيانات متغيرات الدراسة:

جرى استخدام اختبار (Jarque-Bera) لاختبار مدى تطابق البيانات مع التوزيع الطبيعي والجدول يوضح اختبار (Jarque-Bera) لمتغيرات الدراسة.

جدول رقم (3) اختبار (Jarque-Bera) لمتغيرات الدراسة

اسم المتغير	Jarque-Bera إحصاءة اختبار	مستوى المعنوية	القرار الإحصائي
yمعدل الخصوبة الكلي	6.684231	0.035362	توجد فروق جوهرية
x1معدل وَفيَات الرضع	7.124197	0.028379	توجد فروق جوهرية
x2عدد وَفيَات الأمهات	6.818938	0.033059	توجد فروق جوهرية
x3معدل الوَفيَات الخام	6.532624	0.038147	توجد فروق جوهرية
x4معدل انتشار سوء التغذية	6.550346	0.03781	توجد فروق جوهرية
x5معدل وَفيَات الإناث	3.043122	0.218371	لا توجد فروق جوهرية
معدل وَفيَات الأطفال دون سن الخامسة x6	7.270252	0.026381	توجد فروق جوهرية
نسبة وَفيَات الأمهات لكل 1000 مولود حي X7	11.4852	0.003206	توجد فروق جوهرية

المصدر: إعداد الباحثَيْن بالاعتماد على نتائج البرنامج الإحصائي(Eviews9)

يتضح من الجدول رقم (3) أن جميع بيانات متغيرات الدراسة المتمثلة في (معدل الخصوبة الكلي - معدل وَفيَات الرضعx1– عدد وَفيَات الأمهاتx2– معدل الوَفيَات الخامx3- معدل انتشار سوء التغذيةx4- معدل وَفيَات الإناثx5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات (لكل 1000 مولود) حي X7 لا تتوزع توزيعًا طبيعيًا، وهذا ما أوضحه اختبار (Jarque-Bera) حيث بلغ مستوى المعنوية المصاحبة للاختبار لجميع المتغيرات (0.000) باستثناء متغير معدل وَفيَات الإناث وهي أقل من المستوى المعتمد في المقارنة (0.05) وعليه البيانات لا تتوزع توزيعًا طبيعيًا، وبالتالي فإن النموذج المناسب لتحليل العلاقة بين العوامل الديموغرافية المدروسة كتغيرات مستقلة ومعدل الخصوبة الكلي كمتغير تابع هو نموذج الانحدار الحصين.

الكشف عن استقرارية السلاسل الزمنية لمتغيرات الدراسة

استخدم الباحثان في هذا الجز اختبار ديكي فولر الموسع (ADF) لاختبار استقرارية السلاسل الزمنية لمتغيرات الدراسة المتمثلة في (معدل الخصوبة الكلي - معدل وَفيَات الرضعx1– عدد وَفيَات الأمهاتx2– معدل

الوَفيَات الخام x3- معدل انتشار سوء التغذية x4- معدل وَفيَات الإناث x5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات (لكل 1000 مولود) حي X7 وكانت النتائج كالآتي:

أ. نتائج اختبار (ديكي فولر) الموسع (ADF) لمتغيرات الدراسة

يوضح الجدول الآتي رقم (4) نتيجة اختبار (ديكي فولر) الموسع (ADF) للسلاسل الزمنية لمتغيرات الدراسة:

جدول رقم (4) نتائج اختبار جذر الوحدة ADF للسلاسل الزمنية لمتغيرات الدراسة

المتغيرات	القيمة المحسوبة لإحصائية اختبار ديكي- فولر	القيم الحرجة عند مستوى معنوية 0.05	مستوى الاستقرار
yمعدل الخصوبة الكلي	-12.31741	-3.632896	الفرق الاول
x1معدل وَفيَات الرضع	-4.051131	-3.012363	الفرق الاول
x2عدد وَفيَات الأمهات	-3.852084	-3.622033	الفرق الاول
x3معدل الوَفيَات الخام	-12.24833	-3.632896	الفرق الاول
x4معدل انتشار سوء التغذية	-5.211142	-3.644963	الفرق الاول
x5معدل وَفيَات الإناث	-4.494540	-3.632896	الفرق الثاني
معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000x6 مولود) حي	-4.388561	-3.673616	الفرق الاول
X7نسبة وَفيَات الأمهات لكل 1000 مولود حي	-5.132712	-3.004861	الفرق الثاني

المصدر: من إعداد الباحثيْن بالاعتماد على نتائج البرنامج الإحصائي (Eviews9)

يتضح من الجدول أعلاه رقم (4) أن قيم (T) المحسوبة لإحصائية اختبار (ديكي فولر) الموسع للسلاسل الزمنية لمتغيرات الدراسة بعد أخذ الفروق الأولى أكبر من القيم الحرجة عند مستويات المعنوية (0.0.05) وعليه يمكن رفض فرض العدم الذي ينص أن السلسلة الزمنية غير ساكنة، وقبول الفرض البديل الذي ينص أن السلسلة الزمنية ساكنة، ويمكن القول إن السلاسل الزمنية لمتغيرات الدراسة غير ساكنة في المستويات، وتصبح ساكنة أومستقرة بعد أخذ الفروق الثانية أي إن السلاسل الزمنية لهذه المتغيرات متكاملة من الدرجة الثانية، وهذا ينطبق مع النظرية الاقتصادية.

اختبار التكامل المشترك لـ(جوهانسون) لمتغيرات الدراسة

لاختبار طبيعة علاقة الانحدار بين المتغيرات المستقلة والمتغير التابع هل هي علاقة انحدار حقيقي أوعلاقة انحدار زائف؟ بالإضافة إلى أن السلاسل الزمنية قد تكون غير مستقرة عند أخذ كل سلسلة على حدة، ولكن تكون متكاملة عند أخذها كمجموعة واحدة، بمعنى أن التقلبات التي تحصل في إحدى السلسلتين تؤدي إلى إلغاء التقلبات في السلسلة الأخرى، وبالتالي تصبح السلاسل الزمنية متكاملة، وعليه سيقوم الباحثان باستخدام اختبار التكامل المشترك لاختبار طبيعة علاقة الانحدار بين المتغير التابع معدل الخصوبة الكلي والمتغيرات المستقلة (معدل وَفيَات الرضع x1– عدد وَفيَات الأمهات x2– معدل الوَفيَات الخام x3- معدل انتشار سوء التغذية x4- معدل وَفيَات الإناث x5-

https://doi.org/10.20428/jss.v29i3.2160

معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات لكل 1000 مولود) حي X7 ويوضح الجدول التالي اختبار التكامل المشترك (نسبة الإمكان الأعظم واختبار الأثر) لـ (جوهانسون) لمتغيرات الدراسة:

جدول رقم (5) نتائج اختبار (جوهانسون) للتكامل المشترك بين أهم العوامل الديموغرافية ومعدل الخصوبة الكلي

Hypothesized	Maximum-Eigen Test			Test trace		
No. of CE (S)	Max-Eigen Statistic LR	Percent5 Critical value	Prob.	Trace Statistic	Percent5 Critical value	Prob.
None *	68.74825	52.36261	0.0005	328.5391	159.530	0.0000
At most 1 *	56.82680	46.23142	0.0027	259.7909	125.615	0.0000
At most 2 *	44.83077	40.07757	0.0135	202.9641	95.754	0.0000
At most 3 *	41.55651	33.87687	0.0050	158.1333	69.819	0.0000
At most 4 *	38.03677	27.58434	0.0016	116.5768	47.856	0.0000
At most 5 *	35.27467	21.13162	0.0003	78.54000	29.797	0.0000
At most 6 *	26.72389	14.26460	0.0003	43.26533	15.495	0.0000
At most 7 *	16.54144	3.841466	0.0000	16.54144	3.841	0.0000

المصدر: إعداد الباحثيْن بالاعتماد على نتائج البرنامج الإحصائي(Eviews9)

يتضح من الجدول أعلاه رقم (5) أن القيم المحسوبة لنسبة الإمكان الأعظم (LR) واختبار الأثر أكبر من القيمة الحرجة عند مستوى معنوية (0.05) وعليه نرفض فرضية العدم الثانية التي تنص على عدم وجود أي متجه للتكامل مشترك بين سلاسل المتغيرات (معدل وَفيَات الرضع x1– عدد وَفيَات الأمهات x2– معدل الوَفيَات الخام x3- معدل انتشار سوء التغذية x4- معدل وَفيَات الإناث x5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات لكل 1000 مولود حي X7) ، ومعدل الخصوبة الكلي وقبول الفرض البديل الذي ينص على وجود أكثر من متجه للتكامل مشترك بين سلاسل المتغيرات (معدل وَفيَات الرضع x1– عدد وَفيَات الأمهات x2– معدل الوَفيَات الخام x3- معدل انتشار سوء التغذية x4- معدل وَفيَات الإناث x5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات لكل 1000 مولود) حي X7، ومعدل الخصوبة الكلي، وبناءً على ما سبق يمكن التأكيد على وجود على توازنية طويلة الأجل بين المتغيرات أهم المتغيرات الديموغرافية المدروسة ومعدل الخصوبة الكلي، مما يعني أن السلاسل الزمنية للمتغيرات لا تبتعد عن بعضهما كثيراً، بحيث تظهر سلوكًا متشابهًا في الأجل الطويل.

تقدير نموذج الانحدار الحصين:

في هذا الجزء يقوم الباحثان بتقدير نموذج الانحدار الحصين (Robust regression) باستخدام أهم ثلاث طرق هي (طريقة MM الحصينة - طريقة S الحصينة- طريقة M الحصينة) لتحليل العلاقة بين المتغيرات المستقلة المتمثلة في (معدل وَفيَات الرضع x1– عدد وَفيَات الأمهات x2– معدل الوَفيَات الخام x3- معدل انتشار سوء التغذية x4- معدل وَفيَات

الإناث5x- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي 6x- نسبة وَفيَات الأمهات لكل 1000 مولود حي) X7، ومعدل الخصوبة الكلي كمتغير تابع، واعتماداً على برنامج التحليل الإحصائي (EVIEWS 9) توصلنا إلى الآتي:

1. **تقدير معالم نموذج الانحدار الحصين:**

يوضح الجدول رقم (6) نتائج معالم نموذج الانحدار الحصين -استخدام الطرق الحصينة (طريقة MM الحصينة - طريقة SS الحصينة- طريقة M الحصينة) -بين أهم العوامل الديموغرافية ومعدل الخصوبة الكلي بالطرق الحصينة

جدول رقم (6) نموذج الانحدار الحصين بين أهم العوامل الديموغرافية ومعدل الخصوبة الكلي بالطرق الحصينة

S-Estimation Robust			MM-Estimation Robust			M-Estimation Robust			V
Prob	StatisticZ	Parameter	Prob.	StatisticZ	Parameter	Prob.	StatisticZ	Parameter	
0.00	-32.80266	-4.364549	0.000	-13.34565	-0.088004	0.000	-16.0501	-3.997234	**C**
0.000	-26.22330	-0.094877	0.000	-5.319264	-0.000409	0.000	-12.9003	-0.087363	**X1**
0.000	-13.00772	-0.000549	0.000	16.91236	1.835139	0.000	-5.30728	-0.000419	**X2**
0.000	26.63656	1.585830	0.000	6.233233	0.030205	0.000	16.44771	1.832882	**X3**
0.000	12.77315	0.033961	0.000	6.329648	0.014073	0.000	6.131979	0.030516	**X4**
0.000	14.72681	0.017965	0.0002	3.740080	0.017436	0.000	5.880265	0.013427	**X5**
0.00	11.07354	0.028324	0.000	-6.905800	-0.010357	0.0003	3.594168	0.017208	**X6**
0.000	-8.310930	-0.006839	0.000	-16.87354	-4.091882	0.000	-6.62005	-0.010196	**X7**
	0.989477			0.794229			0.846588		**R^2**
	0.014467 (0.000)			0.014467 (0.000)			0.018197 (0.000)		**Scale (Prob.)**

المصدر: إعداد الباحثيْن بالاعتماد على نتائج البرنامج الإحصائي (Eviews9)

بناءً على نتائج نموذج الانحدار الحصين بين المتغيرات (معدل وَفيَات الرضع1x– عدد وَفيَات الأمهات2x– معدل الوَفيَات الخام3x- معدل انتشار سوء التغذية4x- معدل وَفيَات الإناث5x- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي 6x- نسبة وَفيَات الأمهات لكل 1000 مولود) حي X7، ومعدل الخصوبة الكلي كمتغير تابع في الجمهورية اليمنية في الفترة من (1990-2014) والموضحة في الجدول رقم (6) قام الباحثان بمقارنة وتحليل نتائج مقدرات النموذج الحصين باستخدام الثلاث الطرق الحصينة كالآتي:

https://doi.org/10.20428/jss.v29i3.2160

1- مقدرات طريقة (M-Estimation Robust)

توضح مقدرات طريقة (M-EstimationRobust)، أن جميع المتغيرات المستقلة المتمثلة في (معدل وَفيَات الرضعx1– عدد وَفيَات الأمهاتx2– معدل الوَفيَات الخامx3- معدل انتشار سوء التغذيةx4- معدل وَفيَات الإناثx5- معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6- نسبة وَفيَات الأمهات (لكل 1000 مولود حي (X7 تؤثر تأثيراً معنوياً على معدل الخصوبة الكلي، وهذا ما أوضحه اختبار (Z) المصاحب لمعالم النموذج، حيث بلغ مستوى المعنوية لجميع المتغيرات المستقلة (0.000) وهو أقل من المستوى المعتمد في المقارنة (0.05)، من حيث ترتيب تأثير المتغيرات المستقلة المدروسة على معدل الخصوبة الكلي نجد أن معدل الوَفيَات الخام يؤثر في المرتبة الأولى، فيما يأتي في المرتبة الثانية معدل وَفيَات الرضع، أما المرتبة الثالثة يأتي معدل انتشار سوء التغذية، أما معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6 يأتي في المرتبة الرابعة، واحتل معدل وَفيَات الإناثx5 المرتبة الخامسة، أما معدلي نسبة وَفيَات الأمهات لكل 1000 مولود حي X7 وعدد وَفيَات الأمهاتx2 فقد احتلا المرتبة السادسة والسابعة على التوالي. كما تفسر هذه المتغيرات (84٪) من التغيرات التي تحصل في معدل الخصوبة الكلي، بينما (14٪) تعود إلى عوامل أخرى غير مدرجة في النموذج وفقًا لهذه الطريقة.

2- مقدرات طريقة (MM-Estimation Robust)

وفقًا لهذه الطريقة فإن جميع المتغيرات المستقلة المدروسة تؤثر تأثيراً معنوياً على معدل الخصوبة الكلي، وهذا ما أوضحه مستوى المعنوية المصاحب لكل معلمة من معالم النموذج المقدر، ومن حيث ترتيب تأثير المتغيرات المستقلة المدروسة على معدل الخصوبة الكلي نجد أن مؤشر نسبة وَفيَات الأمهات لكل 1000 مولود حي X7 يحتل المرتبة الأولى، فيما يحتل المرتبة الثانية مؤشر عدد وَفيَات الأمهات، فيما احتل معدل الوَفيَات الخامx3 المرتبة الثالثة، أما معدل وَفيَات الإناثx5 ومعدل انتشار سوء التغذية x4 فقد حصلا على المرتبة الرابعة والخامسة على التوالي. أما مؤشر معدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6 فقد حصل على المرتبة السادسة، فيما حصل معدل وَفيَات الأطفال الرضع على المرتبة الأخيرة. وفقًا لهذه الطريقة فإن المتغيرات الديموغرافية المدروسة تؤثر على معدل الخصوبة الكلي بنسبة (79٪) أما (21٪) من التغيرات تعود إلى عوامل غير مدرجة في النموذج.

3- مقدرات طريقة (S-Estimation Robust)

بناءً على هذه الطريقة تؤثر المتغيرات الديموغرافية المدروسة على معدل الخصوبة الكلي بنسبة (98٪) بينما (2٪) من التأثيرات تعود إلى عوامل أخرى من النموذج، فيما احتل المرتبة الأولى في التأثير معدل الوَفيَات الخامx3 يليه في المرتبة الثانية معدل وَفيَات الرضعx1 فيما احتل معدل انتشار سوء التغذيةx4 ومعدل وَفيَات الأطفال دون سن الخامسة (لكل 1000 مولود) حي x6المرتبة الثالثة والرابعة على التوالي. أما معدل وَفيَات الإناثx5 حصل على المرتبة الخامسة، فيما مؤشر نسبة وَفيَات الأمهات لكل 1000 مولود حي X7) على المرتبة السادسة، فيما حصل مؤشر عدد وَفيَات الأمهاتx2 على المرتبة الأخيرة.

2.المفاضلة بين نتائج طرائق الانحدار الحصين:

قام الباحثان في هذا الجزء باستخدام اختبارات تقييم جودة النموذج من الناحية الإحصائية والقياسية والمتمثلة (R-squared-Deviance) من أجل المقارنة بين جودة مقدرات طرائق الانحدار الحصين المتمثلة في (مقدرات طريقة

M-Estimation Robust- مقدرات طريقة MM-Estimation Robust- مقدرات طريقة (S-Estimation Robust والجدول التالي يوضح نتائج اختبارات تقييم جودة النموذجين.

جدول رقم (7) نتائج اختبارات تقييم جودة النموذج

معايير التقييم	M-Estimation Robust	MM-Estimation Robust	S-Estimation Robust
RMSE	0.021851	0.022314	0.024634
MAE	0.015815	0.015988	0.016416
MAPE	0.265346	0.269356	0.276617
Theil Coefficient	0.001718	0.001755	0.001937

المصدر: إعداد الباحثَيْن بالاعتماد على نتائج البرنامج الإحصائي(Eviews9)

يتضح من الجدول رقم (7) أن أفضل طريقة من حيث معايير المفاضلة المتمثلة في (RMSE- MAE-- MAPETheil Coefficient) هي طريقة (M-Estimation Robust) حيث حققت أدنى قيم لمؤشرات المفاضلة، وعليه فإن مقدرات طريقة (M-Estimation Robust) هي مقدرات أكثر دقة وأكثر اتساقًا وفقًا للنموذج المقدر وحجم العينة المحددة في البحث، وليس بالضرورة أن تكون أفضل طريقة.

النتائج:

1- بينت اختبارات القيم الشاذة المتمثلة في (H) Hat-Matrix- Covratio - DFFITS)) أن بيانات الدراسة تعاني من قيم شاذة.
2- أظهرت نتائج تحليل استقرارية السلاسل الزمنية لمتغيرات الدراسة أن السلاسل الزمنية لمتغيرات الدراسة عند الفرق الأول والثاني.
3- أوضحت نتيجة اختبار التوزيع الطبيعي Jarque-Bera)) أن بيانات متغيرات الدراسة لا تتوزع توزيعًا طبيعيًا.
4- تعد أفضل طريقة لتقدير معالم الانحدار الحصين هي طريقة M الحصينة، حيث حققت أفضل قيم لمعايير المفاضلة المتمثلة في (RMSE- MAE-- MAPETheil Coefficient)
5- وجود علاقة انحدار حقيقية بين أهم المتغيرات الديموغرافية المدروسة ومعدل الخصوبة الكلي وفقًا لاختبار التكامل المشترك لـ (جاهنسون).
6- تؤثر المتغيرات الديموغرافية المدروسة على معدل الخصوبة الكلي بنسبة (84٪) بينما (14٪) تعود لعوامل أخرى غير مدرجة في النموذج المقدر.

التوصيات:

1- ضرورة اختبار التوزيع الطبيعي للبيانات، وإجراء اختبارات الكشف عن القيم الشاذة قبل تطبيق نماذج الانحدار.
2- استخدام الطرق الحصينة، وأبرزها نموذج الانحدار الحصين في حالة وجود بيانات شاذة ولا تتوزع توزيعًا طبيعيًا.
3- يفضل استخدام أكثر من طريقة تقدير عند تحليل علاقة الانحدار الحصين بين عدد من المتغيرات بهدف الوصول إلى مقدرات أقل خطأ وأكثر دقة.

https://doi.org/10.20428/jss.v29i3.2160

قائمة المراجع:

المصادر والمراجع العربية

البدري، باسم حزام وصالح، زياد زكي (2010)، استخدام أسلوب طريقة التقدير الحصينة (M) في تقدير معلمات أنموذج الانحدار الخطي لتحليل واقع إنتاج واستهلاك محصولي القمح والشعير في العراق للمدة (1975-2009)، *مجلة الإدارة والاقتصاد*، (82)، الجامعة المستنصرية، العراق.

حسن، تارا أحمد ورضا، مهدي صابر. (2011)، استخدام الانحدار الحصين لا يجاد أنسب نموذج لتمثيل بيانات الأنواء الجوية في مدينة أربيل في الفترة (1998-2010)، *مجلة الإدارة والاقتصاد*، (34) جامعة المستنصرية، العراق.

خطار، جبران عبد الحميد ومحمد، كوريس شهيد واسماعيل، محمد سالم، (2016)، تقدير معممات الانحدار الخطي المتعدد باستخدام الطرائق الحصينة (دراسة مقارنة)، *مجلة القادسية لعلوم الحاسوب والرياضيات 8* (1)، كلية العوم الحاسبات والرياضيات، جامعة القادسية، العراق.

السواعي، خالد محمد، داود حسام علي (2013)، *الاقتصاد القياسي بين النظرية والتطبيق باستخدام EVIEWS7*، دار المسيرة للنشر والتوزيع، الأردن.

شاكر، صالح مؤيد (2009)، تحسين أسلوب Mالحصين في تقدير معلمات نموذج الانحدار الخطي، *المجلة العراقية للعلوم الإحصائية*، (16) جامعة الموصل، العراق.

شعراوي، (2005)، سمير مصطفى، *مقدمة في التحليل الحديث للسلاسل الزمنية*، مركز النشر العلمي جامعة الملك عبد العزيز، المملكة العربية السعودية.

المصادر والمراجع الأجنبية

Abd-Almonem, M. (2015). Robust Methods in Regression Analysis: Comparison and Improvement. *Journal of Science*, 38–53.

Alm, Özlem Gürünlü, (2011), Comparison of Robust Regression Methods in Linear Regression, *Int. J. Contemp. Math. Sciences*, *6*(9), 409– 421.

Belsley,D. et al,(1980). *Regression Diagnostics; Identifying Influence Data and Source of Collinearity*. Wiley, New York.

Brown, J. D., (2014). *Linear Models in Matrix Form.* Springer international Publishing Switzerland. New York

Cetin, M., &Toka, O. (2011). The Comparison of S-estimator and M-estimators in Linear Regression. *Journal of Science, 24*(4), 1–6.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression.* New York: John Wiley

Dehnel, G., (2015), Robust Regression in Monthly Business Survey, *Statistics In Transitions new series*, *16*(1), pp. 137–152

Netter et al, (1996). *Applied Linear Statistical Models*. WCB McGraw Hill. Fourth Edition.

Schumacker, R. E., Monahan, M. P., & Mount, R. E. (2002). A Comparison of OLS and Robust Regression using S-Plus. *Multiple Linear Regression Viewpoints*, *28*(2), 10–13

Wen, Y. W., Tsai, Y. W., Wu, D. B. C., & Chen, P. F. (2013). The Impact of Outliers on Net-Benefit Regression Model in Cost- Effectiveness Analysis. *PLoS ONE*, *8*(6), 1–9

المواقع الإلكترونية

البنك الدولي، المؤشرات الاقتصادية لليمن، مسترجع من http://databank.albankaldawli.org/data
المركزي الوطني للمعلومات، الجمهورية اليمنية. مسترجع من http://www.yemen-nic.info/sectors/economics

https://doi.org/10.20428/jss.v29i3.2160