مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

المقارنة بين طريقة المربعات الصغرى الاعتيادية (ols) وطريقة مقدرات ليو لتقدير انموذج الانحدار الهندسي مع تطبيق عملي

م.د. فائز حامد سلمان
كلية الادارة والاقتصاد، جامعة بغداد
Faiz.H@coadec.uobaghdad.edu.iq

المستخلص :

يوجد الكثير من الاساليب التي تستخدم للتنبؤ وحسب طبيعة البيانات ومن ابرزها نماذج الانحدار وخصوصا النماذج المحولة من التوزيعات الاحصائية التي تستخدم للتنبؤ ,يعد التنبؤ المستند على طرق رياضية رصينة اساسا للتخطيط , واعتماد الجهات ذات العلاقة على التخمين حسب علم الباحث يؤدي الى عدم الثقة بالدقة التنبؤية وظهور مشاكل عند تنفيذ الخطط في المستقبل مما يصاحبه انعكاسات سلبية تتمثل بهدر الوقت والاموال وعدم الاستخدام الامثل للموارد البشرية , وبالتالي خسارة المشروع المنفذ او تنفيذه بصورة لا تلبي الطموح , وهدف البحث الى تقدير معادلة التوزيع الهندسي التي تخضع لها بيانات البحث باستخدام طريقة المربعات الصغرى الاعتيادية وطريقة مقدرات ليو الحصينة , وتحديد الطريقة المثلى واستخدامها في التنبؤ باعداد طلبة الثانوي في محافظة النجف الاشرف للمدة الزمنية (2025-2030) باعتماد بيانات السلسلة الزمنية الممتدة (2005-2022)م ,وتوصل الباحث الى افضلية طريقة (liu=0.50) رغم ان البيانات لاتحتوي على قيم متطرفة او شاذة وتبين ان اعداد طلبة الثانوي في محافظة النجف الاشرف في تزايد مستمر للمدة (2023-2025).

الكلمات المفتاحية : انحدار العد ,التوزيع الهندسي , مقدرات ليو , الحصين , التنبؤ

DOI: https://doi.org/10.36325/ghjec.v20i3.17270

Al-Ghary Journal of Economic and Administrative Sciences Vol. 20 (No.3) 2024 P. 423-437

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

Comparison between (OLS) method and the Liu estimators method for estimating the geometric regression model with a practical application

Faez Hamed Salman
University of Baghdad, College of Administration and Economics
Faiz.H@coadec.uobaghdad.edu.iq

Abstract:

Forecasting based on solid mathematical methods is the basis of planning, and the relevant authorities' reliance on guesswork according to the researcher's knowledge leads to a lack of confidence in predictive accuracy and the emergence of problems when implementing plans in the future, which is accompanied by negative repercussions represented by the waste of time and money and the lack of optimal use of human resources, and thus the loss of the implemented project. Or implementing it in a way that does not meet the ambition. The aim of the research is to estimate the geometric distribution equation using the ordinary least squares method and the Liu parametric estimators method, and to determine the optimal method and use it in predicting the numbers of secondary students in the Najaf Governorate for the period of time (2025-2030) by adopting extended time series data ((2005-2022) AD.

المقدمة :

يوجد الكثير من النماذج الرياضية ومن ابرزها نماذج الانحدار وخصوصا النماذج المحولة من التوزيعات الاحصائية التي تستخدم للتنبؤ ,لذا اختار الباحث استخدام انموذج الانحدار الهندسي للتنبؤ باعداد الطلبة الثانوي في محافظة النجف المقدسة بعد خضوع البيانات للتوزيع الهندسي , ويعد التنبؤ المستند على طرق رياضية رصينة اساسا للتخطيط , واعتماد الجهات ذات العلاقة على التخمين حسب علم الباحث يؤدي الى عدم الثقة بالدقة التنبؤية وظهور مشاكل عند تنفيذ الخطط في المستقبل مما يصاحبه انعكاسات سلبية تتمثل بهدر الوقت والاموال وعدم

DOI: https://doi.org/10.36325/ghjec.v20i3.17270

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

الاستخدام الامثل للموارد البشرية , وبالتالي خسارة المشروع المنفذ او تنفيذه بصورة لا تلبي الطموح ,

هدف البحث : يهدف البحث الى تقدير معادلة التوزيع الهندسي باستخدام طريقة المربعات الصغرى الاعتيادية وطريقة مقدرات ليو الحصينة وجدير بالذكر ان الطرائق الحصينة تتعامل بمرونة عالية مع فرضيات الخطا العشوائي ولايؤثر ذلك على كفاءة المقدرات, اذ انها تستند بشكل رئيسي على البيانات وان نوع البيانات يفسر الشكل الفعلي لمنحنى الانحدار,ويهدف البحث ايضا الى المفاضلة بين نتائج تقدير الطريقتين باستخدام معايير المفاضلة (AIC,BIC,MSE) وتحديد الطريقة المثلى لتقدير معادلة التوزيع الهندسي , ومن ثم التنبؤ باعداد طلبة الثانوي في محافظة النجف الاشرف للمدة الزمنية (2025-2030) باعتماد بيانات السلسلة الزمنية الممتدة ((2005-2022)م .

1- الجانب النظري :

1-1: التوزيع الهندسي (Geometric Distribution)[3],[8]

ينتمي هذا التوزيع الى عائلة التوزيعات المتقطعة ويعد التوزيع مهما في التطبيقات الاحصائية وخصوصا في موضوع الاحصاء السكاني عند دراسة معدلات نمو السكان ومعدلات الوفيات والولادات , والتوزيع الهندسي هو جزء من التوزيع الاحتمالي المتعلق بتجارب بيرنولي ، اذ يمكن ان نعرف دالة الكثافة الاحتمالية للتوزيع الهندسي هو عدد المحاولات الفاشلة لتجربة ما لحين الحصول على اول نجاح تلك التجربة. فإذا كان المتغير (Y) يشير إلى عدد مرات تكرار التجربة ,(θ)يشير إلى احتمال نجاح التجربة, ($1-\theta$) احتمال فشل التجربة و عليه فإن الدالة الاحتمالية لهذا التوزيع ستكون:

$$P(Y,P) = \theta.(1-\theta)^{y} \quad ,0 < P < 1 \,, Y = 0,1,2,\ldots \;, \qquad (1)$$

2-1 [10] :بعض خصائص التوزيع الهندسي :

1- الوسط للمتغير Y هو :

$$M_1(Y) = EY = \frac{\theta}{(1-\theta)} \quad \ldots (2)$$

2- المنوال للمتغير Y هو :

$$M_O(Y)=0 \qquad \ldots(3)$$

3- الوسيط للمتغير Y هو :

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

$$M_e(Y) = \left[\frac{-1}{log_2(1-\theta)}\right] - 1 \quad \dots (4)$$

4- التباين للمتغير Y هو :

$$V(Y) = \frac{\theta}{(1-\theta)^2} \qquad \dots (5)$$

5- الالتواء للمتغير Y هو :

$$S_k(Y) = \frac{2-\theta}{\sqrt{1-\theta}} \quad \dots (6)$$

6- التفلطح للمتغير Y هو :

$$E_x(Y) = 6 + \frac{\theta^2}{1-\theta} \quad \dots (7)$$

7- الدالة التجميعية للمتغير Y هي :

$$F(Y) = 1 - \theta^y \quad , Y = 0,1,2, \dots n \qquad (8)$$

8- الدالة المولدة للعزوم هي :

$$M_{Y^{(t)}} = \frac{\theta}{1-\theta e^t} \quad \dots (9)$$

3-1: تحليل الانحدار : Regression Analysis [4],[1]

يعد اسلوب تحليل الانحدار من الطرق او الاساليب الاحصائية التي تستخدم بكثرة في معظم البحوث بصورة عامة لا نها تصف العلاقة بين المتغيرات على هيئة معادلة رياضية , ويعرف تحليل الانحدار بانه عبارة عن وسيلة او أداة احصائية $(Statistical\ Tool)$ تستخدم لمعرفة العلاقة $(Relationship)$ بين متغير مستقل واحد او اكثر $(Independent\ Variables)$ ومتغير تابع $(dependent\ Variables)$.

4-1: نماذج الانحدار : Regression Models [41],[37],[2]

تنقسم نماذج الانحدار الى قسمين وكالاتي : -

اولا/ انموذج الانحدار الخطي : $Linear\ Regression\ Models$

ويتضمن نوعين هما :-

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

1- انموذج الانحدار الخطي البسيط : $Simple\ Linear\ Regression\ Models$

يحتوي هذا الانموذج متغير مستقل واحد و متغير معتمد الذي يتصل خطيا بالمتغير المستقل , وصيغة الأنموذج الرياضية هي :

$$Y_i = B_0 + B_1 X_i + u_i \qquad \ldots(1) \qquad , i=1,2,\ldots,n$$

2- أنموذج الانحدار الخطي المتعدد: $The\ Multiple\ \ Regression\ Models$

يحتوي هذا الانموذج على اكثر من متغير مستقل واحد بالاضافة الى المتغير المعتمد الذي يتصل خطيا بالمتغيرات المستقلة , وصيغة الانموذج الرياضية هي :

$$Y_i = B_0 + B_1 X_{i1} + B_2 X_{i2} + \cdots + B_p X_{ip} + u_i \qquad \ldots(2)$$

$i=1,2,\ldots,n \qquad , \ j= 1,2,\ldots,p$

اذ ان :

$\boldsymbol{Y_i}$ **:** المتغير المعتمد (التابع).

X_i : المتغيرات المستقلة (التفسيرية).

B_0 :قيمة الحد الثابت.

u_i **:** قيمة الخطا العشوائي.

$B_1, B_2, \ldots, B_P$: معاملات الانحدار الجزئي.

ثانيا / نماذج الانحدار اللاخطية :$Non-linear\ Regression\ Models$

هناك العديد من الظواهر تسلك سلوكا لاخطيا , وقد يكون الانموذج (غير خطي من جهة المتغيرات وخطي بدلالة المعلمات), او يكون الانموذج (غير خطي بدلالة المعلمات وخطي بدلالة المتغيرات), واخيرا (غير خطي بدلالة المعلمات والمتغيرات معا), ويدعى الانموذج الذي يحتوي متغير مستقل واحد انموذج غير خطي بسيط , ويدعى متعدد اذا احتوى عدة متغيرات مستقلة , ومفهوم الانموذج اللاخطي هو ان الانموذج يحتوي على معلمة واحدة او اكثر عند الاشتقاق الجزئي, والاتي انموذج الانحدار الهندسي .

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (20) عدد (3) 2024

1-5: انموذج الانحدار الهندسي . [8][10]

يعد انموذج الانحدار الهندسي احد نماذج الانحدار الغير خطية , ويتم تحويله الى احد النماذج الخطية – اللوغاريتمية $(Log - Linear\ Models)$ وذلك من خلال اخذ اللوغارتم الطبيعي لمعادلة التوزيع الهندسي فانها تتحول الى صيغة خطية .

كما يمكن تعريف الانحدار الهندسي بانه الاسلوب الذي بواسطته يتم نمذجة المتغير التابع كونه متغير استجابة $(Response\ Variable)$ عندما تكون قيم المتغير بصفة قيم معدودة $(Count\ Data)$ او بصفة معدلات $(Rate\ Data)$. وممكن الحصول على معادلة الانحدار الهندسي من التوزيع الهندسي كالاتي :

ليكن المتغير(Y) يتبع التوزيع الهندسي بالمعلمة(θ), وان (X) يمثل المتغير التوضيحي (الزمن) نستطيع كتابة معادلة التوزيع الهندسي بالصيغة الاتية :

$$Y = \theta(1-\theta)^x \qquad \dots(3)$$

باخذ اللوغارتم لطرفي المعادلة نحصل على معادلة خطية من الدرجة الاولى :

$$\ln(Y) = \ln(\theta) + X\ln(1-\theta)$$

وبافتراض الادناه نحصل على معادلة (4) تمثل انموذج الانحدار الخطي التقديري للتوزيع الهندسي وكالاتي :

- $\ln(\theta) = b_0$
- $\ln(1-\theta) = b_1$

$$\ln(y) = b_0 + b_1 x_i + e_i \qquad \dots(4)$$

وفي حالة اخذ الدالة الاسية لطرفي المعادلة (4) نحصل على انموذج الانحدار التقديري الغير خطي للتوزيع الهندسي وكالاتي :

$$y = e^{b_0 + b_1 x_i + e_i} \qquad \dots(5)$$

1-6: شروط تحليل الانحدار الخطي . [6][7]

1- يتوزع الخطأ العشوائي (u_i) توزيعاً طبيعياً بمتوسط يساوي صفراً $E(u_i)=0$ وتباين ثابت يساوي $V(u_i)= \sigma^2$

$$u_i \sim N(0, \sigma^2) \qquad \dots(6)$$

DOI: https://doi.org/10.36325/ghjec.v20i3.17270

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

2- الاخطاء العشوائية المتتالية يجب ان لاتكون مرتبطه مع بعضها البعض في أنموذج الإنحدار الخطي البسيط أي أن :

$$\text{Cov}\,(u_i\ , u_j)=0 \qquad \ldots(7)$$

3- التباين المشترك بين قيم المتغير التوضيحي و بين حدود الخطأ المقابلة له مساوية الى الصفر أي أن:

$$\text{Cov}\,(x_i\ , u_i)=0 \qquad \ldots(8)$$

7-1: طرائق تقدير انموذج الانحدا الهندسي:

هناك الكثير من الطرائق لتقدير انموذج الانحدار الهندسي سنتطرق في هذه الفقرة الى طريقتين وهي :

طريقة المربعات الصغرى الاعتيادية :

يرمز لها (OLS) اختصارا لكلمة Ordinary Least Square وهي من اكثر الطرق استخداما في تقدير المعلمات , وتعتمد هذه الطريقة على مبدا (تصغير مجموع مربعات الأخطاء) فهي تهدف على ايجاد المعلمات التي تجعل مجموع مربعات الخطأ اقل ما يمكن وكالاتي.

$$Y_t = B_0 + B_1 * t_i + u_i$$

اذ ان :

Y_t : القيم الاتجاهية لسلسلة الزمنية t.

b_0 : نقطة تقاطع خط الاتجاه مع المحور الصادي.

b_1 : ميل خط الاتجاه العام .

e_i : الخطأ العشوائي .

$$\hat{u}_i = Y_i - \hat{Y}_i \qquad \ldots(6)$$

$$\hat{u}_i = \sum_{t=1}^{n}\left(Y_i - \hat{B}_0 - \hat{B}_1 t\ \right)^2 \qquad \ldots(7)$$

وباشتقاق المعادلة اعلاه جزئياً بالنسبة الى $\hat{B}_1, \hat{B}_0$ على التوالي ومساواتهما بالصفر نحصل على تقدير غير متحيز للمعلمات $\hat{B}_1, \hat{B}_0$ ، اي ان:

$$\frac{\partial ei}{\partial \hat{B}_0} = -2\sum_{i=1}^{n}\left(Yi - \hat{B}_0 - \hat{B}_1\, t_i\right) = 0 \qquad \ldots(8)$$

Al-Ghary Journal of Economic and Administrative Sciences Vol. 20 (No.3) 2024 PP. 429-437

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

وبتقسيم طرفي المعادلة على حجم العينة n نحصل على تقدير ($\hat{B}_0$) كما في الصيغة التالية :

$$\hat{B}_0 = \bar{y} - \hat{B}_1 \bar{t} \qquad ....(8)$$

$$\frac{\partial \hat{u}_i}{\partial \hat{B}_1} = -2 \sum_{i=1}^{n} \left(Yi - \hat{B}_0 - \hat{B}_1 t_i \right) \times t_i = 0 \qquad ...(9)$$

ومنها نحصل على تقدير معلمة الميل الحدي لخط الاتجاه العام:

$$\hat{B}_1 = \frac{\sum_{t=1}^{n} t y_i - n\bar{t}\bar{y}}{\sum_{t=1}^{n} t^2 - n\bar{t}^2} \qquad ....(10)$$

2- طريقة مقدرات ليو:Liu Estimators Method [19],[14],[13]

وهي من طرق التقدير الحصينة وتستخدم لمعالجة مسألة تضخم تباين معلمات الإنموذج المقدرة اذ انها تتعامل بمرونة عالية مع فرضيات الخطا العشوائي ولايؤثر ذلك على كفاءة المقدرات, اذ انها تستند بشكل رئيسي على البيانات وان نوع البيانات يفسر الشكل الفعلي لمنحنى الانحدار, وان العالم Liu أسس هذه الطريقة عام (1993) م حين أوجدها لإنموذج الإنحدار الخطي , تنشأ هذه الطريقة بذات الأسلوب الذي اتبعه (Liu) مع اختلاف طبيعة نوع إنموذج الانحدار الهندسي كونه إنموذجاً غير خطي عند متغير الأستجابة (y)، عند العودة إلى صيغة الإنموذج الهندسي الواردة في المعادلة (4) :

$$Y^* = b_0 + b_1 X_i + u_i$$

وباعتماد القيد الأتي الذي وضعه (Liu)

$$C'C = (S\underline{\hat{\beta}} - \underline{\beta}^*)'(S\underline{\hat{\beta}} - \underline{\beta}^*) \qquad ...(11)$$

إذ أن:

$C'C$: مقدار الزيادة في متوسط مربعات الخطأ الموزون عند استبدال موجه المعلمات المقدرة بطريقة الإمكان الأعظم ($\underline{\hat{\beta}}$) بموجه المعلمات المقدرة وفقا لطريقة ليو ($\underline{\beta}^*$).

$\underline{\hat{\beta}}$: مقدرات الأمكان الأعظم عند تحويل الإنموذج الى إنموذج خطي

$\underline{\beta}^*$: متجه معلمات الإنموذج المقدرة وفق طريقة (Liu)

S: معلمة التحيز

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

بعد عملية تحويل الإنموذج الى خطي ووضع القيد، يتم أخذ مربع الأخطاء الموزونه للإنموذج المبحوث مع إضافة القيد ومن ثم اشتقاق الناتج بالنسبة الى($\underline{\beta}^*$) وكما يأتي:

$$\underline{C}'W\underline{C} = (Y^* - X\underline{\beta}^*)'\widehat{W}(Y^* - X\underline{\beta}^*) + (s\underline{\hat{\beta}} - \underline{\beta}^*)'(s\underline{\hat{\beta}} - \underline{\beta}^*)$$

$$= \underline{Y}^{*'}\widehat{W}\underline{Y}^* - 2\underline{\beta}^*X'\widehat{W}\underline{Y}^* + \underline{\beta}^{*'}X'\widehat{W}X\underline{\beta}^* + S'\underline{\hat{\beta}}'\underline{\hat{\beta}}S - 2S\underline{\hat{\beta}\beta}^* + \underline{\beta}^{*'}\underline{\beta}^* \quad \ldots(12)$$

بإشتقاق المعادلة (12) بالنسبة لموجه المعلمات($\underline{\beta}^*$) نحصل على:

$$\frac{\partial \underline{c}'W\underline{c}}{\partial \underline{\beta}^*} = -2X'\widehat{W}\underline{Y}^* + 2X'\widehat{W}X\underline{\beta}^* - 2S\underline{\hat{\beta}} + 2\underline{\beta}^* \qquad \ldots(13)$$

وعندمساواة نتيجة الاشتقاق الى الصفر نحصل على مقدرات (Liu) لمعلمات أنموذج الانحدار الهندسي.

$$\underline{\hat{\beta}}_{Liu} = (X'\widehat{W}X + 1)^{-1}(X'\widehat{W}\underline{Y}^* + S\underline{\hat{\beta}}) \qquad \ldots(14)$$

وبما أن:

$$\underline{\hat{\beta}} = \underline{\hat{\beta}}_{ML} \qquad \ldots(15)$$

والتي تم تعريفها سابقاً في الصيغه المرقمة (11) على أنها مقدرات الإمكان الأعظم عند تحويل إنموذج الانحدار الهندسي الى الصيغه الخطية.

$$\underline{\hat{\beta}} = \underline{\hat{\beta}}_{ML} = (X'WX)^{-1}(X'W\underline{Y}^*) \qquad \ldots(16)$$

إذ أن :

$$\left(X'W\underline{Y}^*\right) = (X'WX)\underline{\hat{\beta}}_{ML} \qquad \ldots(17)$$

وعند تعويض قيمة $\left(X'W\underline{Y}^*\right)$ بما يساويها في الصيغة المرقمة (14) نحصل على شكل آخر لمقدرات (liu) لمعلمات أنموذج الانحدار الهندسي.

$$\underline{\hat{\beta}}_{Liu} = [(X'\widehat{W}X + 1)^{-1}\left(X'\widehat{W}X\right)\underline{\beta}_{ML} + S\underline{\hat{\beta}}_{ML}]$$

ومنها نحصل على :

$$\underline{\hat{\beta}}_{Liu} = (X'\widehat{W}X + 1)^{-1}\left(X'\widehat{W}X + SI\right)\underline{\hat{\beta}}_{ML} \qquad \ldots(18)$$

1- 8 : معايير المفاضلة:[17],[9],[5]

1- معيار أكايكي (AIC) : Akaike Information Criterion

Al-Ghary Journal of Economic and Administrative Sciences Vol. 20 (No.3) 2024 PP. 431-437

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

$$AIC = nln\hat{\sigma}_{\epsilon}{}^{2} + 2M \qquad \dots (19)$$

2- معيار شوارتز Schwartz Bayesian criterion :(SBC)

$$\mathrm{SBC} = \mathrm{nln}\hat{\sigma}_{\epsilon}{}^{2} + \mathrm{Mln(n)} \qquad \dots (20)$$

3- معيار حنان – كوين Hannan – Quinn Criterion : (H-Q)

$$H - Q(M) = ln(\sigma_{\epsilon}{}^{2}) + 2MC\ \frac{ln[ln(n)]}{n} \qquad ; C > 2 \qquad \dots (21)$$

وبصورة عامة فان :

$\hat{\sigma}_{\epsilon}{}^{2}$: مقدار التباين وصيغته: $\hat{\sigma}_{\epsilon}{}^{2} = \frac{\sum_{t=0}^{n}(e_i - \bar{e}_i)^2}{n}$

n: تمثل عدد المشاهدات.

M : عدد معلمات الانموذج المختار او رتبة الانموذج.

C : ثابت .

2- الجانب التطبيقي .

2-1: **وصف وتعريف عينة الدراسة :-**

تم الأعتماد على بيانات اعداد طلبة الثانوي السنوية في محافظة النجف الاشرف للمدة الزمنية (2005-2022) لمعادلة الانحدار الهندسي والمتغير المعتمد(Y) يمثل اعداد طلبة الثانوي خلال الزمن والمتغير المستقل (X) يمثل الزمن (بالسنوات) .

الجدول (1) اعداد طلبة الثانوي

ت	السنة	المشاهدات	ت	السنة	المشاهدات
1	2005	56405	10	2014	121560
2	2006	58048	11	2015	128304
3	2007	65263	12	2016	132595
4	2008	72504	13	2017	135614
5	2009	81800	14	2018	138185
6	2010	89480	15	2019	140608
7	2011	94235	16	2020	145469
8	2012	108548	17	2021	167262
9	2013	117436	18	2022	169532

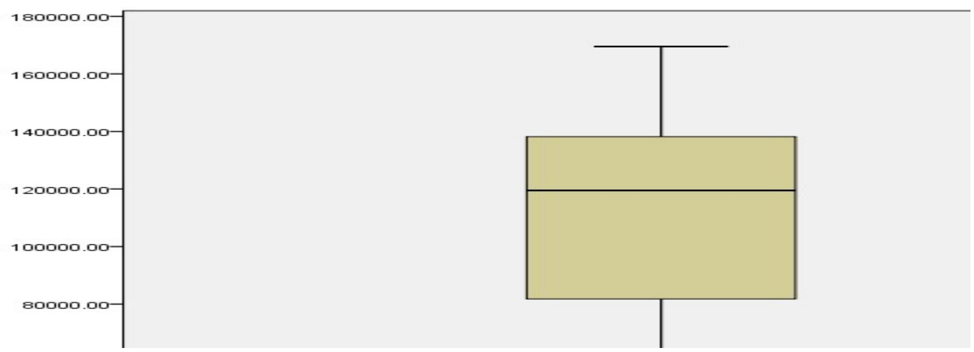
2-2 : ملائمة التوزيع الهندسي للبيانات :

DOI: https://doi.org/10.36325/ghjec.v20i3.17270

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (20) عدد (3) 2024

استخدم اختبار حسن المطابقة (Goodness of Fit) للبيانات باستخدام البرنامج الاحصائي (Easy Fit) وتبين ان البيانات تتبع التوزيع الهندسي بمعلمة ($\theta = 2.5821E - 4$) وكانت قيمة (p-value=0.001).

2-3 : التاكد من وجود قيم متطرفة وتضخم البيانات:



شكل (1) يمثل الرسم الصندوقي لبيانات سلسلة اعداد طلبة الثانوي

من خلال النظر الى الرسم الصندوقي نشاهد لايوجد قيم متطرفة خارج جدران الرسم الصندوقي , وان الوسيط يقسم الصندوق الى منطقتين غير متساويتين , ونشاهد تضخم البيانات في الربع الاول من البيانات اكبر من الربع الثالث .

2-4 : التحليل الاحصائي للبيانات :

تم تحليل البيانات احصائيا باستخدام البرنامج الاحصائي (Gretl) وكالاتي :

اولا : تقديرانموذج الانحدار الهندسي باستخدام طريقة المربعات الصغرى الاعتيادية وطريقة ليو وكالاتي :

1- انموذج الانحدار الهندسي التقديري باستخدام طريقة (ols) هي :

$$\text{lny} = -117.591 + 0.0641500\ \text{x}$$

2- انموذج الانحدار الهندسي التقديري باستخدام طريقة (liu) عند نطاق (0.5)هي:

$$lny = -122.185 + 0.0664189\ x$$

2 - انموذج الانحدار الهندسي التقديري باستخدام طريقة (liu) عند نطاق (0.25)هي:

$$lny = 11.2478 + 0.0551929\ x$$

4- انموذج الانحدار الهندسي التقديري باستخدام طريقة (liu) عند نطاق (0.75)هي:

$$lny = -118.470 + 0.0645857\ x$$

ثانيا : اختبار معنوية العلاقة الخطية المفترضة وكالاتي :

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (20) عدد (3) 2024

1- اختبار معنوية العلاقة الخطية لانموذج الانحدار الهندسي المقدرة باستخدام ols .

$H_0: B_1 = 0$

$H_1: B_2 \neq 0$

يتضح من الجدول (2) قيمة (F=285.572) وبقيمة احتمالية [p-value= 1.26e-011] وهي اقل من مستوى المعنوية (0.05) وعليه نقبل الفرضية البديلة التي تنص ان متغير الزمن يمارس تاثيره في متغير اعداد الطلبة ونسبة (95%).

جدول (2) تحليل التباين

```
Analysis of Variance:

Sum of squares       df      Mean square

Regression              1.99383       1         1.99383
Residual                0.11171      16      0.00698188
Total                   2.10554      17        0.123855

R^2 = 1.99383 / 2.10554 = 0.946945
F(1, 16) = 1.99383 / 0.00698188 = 285.572 [p-value 1.26e-011]
```

2- اختبار معنوية العلاقة الخطية لانموذج الانحدار الهندسي المقدرة باستخدام طريقةliu

- عند نطاق (0.50)

$H_0: B_1 = 0$

$H_1: B_2 \neq 0$

يتبين من الجدول رقم (3) قيمة (F = 194.674) وقيمتها الاحتمالية [p-value= 2.25e-010] وهي اصغر من مستوى المعنوية (0.05) وعليه نقبل الفرضية البديلة التي تنص ان متغير الزمن يمارس تاثيره في متغير اعداد الطلبة ونسبة (92%).

جدول رقم (3)

```
Analysis of Variance:

                     Sum of squares       df      Mean square

  Regression                2.13736        1          2.13736
  Residual                 0.175667       16        0.0109792
  Total                     2.31302       17          0.13606

  R^2 = 2.13736 / 2.31302 = 0.924053
  F(1, 16) = 2.13736 / 0.0109792 = 194.674 [p-value 2.25e-010]
```

- عند نطاق (0.25)

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

يتضح من الجدول (4) قيمة (F=285.572) وبقيمة احتمالية [p-value= 249.766] وهي اقل من مستوى المعنوية (0.05) وعليه نقبل الفرضية البديلة التي تنص ان متغير الزمن يمارس تاثيره في متغير اعداد الطلبة ونسبة (95%).

جدول (4) تحليل التباين

```
Analysis of Variance:

                      Sum of squares       df      Mean square

  Regression                 0.85295        1          0.85295
  Residual                 0.0443949       13       0.00341499
  Total                     0.897345       14        0.0640961

  R^2 = 0.85295 / 0.897345 = 0.950526
  F(1, 13) = 0.85295 / 0.00341499 = 249.766 [p-value 7.24e-010]
```

- عند نطاق (0.75) يتضح من الجدول (5) قيمة (F=385.06) وبقيمة احتمالية [p-value= 1.28e-012] وهي اقل من مستوى المعنوية (0.05) وعليه نقبل الفرضية البديلة التي تنص ان متغير الزمن يمارس تاثيره في متغير اعداد الطلبة ونسبة (96%).

جدول (5) تحليل التباين

```
Analysis of Variance:

                      Sum of squares       df      Mean square

  Regression                   2.021        1            2.021
  Residual                 0.0839764       16       0.00524852
  Total                      2.10497       17         0.123822

  R^2 = 2.021 / 2.10497 = 0.960106
  F(1, 16) = 2.021 / 0.00524852 = 385.06 [p-value 1.28e-012]
```

5- 2 : معايير المفاضلة :

بعد ان تم تقدير نماذج الانحدار الهندسي والتاكد من معنوية العلاقة الخطية , يتم اللجوء الى معايير المفاضلة لاختيار الطريقة التي تحقق افضل انموذج وكالاتي :

جدول (6) معايير المفاضلة

	AIC	BIC	H-Q
ols	−36.3981	−34.61744	−36.15264
l iu 0.25	−40.77205	−39.35595	−40.78714

مجلة الغري للعلوم الاقتصادية والادارية

مجلد (20) عدد (3) 2024

Liu0.50	−28.24988	−26.46914	−28.00434
Liu 0.75	−41.53486	−39.75411	−41.28932

من خلال النظر الى الجدول (6) نلحظ ان طريقة liu عند نطاق (0.50)هي الافضل لكونها تمتلك اصغر معايير المفاضلة .

6 - 2 : التنبؤ :

التنبؤ باعداد طلبة الثانوي لمحافظة النجف الاشرف للمدة الزمنية المستقبلية (2025-2030) باستخدام انموذج الانحدار الهندسي المقدر وفق طريقة (liu=0.50) استنادا الى بيانات السلسلة الزمنية الجدول (1) .

جدول رقم (7) يمثل القيم التنبؤية لاعداد طلبة المرحلة الثانوية للمدة الزمنية (2030-2025)

Years	2025	2026	2027	2028	2029	2030
Predictive Values	222530	237274	252993	269755	287627	306683

4 – النتائج والتوصيات :

1. تبين ان مشاهدت اعداد طلبة الثانوي تخضع للتوزيع الهندسي .
2. ظهرت نماذج الانحدار المقدرة جميعها معنوية .
3. افضلية طريقة انحدار(liu=0.50) لكونها تمتلك اقل معايير مقارنة.
4. القيم المتنبئ بها لوحظ ان هناك تزايد تدريجي في اعداد الطلبة.
5. نوصي باستخدام طرائق اخرى لتقدير معادلة الانحدارالهندسي للتوصل الى افضل نموذج للتنبؤ .
6. نوصي الاستفادة من نتائج البحث واعتمادها لاعداد الخطط المستقبلية من الجهات ذات العلاقة .
7. نوصي بدراسة وضع الخطط المستقبلية وفقا لتخطيط الاستراتيجي .
8. نوصي بمواجهة الزيادة في اعداد الطلبة وفق رؤئ ستراتيجية لتجنب المشاكل في المستقبل

5 – المصادر .

1- بري , عدنان ماجد عبد الرحمن (2003) " تحليل الانحدار الخطي " جامعة الملك سعود .

2- التميمي , زهرة حسن عباس واخرون , (2014)" تحليل الانحدار " مديرية دار الكتب للطباعة والنشر- جامعة البصرة / كلية الادارة والاقتصاد .

3- التوزيع الهندسي / https://marefa.org

4- الراوي , خاشع محمود (1987)" المدخل الى تحليل الانحدار " مديرية دار الكتب للطباعة والنشر - جامعة الموصل / كلية الزراعة والغابات .

5- الصراف،نزار مصطفى وشومان، عبد اللطيف حسن،(2013م)، "السلاسل الزمنية والارقام القياسية,دار الدكتور للعلوم الادارية والاقتصادية ,الطبعة الاولى ,بغداد,العراق.

6- عبد , حميد عبيد (2017) " الاقتصاد القياسي " دار الكتب – كربلاء المقدسة .

7- كاظم , اموري هادي ومسلم , باسم شليبة (2002) "القياس الاقتصادي المتقدم النظرية والتطبيق" بغداد مكتبة دنيا الامل .

8- هرمز , امير حنا (1990) " الاحصاء الرياضي " مديرية دار الكتب للطباعة والنشر – جامعة الموصل/ كلية الادارة والاقتصاد .

9- Akaike, H. (1973),"Information theory and extension of the maximum likelihood principle" , In : B.N. petrov and F.Csaki,eds, 2 nd International symposium on Information Theory ,Academia Kiado, Budapest ,pp.267-281.

10- kadhim Al-Quraishi, O. A. (2021, May). Choosing the best eestimated regression equation for data subject to geometric distribution (Student data as a case study). In Journal of Physics: Conference Series (Vol. 1879, No. 3, p. 032045). IOP Publishing.,

11- Frome, E. L. (1983). The analysis of rates using Poisson regression models. Biometrics, 665-674.

12- Hwang, C. H., & Shim, J. Y. (2008). Semiparametric kernel Poisson regression for longitudinal count data. Communications for Statistical Applications and Methods, 15(6), 1003-1011.

13- Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (Eds.). (1998). Applied regression analysis: a research tool. New York, NY: Springer New York.

مجلة الغري للعلوم الاقتصادية والادارية
مجلد (20) عدد (3) 2024

14- Long, J . S (1997). "Regression Models for Categorical and Limited Independent Variables", SAGE Publicayion Inc, USA.

15- Månsson, K., Kibria, B. G., Sjölander, P., & Shukur, G. (2012). Improved Liu estimators for the Poisson regression model. International Journal of Statistics and Probability, 1(1), 2.

16- Robert.S.Pindyck & Daniel. I. Rubinfeld, (2000), "Econometric models and economic forecasts), New York , McGraw-Hill Book company . Second

17- Chatterjee, S., & Hadi, A. S. (2015). Regression analysis by example. John Wiley & Sons.

18- Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 461-464.

19- Shim, J. (2012). Kernel Poisson regression for mixed input variables. Journal of the Korean Data and Information Science Society, 23(6), 1231-1239.

20- Raudenbush, S. W. (1988). Educational applications of hierarchical linear models: A review. Journal of Educational Statistics, 13(2), 85-116.