

تحليل السلسلة الزمنية باستخدام منهجية بوكس - جنكنز (دراسة تطبيقية على طلاب كلية العلوم في جامعة تشرين)

د. أحمد الوسوف*

د. تميم عليا**

نديم ابراهيم***

(تاريخ الإيداع 23 / 9 / 2020. قُبل للنشر في 21 / 6 / 2021)

□ ملخص □

تحرص وزارة التعليم العالي على تطوير سياسات القبول الجامعي بما ينسجم مع مهارات الطالب وبما يحقق أفضل المدخلات للعملية التعليمية وتحسين كفاءة الخريجين بما يناسب متطلبات سوق العمل ، وهذا التطوير يكون أيضا بمواجهة التحديات التي تواجه قطاع التعليم وخاصة في الأوضاع الراهنة وفي ظل الأحداث التي تشهدها البلاد وأثرها على المؤسسات التعليمية ومنها جامعة تشرين من نواحي متعددة ومنها التغيرات الكبيرة في أعداد الطلاب كان من الضروري اجراء دراسة تهدف الى :

1- معرفة السلوك الذي تسلكه السلسلة الزمنية المُمَثَّلة لأعداد المقبولين وتلك المُمَثَّلة لأعداد الخريجين ومدى استقرارها مع الزمن .

2- ايجاد نموذج إحصائي للتنبؤ بأعداد الطلاب المتوقع قبولهم في كلية العلوم في السنوات الدراسية القادمة وكذلك أعداد الخريجين من الكلية نفسها .

وتوصلت الدراسة الى أن أفضل نموذج لتنبؤ لكل من أعداد المقبولين في كلية العلوم وأعداد الخريجين باستخدام منهجية بوكس - جنكنز هو النموذج $ARIMA(1.1.0)$ ، كما أن الأزمة السورية أثرت سلباً على استقرار أعداد الطلاب مع الزمن .

الكلمات المفتاحية : بوكس - جنكنز _ نماذج أريما _ التنبؤ _ الاستقرارية .

* استاذ مساعد - قسم الإحصاء الرياضي - كلية العلوم - جامعة تشرين - اللاذقية - سورية .

** استاذ - قسم الكيمياء البيئية - المعهد العالي لبحوث البيئة - جامعة تشرين - اللاذقية - سورية .

*** طالب ماجستير - قسم الإحصاء الرياضي - كلية العلوم - جامعة تشرين - اللاذقية - سورية.

Time Series Analysis Using Box-Jenkins Methodology (Applied Study On Students Of The Faculty Of Science at Tishreen University)

Dr. Ahmad ALWassouf*
Dr. Tamim Alia**
Nadeem Ibraheem***

(Received 23 / 9 / 2020. Accepted 21 / 6 /2021)

□ ABSTRACT □

The ministry of Higher education is keen to develop university admission procedures in line with the student's skills , and to achieve the best inputs to the educational process and improve the graduates' qualifications to suit the requirements of the labor market .

This development occurs by addressing the challenge facing the education sector especially in the current situation in the country and its impact on educational institutions including Tishreen University in many respects including significant changes in the number of students so it was necessary to conduct a study aimed at :

1-Knowing the behavior that the time series represented to number of students accepted and those that represented to number of graduates and the extent of their stationary over time.

2-to find a statistical model to predict the number of students expected to be admitted to Faculty of science for the coming year, as well as the number of graduates from the same college .

The study found that the best model for each of the numbers accepted in the Faculty of science and the number graduates using the methodology of the Box-Jenkins is autoregressive model of the first rank ARIMA(1.1.0) .

The Syrian crisis has also adversely affected the stability of student numbers over time.

Keywords : Box-Jenkins - ARIMA Models – Forecasting – Stability .

* Associate Professor , Department of Statistics, Faculty of science, Tishreen University, Lattakia, Syria .

** Professor , Higher Institute of Environmental Research, Tishreen University, Lattakia, Syria.

*** Master Student , Department of Statistics , Faculty of science , Tishreen University, Lattakia, Syria.

مقدمة:

تعمل منظومة التعليم العالي في سورية على تأمين فرص تعليمية للراغبين بها وتأمين الكوادر العلمية المؤهلة والمدرية لتلبية احتياجات سوق العمل ، وهي بذلك تؤدي مهمة أساسية في عملية البناء والتطوير والتغيير التي تضع سورية على خريطة الدول المتقدمة والقادرة على المنافسة عالمياً، و نظراً لزيادة الطلب على التعليم العالي والأزمة التي تمر بها البلاد وغيرها من العوامل التي تؤدي الى ظهور صعوبات تواجه الطلاب وتؤثر على أعداد الخريجين و أعداد المنتسبين الى الكليات في الجامعات السورية ، لابد من التخطيط لمواجهة تلك الصعوبات والتحديات من خلال التعرف على سلوك تلك الأعداد مع الزمن والتنبؤ بها ومعرفة مدى تأثير الأزمة السورية على استقرار تلك الأعداد ، و تعتبر منهجية بوكس - جنكيز أو ما تعرف بنماذج أريما (Autoregressive Integrate Moving Average)ARIMA احدى الطرق الشائعة في التنبؤ بالسلاسل الزمنية ، وتوجد العديد من الدراسات في تحليل السلاسل الزمنية والتنبؤ في شتى المجالات سواء اعتماداً على تلك المنهجية أو غيرها ونذكر منها على سبيل المثال دراسة نقار والعواد في [1] عام 2011 قام الباحثان بإجراء دراسة تطبيقية على تلاميذ الصف الأول الأساسي في سورية و قدّم الباحثان البلخي وجان قرما في [2] عام 2014 دراسة تطبيقية على أسهم بنك التجارة والتمويل الدولي باستخدام نماذج أريما ونماذج الانحدار للسلاسل الزمنية ، كما قام الباحثان الصالح وزينة الأحمد في [3] عام 2016 باختبار نموذج السير العشوائي على عوائد أسهم الشركات المدرجة في سوق دمشق للأوراق المالية ، حيث بينت الدراسة أن الأزمة السورية أثرت سلباً على كفاءة أسهم معظم الشركات المدرسة ، قام الباحث Charafi وآخرون في [4] عام 2016 بتقديم توقعات مناسبة لاتجاه مرض الليشمانيا بمنطقة فارس في إيران للفئات العمرية فوق 15 سنة والأقل من 15 سنة كما قام الباحثان Abonazel و Abd-Elfatah في [5] عام 2019 ببناء نموذج لبيانات الناتج الإجمالي المصري باستخدام نماذج أريما كما قدم الباحث Nyoni في [6] عام 2019 نموذجاً للتنبؤ ببيانات مؤشر أسعار المستهلك في ألمانيا.

تعريف ومفاهيم أساسية :

1- السلسلة الزمنية Time series :

تُعرف السلسلة الزمنية بأنها مجموعة من المشاهدات التي تأخذها ظاهرة عشوائية على فترات زمنية متتابعة متساوية الطول .

2- الاستقرار Stationary :

تُعتبر الاستقرار مفهوم مهم في تحليل السلاسل الزمنية فهو الافتراض الأول الذي يجب اختباره لتطبيق منهجية بوكس_جنكيز والسلسلة الزمنية قد تكون مستقرة أو غير مستقرة ، فالسلاسل المستقرة هي تلك التي لا تتغير خصائصها الإحصائية بتغير الزمن بمعنى أنها تتميز بمتوسط ثابت وتباين ثابت [7] .

من السلاسل الشهيرة التي تحقق شرط الاستقرار سلسلة الضجة البيضاء وهي سلسلة مستقرة ومتماثلة التوزيع بتباين ثابت ومتوسط يساوي الصفر .

2-1 . اختبار جذر الوحدة Unit Root :

يمكن تجزئة سلسلة زمنية (X_t) غير مستقرة بالشكل :

$$X_t = \delta_t + S_t + C_t + e_t$$

حيث δ_t : مركبة الاتجاه العام .

S_t : المركبة الموسمية .

c_t : المركبة الدورية .

e_t : المركبة العشوائية .

إن عدم استقرار السلسلة قد يكون بسبب أن السلسلة لها اتجاه عام حتمي (determinatic trend) أو اتجاه عام عشوائي (stochastic trend) و يمكننا في هذه الحالة جعلها مستقرة بأخذ الفروق (differences)

$$\nabla y_t = y_t - y_{t-1}$$

ومن السلاسل الغير المستقرة سلسلة السير العشوائي Random Walk وهي عبارة عن عملية انحدار ذاتي تحوي

$$y_t = y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0,1) \quad [8]$$

وهناك اختبارات يمكن إجرائها لتحديد ما اذا كانت السلسلة مستقرة أم لا ومنها : [9]

2-2. اختبار ديكي-فولر (Dickey-Fuller) : يعتمد هذا الاختبار على ثلاثة نماذج بسيطة تفترض وجود عملية

عشوائية من نمط انحدار ذاتي من المرتبة الأولى:

النموذج الأول: وهو نموذج دون ثابت أو اتجاه عام:

$$\nabla X_t = \alpha_1 X_{t-1} + e_t$$

النموذج الثاني : وهو نموذج مع ثابت و دون اتجاه عام:

$$\nabla X_t = \alpha_0 + \alpha_1 X_{t-1} + e_t$$

النموذج الثالث: وهو نموذج مع ثابت واتجاه عام:

$$\nabla X_t = \alpha_0 + \alpha_1 X_{t-1} + \delta_t + e_t$$

حيث e_t عملية الضجة البيضاء و δ_t مركبة الاتجاه العام و α_0 ثابت و α_1 جذر الوحدة .

يتم تقدير معالم النماذج السابقة وفق طريقة المربعات الصغرى، وتكون إحصائية الاختبار:

$$t^* = \frac{\hat{\alpha}_1}{SE(\hat{\alpha}_1)} \quad \text{حيث أن } SE(\hat{\alpha}_1) \text{ الخطأ للمعامل } \hat{\alpha}_1 .$$

تكون فرضية العدم والفرضية البديلة بالشكل التالي:

$H_0: \alpha_1 = 0$ السلسلة لها جذر وحدة (السلسلة غير مستقرة).

$H_1: \alpha_1 \neq 0$ السلسلة ليس لها جذر وحدة (السلسلة مستقرة).

إن إحصائية الاختبار لا تتوزع وفق توزيع ستودنت لذلك قام ماكينون Mackinnon بحساب القيم الحرجة لهذا الاختبار ووضعها بجدول سميت باسمه، وتقارن القيمة المحسوبة لاختبار ديكي فولر Dickey Fuller مع القيمة الحرجة من جداول ماكينون فاذا كانت القيمة المحسوبة أصغر من القيمة الحرجة المطلقة عندها نرفض فرضية العدم أي أن السلسلة مستقرة [10] .

2-3 . اختبار ديكي فولر - الموسع ADF (Augmented Dickey and Fuller) :

يعتمد الاختبار على ثلاثة نماذج تفترض وجود عملية عشوائية من نمط انحدار ذاتي من المرتبة p .

النموذج الأول: وهو نموذج دون ثابت أو اتجاه عام:

$$\nabla X_t = \alpha_1 X_{t-1} + \sum_{j=1}^p B_j X_{t-j} + e_t$$

النموذج الثاني : وهو نموذج مع ثابت و دون اتجاه عام:

$$\nabla X_t = \alpha_0 + \alpha_1 X_{t-1} + \sum_{j=1}^p B_j X_{t-j} + e_t$$

النموذج الثالث: وهو نموذج مع ثابت واتجاه عام:

$$\nabla X_t = \alpha_0 + \alpha_1 X_{t-1} + \sum_{j=1}^p B_j X_{t-j} + \delta_t + e_t$$

يعتمد الاختبار على نفس احصائية اختبار ديكي فولر ، كذلك نفس فرضية العدم والفرضية البديلة .

2-4 . نماذج Autoregressive Integrate Moving Average ARIMA :

إن النماذج التي اقترحها بوكس وجنكينز [11] لتمثيل السلاسل الزمنية المستقرة هي :

1- نماذج الانحدار الذاتي Autoregressive AR :

نقول عن $\{X_t\}$ أنها تمثل عملية انحدار ذاتي من المرتبة p إذا أمكن التعبير عنها بالصورة التالية :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

. ϕ_1, \dots, ϕ_p هي معاملات النموذج .

ε_t : الاخطاء العشوائية وهي عملية ضجة بيضاء .

2- نماذج المتوسطات المتحركة Moving Average MA :

نقول عن $\{X_t\}$ أنها عملية متوسطات متحركة من المرتبة q إذا أمكن التعبير عنها بالعلاقة التالية :

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

$\theta_1, \dots, \theta_q$ معاملات النموذج

3- نماذج الانحدار الذاتي والمتوسطات المتحركة ARMA (p,q)

(Autoregressive Moving Average)

هي مزيج بين نماذج الانحدار الذاتي ونماذج المتوسطات المتحركة، تعرف بالعلاقة:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

4- نماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية ARIMA(p,d,q)

(Autoregressive Moving Average Integrated)

عندما تكون السلسلة الزمنية $\{X_t\}$ غير مستقرة يتم أخذ الفرق الأول من الشكل

$$X_t - X_{t-1} = \nabla X_t$$

فإذا حصلنا على سلسلة مستقرة عندئذ يُقال عن السلسلة أنها متكاملة من المرتبة الأولى وفي بعض الأحيان نحتاج

لأخذ الفرق (d) مرة للحصول على الاستقرار عندئذ نقول أن السلسلة متكاملة من المرتبة d أي أن الفرق d حوّل

السلسلة الأصلية الى عملية ARMA المستقرة وتُكتب ARIMA (p,d,q) .

2-5 . منهجية بوكس - جنكينز Box-Jenkins methodology :

اقترح الباحثان بوكس و جنكينز إجراءً تكرارياً لنمذجة السلاسل الزمنية وهذه النمذجة تشمل ثلاث مراحل : [12]

أولاً : مرحلة التعريف (تحديد النموذج) (Identification)

ثانياً : مرحلة التقدير (Estimation)

ثالثاً : الفحص التشخيصي (Diagnostic Checking)

6-2 . معايير المعلومات أكايكي (Akaike Information criterions)

هناك نهج بديل في اختبار مجموعة من النماذج بناءً على اعتبارات سابقة لأقصى قيم p & q والقيام بتقدير كل نموذج واختيار النموذج الذي يقلل من معيار الاختبار ، وهناك مجموعة من معايير الاختبار للنموذج المناسب وربما الأكثر شعبية هو معيار المعلومات أكايكي AIC المُعرّف بالعلاقة :

$$AIC = \ln \left(\frac{\sum_{t=1}^T e^2_t}{T} \right) + \frac{2P}{T}$$

T: طول السلسلة الزمنية والمتمثل بعدد المشاهدات .

والنموذج التي يحوي على قيم صغيرة للمعيار AIC هو النموذج الأفضل . [13]

أهمية البحث وأهدافه:

التخطيط السليم يعتمد على دراسة الواقع والتغيرات التي حدثت أو يمكن أن تحدث في المستقبل استناداً إلى المعطيات والمعلومات التي يمكن الحصول عليها وتحليلها للمساعدة في اتخاذ قرارات سليمة وبالتالي إن الأحداث التي تشهدها البلاد كان لها أثر كبير على المؤسسة التعليمية ومنها جامعة تشرين لذلك كان لابد من تحليل وفهم التغيرات التي حدثت على أعداد الطلاب المسجلين وكذلك أعداد الخريجين خلال السنوات السابقة وحتى الوقت الحاضر والتحقق من إمكانية ربط التغير بتلك الأعداد مع تغير الزمن والتنبؤ بأعداد المُسجلين والخريجين لسنوات قادمة . وكان الهدف من هذا البحث :

1. وضع نموذج إحصائي للتنبؤ بأعداد الطلاب المقبولين في كلية العلوم.
2. وضع نموذج إحصائي للتنبؤ بأعداد الخريجين من كلية العلوم .

طرائق البحث ومواده:

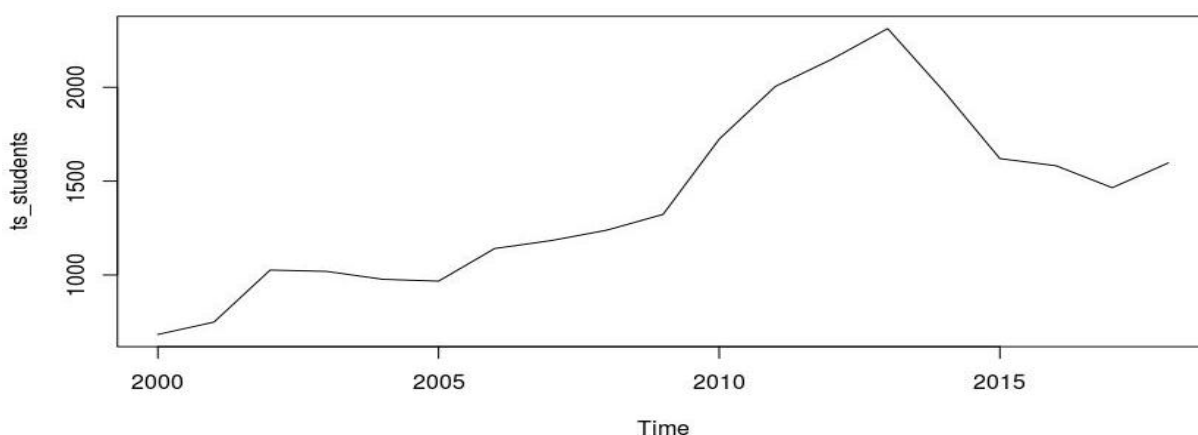
في هذا البحث تم استخدام المنهج التحليلي الإحصائي للسلاسل الزمنية من خلال تطبيق منهجية بوكس-جنكيز على سلسلة أعداد الطلاب المقبولين في كلية العلوم للفترة من 2000 حتى 2018 وأعداد الخريجين منها للفترة من 2000 حتى 2017 ، تم الحصول على البيانات من المكتب المركزي للإحصاء ومن شعبة الدراسات العليا بكلية العلوم .

النتائج والمناقشة:

الجدول (1) يمثل 19 مشاهدة تمتد من عام 2000 الى عام 2018 لأعداد الطلاب المقبولين في كلية العلوم.
- إن نماذج ARIMA يمكن أن توفر نتائج مقبولة عندما تتوفر على الأقل من 16 الى 20 مشاهدة . [14]

جدول (1) : أعداد الطلاب المقبولين في كلية العلوم بجامعة تشرين من عام 2000 الى عام 2018 .

قبل الأزمة		خلال الأزمة	
السنة	الأعداد	السنة	الأعداد
2000/2001	683	2011/2012	2005
2001/2002	749	2012/2013	2149
2002/2003	1026	2013/2014	2313
2003/2004	1019	2014/2015	1979
2004/2005	977	2015/2016	1620
2005/2006	967	2016/2017	1582
2006/2007	1141	2017/2018	1465
2007/2008	1183	2018/2019	1597
2008/2009	1239		
2009/2010	1323		
2010/2011	1724		



الشكل (1) : الرسم البياني لأعداد الطلاب المسجلين في جامعة تشرين للفترة (2000 - 2018).

إن التمثيل البياني للسلسلة الشكل (1) يُبين أن للسلسلة اتجاه عام فهي متزايدة من عام 2000 حتى 2013 ، ونلاحظ أن الأعداد بدأت بالتزايد الكبير من عام 2011 حتى بلغت أعلى قيمة 2313 عام 2013 هذا يدل على وجود عامل أو عدة عوامل كان لها أثر كبير على تلك الأعداد في تلك الفترة قد يكون الوضع الأمني الذي دفع الكثير من الطلاب للتسجيل والدراسة في جامعة تشرين ثم بدأت الأعداد بالتناقص تدريجياً وفي العام الدراسي 2018/2019 حدثت زيادة بالأعداد وهذا قد يدل على أن هذا العامل لا يزال له تأثير .

- في البداية لمعرفة فيما إذا كان للأزمة السورية أثر على ذلك التزايد في الأعداد سنقسم فترة الدراسة الى فترتين : فترة ما قبل الأزمة وفترة خلال الأزمة الجدول (1) ، ثم اختبار فيما إذا كان هناك فرق معنوي بين متوسطي أعداد الطلاب المقبولين في الفترتين حيث بلغ متوسط أعداد الطلاب قبل الأزمة 1093 طالب وبلغ 1838 طالب خلال الأزمة ، بتطبيق اختبار ستودنت للفرق بين متوسطي عينتين

(Independent Samples T-Test) المعطى بالعلاقة التالية : [15]

$$T = (\bar{x}_1 - \bar{x}_2) \mp t_{\alpha/2} \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} ; \quad s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

s_1^2 : تباين العينة الأولى ، s_2^2 : تباين العينة الثانية

نجد أن $P.value=0.000 > 0.05$ (مستوى المعنوية) أي أنه يوجد فرق معنوي بين متوسطي أعداد الفترتين ومنه للأزمة تأثير على متوسط أعداد الطلاب المقبولين في كلية العلوم .

-لايجاد نموذج رياضي لأعداد الطلاب المقبولين في كلية العلوم في البداية تم رسم دالتي الارتباط الذاتي والذاتي الجزئية والتي سنعتمد عليهما في بناء النموذج ، من الشكل (2) نلاحظ أن الارتباط معنوي عند الفجوتين الأولى والثانية لأن معاملات الارتباط الذاتي عندهما تقع خارج حدي 95% لمجال الثقة $\mp \frac{1.96}{\sqrt{n}}$ حيث n حجم العينة

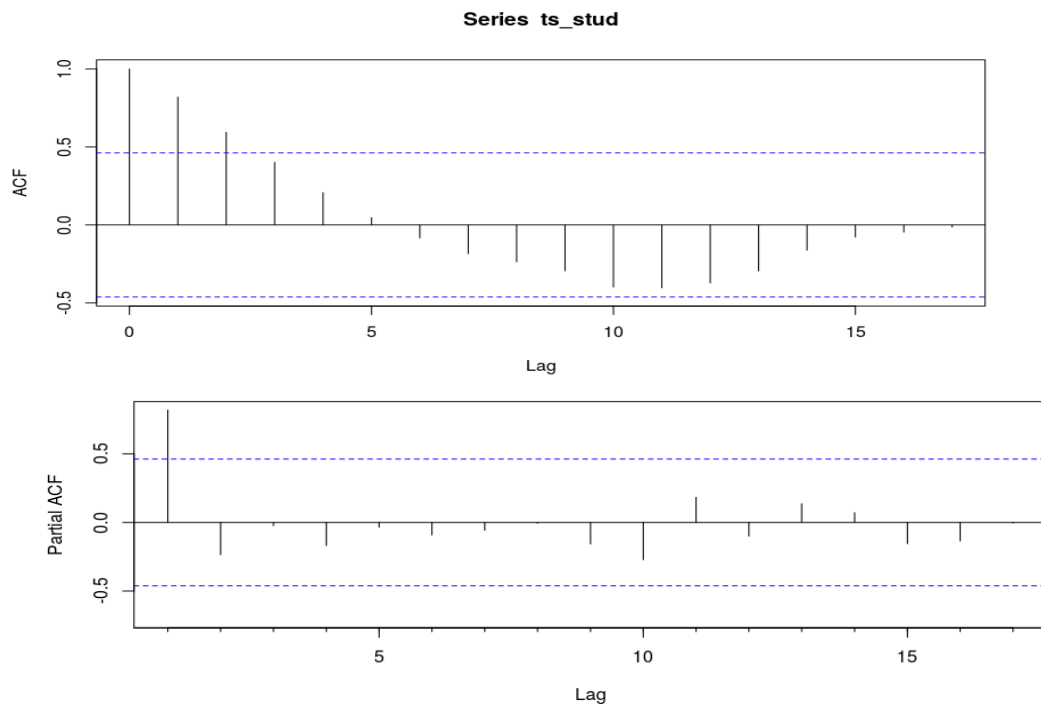
وهذا يدل على أنه يوجد تأثير للقيمة عند الفجوتين الأولى والثانية على القيمة الحالية للسلسلة ، لفحص استقرار السلسلة نطبق اختبار نطيق اختبار ديكي فولر الموسع ونتائج الاختبار الجدول (2) تبين أن :

$P.value = 0.5301 < 0.05$ (مستوى المعنوية) ومنه نقبل فرضية العدم أي أن السلسلة غير مستقرة .

الجدول (2) : نتائج اختبار Augmented Dickey-fuller باستخدام برنامج EViews V.9 .

Null Hypothesis: SER01 has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on AIC, maxlag=3)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.457104	0.5301
Test critical values:		
1% level	-3.886751	
5% level	-3.052169	
10% level	-2.666593	

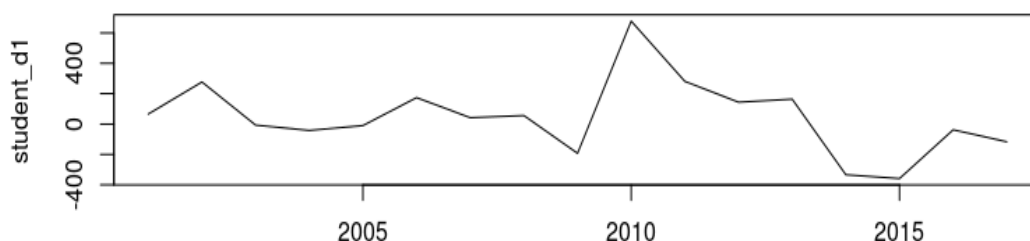
*MacKinnon (1996) one-sided p-values.



الشكل(2) : دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي(PACF).

نأخذ الفرق الأول ، الشكل (3) يوضح شكل السلسلة بعد أخذ الفرق الأول ونلاحظ أن الفروق ساعدت في تثبيت المتوسط عن طريق ازالة التغيرات في مستوي السلسلة وبالتالي التخلص من الاتجاه العام نُشير إلى أنّ سلسلة الفروق تُعطى كما يلي :

$$\nabla X_2 = X_2 - X_1 , \nabla X_3 = X_3 - X_2 , \dots , \nabla X_{18} = X_{18} - X_{17}$$



الشكل(3): الرسم البياني لأعداد الطلاب بعد أخذ الفرق الأول .

نختبر استقرار سلسلة الفروق بتطبيق اختبار ديكي_فولر ونتائج الاختبار التالية باستخدام برنامج EVIEWS9 الجدول (3) تبين أن $p.value = 0.0471 > 0.05$ ومنة السلسلة مستقرة.

الجدول (3) : نتائج اختبار Augmented Dickey-fuller لسلسلة الفروق باستخدام برنامج EVIEWS V.9 .

Null Hypothesis: D(SER01) has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on AIC, maxlag=3)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.098805	0.0471
Test critical values:		
1% level	-3.920350	
5% level	-3.065585	
10% level	-2.673459	

*MacKinnon (1996) one-sided p-values.

تحديد النموذج

يمكن استخدام دالتي الارتباط الذاتي والذاتي الجزئية لتحديد النموذج $ARIMA(p, q)$ أي لتحديد قيمة p & q ، يجب علينا اختبار رتب نماذج $ARIMA(\alpha, \beta)$; $0 \leq \alpha \leq p$, $0 \leq \beta \leq q$ للحصول على أفضل نموذج وللقيام بذلك يجب استخدام معيار المعلومات AIC (Akaike information criterion) [16] ، من الشكل (2) نجد أن دالة الارتباط الذاتي ACF تنقطع بعد الفجوة الثانية $0 \leq q \leq 2$ ، وأن الفجوة الزمنية لدالة الارتباط الذاتي الجزئية تنقطع بعد الفجوة الاولى $0 \leq p \leq 1$ ، ولاختيار الرتب المناسبة نعتمد على معيار المعلومات AIC لأنه يقوم على تدنئة تباين النموذج مقارنة مع زيادة عدد المعلمات المقدره فالرتب التي تعطي أقل قيمة للمعيار AIC هي الأفضل ومن الجدول (4) نلاحظ أن أقل قيمة يأخذها المعيار هي عند الرتب $q=0$ & $P=1$ أي أنه للنموذج الصيغة $ARMA(1,0)$ ولكن بما أن السلسلة استقرت عند الفرق الأول وبالتالي يصبح للنموذج الصيغة $ARIMA(1,1,0)$

$$\nabla X_t = \delta + \phi_1 \nabla X_{t-1} + \varepsilon_t \quad \text{أي الشكل :}$$

الجدول(4): مقارنة النماذج باستخدام معيار المعلومات AIC .

p	q	AIC
0	1	240.97
1	1	242.81
1	0	240.96
0	2	242.87
1	2	244.81
0	3	244.45
1	3	246.39

تقدير معاملات النموذج : إن نتائج التقدير باستخدام برنامج SPSS هي :

		Estimate
عدد_المسجلين-Model	Constant	54.147
	AR Lag 1 Difference	.403 1

فيكتب النموذج بالشكل :

$$\nabla X_t = 54.147 + 0.403\nabla X_{t-1} + \varepsilon_t$$

$$(X_t - X_{t-1}) = 54.147 + 0.403(X_{t-1} - X_{t-2}) + \varepsilon_t$$

$$X_t = 54.147 + 1.403X_{t-1} - 0.403X_{t-2} + \varepsilon_t$$

- يتم اختبار مدى كفاءة النموذج من خلال تحليل الأخطاء ، يجب أن تسلك الأخطاء سلوك الضجة البيضاء تم تطبيق اختبار Ljung-Box [17] للتأكد من كون الأخطاء مستقلة ويختبر الفرضية التالية :

$$H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0 \quad (\text{سلسلة البواقي غير مرتبطة ذاتياً})$$

$$H_1: \rho_i \neq 0 ; 1 \leq i \leq k \quad (\text{سلسلة البواقي مرتبطة ذاتياً})$$

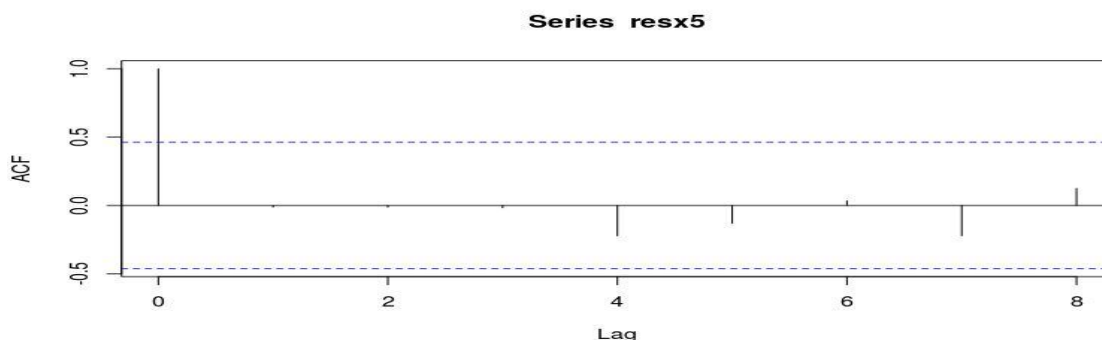
وتظهر نتائج الاختبار للأخطاء الجدول (5) أن $P.value = 0.9584 < 0.05$ (مستوى المعنوية) ومنة نقبل فرضية العدم أي لا يوجد ارتباط ذاتي بين الأخطاء ، وأيضا يمكن ملاحظة ذلك من الرسم البياني الشكل (4) لدالتي الارتباط الذاتي للأخطاء حين أنه جميع الفجوات الزمنية غير داله احصائيا فهي تقع داخل حدي الثقة

$$\pm \frac{1.96}{\sqrt{n}} = \mp \frac{1.96}{\sqrt{19}} = 0.4496$$

Box-Pierce test

data: resx5
X-squared = 0.0027144, df = 1, p-value = 0.9584

الجدول (5) : نتائج اختبار Ljung-Box للأخطاء



الشكل(4): دالة الارتباط الذاتي للأخطاء .

أما التوزيع الطبيعي للأخطاء يظهر من خلال اختبار Shapiro-Wilk ، نتائج الاختبار الجدول (6) تظهر أن P.value $0.05 < = 0.7268$ (مستوى المعنوية) ومنه نقبل فرضية العدم أي أن الأخطاء تتوزع طبيعياً.

Noise residual from -المسجلين_عدد	Statistic	Sig.
Data : resx5	0.9675	0.7268*

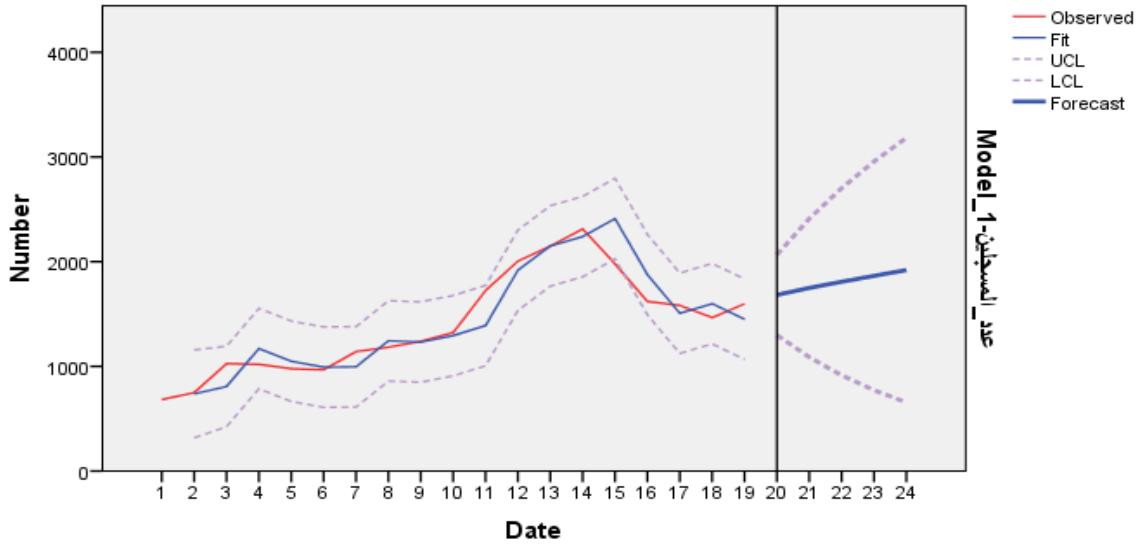
الجدول (6) : نتائج اختبار Shapiro-Wilk للأخطاء .

التنبؤ :

تم استخدام النموذج الناتج للتنبؤ بأعداد الطلاب خلال السنوات القادمة حيث يظهر الجدول (7) نتائج التنبؤ لخمس سنوات قادمة ومقارنتها مع طريقة هولت الخطية :

الجدول(7) : التنبؤات للأعوام الخمسة القادمة .

السنوات	ARIMA(1,1,0)	التمهيد الأسي (طريقة هولت الخطية)
2019/2020	1683	1643
2020/2021	1749	1713
2021/2022	1809	1783
2022/2023	1865	1853
2023/2024	1920	1923



الشكل (5): التمثيل البياني لأعداد الطلاب الحقيقية والمنتبأ بها

للمقارنة بين القدرة التنبؤية للنموذجين ونظراً لصغر حجم العينة سنستخدم أول 17 مشاهدة لتقدير النموذجين ومنتبأ بالأعداد للعامين 2019/2020 & 2018/2019 .

الجدول (8) : التنبؤ بالمشاهدتين 2018 & 2019 باستخدام التمهيد الآسي والنموذج $ARIMA(1.1.0)$

السنة	القيمة الفعلية	التنبؤات	
		$ARIMA(1.1.0)$	هولت الخطية
2018/2019	1597	1436	1357
2019/2020	1444	1683	1643

الجدول (8) يوضح القيم المنتبأ بها باستخدام الطريقتين كما أنه يمكن الاعتماد عليهما في التنبؤ بأعداد الطلاب لأن التنبؤات تقع ضمن 95% فترة ثقة التي تُعطى بالعلاقة (1) لكل من المشاهدين [18]

$$(1) \quad \hat{X}_{t+k} \mp Z_{\frac{\alpha}{2}} \sqrt{S^2(1 + \sum_{j=1}^{n-1} \hat{\theta}_j)}$$

S^2 : تباين الخطأ . $\hat{\theta}_j$: معاملات النموذج .

95% فترة ثقة للمشاهدة 2018/2019 هي [1066 , 1834] .

95% فترة ثقة للمشاهدة 2019/2020 هي [1266 , 2067] .

من خلال النتائج إن الاختلاف أو الفرق الحاصل بين القيمتين الفعلية والمقدرة كان نتيجة عوامل عديدة أثرت في أعداد الطلاب المقبولين في الكلية ومن أهمها قرارات وزارة التعليم المتعلقة بالاستيعاب الجامعي وكذلك الأزمة وتأثيراتها على الحالة الاقتصادية والاجتماعية .

يمكن مقارنة القدرة التنبؤية باستخدام المعيارين جذر متوسط الخطأ التربيعي RMSE

(Root Mean Square Error) ومتوسط الأخطاء المطلق النسبي MAPE

(Mean Absolute Percent Prediction Error) كما في الجدول (9) التالي :

الجدول(9): نتائج المقارنة بين طريقة هولت الخطية والنموذج ARIMA(1.1.0) .

	ARIMA(1,1,0)	التمهيد الأسّي
RMSE	186.259	189.540
MAPE	8.965	13.846

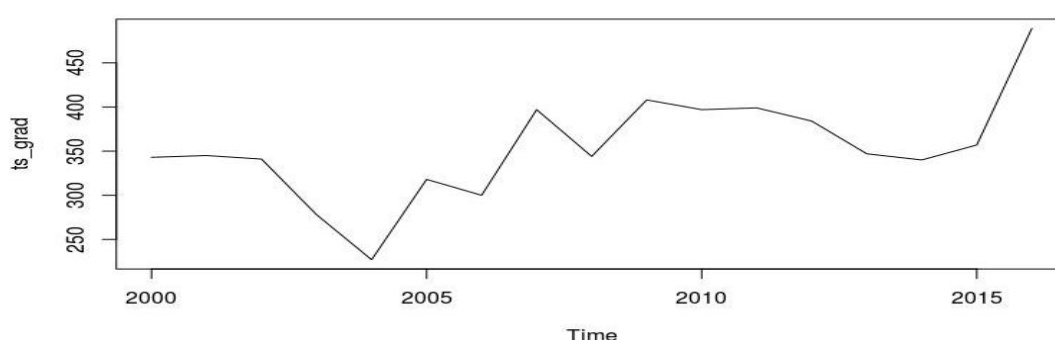
يتضح من الجدول(9) أنّ نموذج ARIMA(1,1,0) يُعطي أفضل تنبؤات من طريقة التمهيد الأسّي لأنّ له أقل قيمة لكل من المعيارين RMSE & MAPE .

أعداد الخريجين :

لدينا البيانات الآتية تمثل أعداد الطلاب الخريجين من كلية العلوم بجامعة تشرين.

الجدول(10) : أعداد الطلاب الخريجين من كلية العلوم قبل وخلال الأزمة من عام 2000 حتى 2017 .

قبل الأزمة		خلال الأزمة	
السنة	أعداد الخريجين	السنة	أعداد الخريجين
2000/2001	343	2011/2012	399
2001/2002	345	2012/2013	384
2002/2003	341	2013/2014	347
2003/2004	278	2014/2015	340
2004/2005	227	2015/2016	357
2005/2006	318	2016/2017	489
2006/2007	300	2017/2018	641
2007/2008	397		
2008/2009	344		
2009/2010	408		
2010/2011	397		



الشكل(6) : الرسم البياني لأعداد الخريجين للفترة (2000 - 2016) .

الرسم البياني للسلسلة الزمنية لأعداد الخريجين الشكل(6) يبين أنّ لها اتجاه عشوائي فهي متزايدة ومتناقصة من العام 2000 الى 2010 وكما أنها تظهر استقرار في الأعداد من العام 2010 الى 2015 لتعود الى التزايد في العام

2016 و 2017 ، إن ثبات أعداد الخريجين في فترة الأزمة بالرغم من تزايد أعداد الطلاب المسجلين يدفعنا للقول أن للأزمة قد يكون لها تأثير على تلك الأعداد ويمكن التأكد من ذلك من خلال تقسيم فترة الدراسة الى فترتين : فترة ما قبل الأزمة وفترة خلال الأزمة الجدول (10) ثم اختبار فيما اذا كان هناك فرق معنوي بين متوسطي الأعداد في الفترتين حيث بلغ متوسط أعداد الخريجين من كلية العلوم قبل الأزمة 331 طالب وبلغ 422 طالب خلال الأزمة، بتطبيق اختبار ستودنت للفرق بين متوسطي عينتين نجد أن $P.value=0.039 > 0.05$ (مستوى المعنوية) أي أنه يوجد فرق معنوي بين متوسطي أعداد الفترتين ومنه للأزمة تأثير على متوسط أعداد الطلاب الخريجين من كلية العلوم

تحديد النموذج :

الخطوة الأولى تتمثل بفحص استقرار السلسلة الزمنية ، وهنا نطبق اختبار ديكي-فولر الموسع (Augmented Dickey-Fuller Test) والجدول التالي يبين أن:

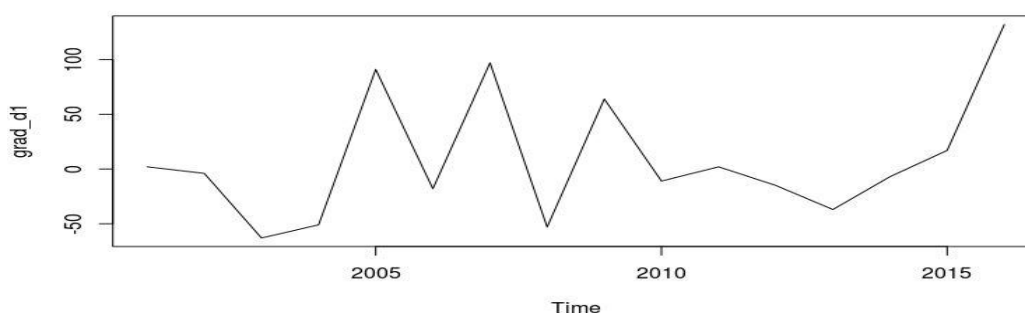
$P.value = 0.5280 < 0.05$ أي أن السلسلة غير مستقرة فكانت النتائج كما يلي الجدول (11) :

الجدول (11) : نتائج اختبار Augmented Dickey-fuller للسلسلة الأصلية لأعداد الخريجين .

Null Hypothesis: SER01 has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on AIC, maxlag=3)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.458076	0.5280
Test critical values:		
1% level	-3.920350	
5% level	-3.065585	
10% level	-2.673459	

*Mackinnon (1996) one-sided p-values.

نأخذ الفرق الأول $\nabla Y_t = Y_t - Y_{t-1}$



الشكل(7): الرسم البياني لأعداد الطلاب بعد أخذ الفرق الأول .

نختبر استقرار سلسلة الفروق الشكل(7) بتطبيق اختبار ديكي- فولر والجدول (12) التالي يبين أن $P.value = 0.0078 > 0.05$ ومنة السلسلة مستقرة.

الجدول (12) : نتائج اختبار Augmented Dickey-fuller لسلسلة الفروق لأعداد الخريجين .

Null Hypothesis: D(SER01) has a unit root

Exogenous: Constant

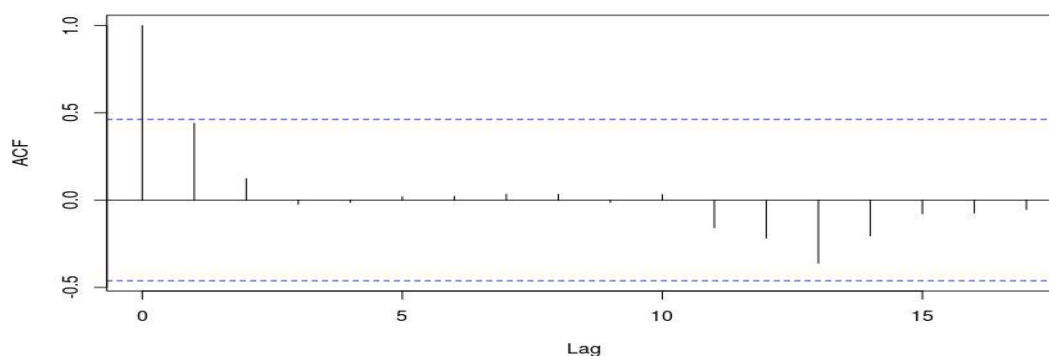
Lag Length: 0 (Automatic - based on AIC, maxlag=3)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-4.095335	0.0078
Test critical values:		
1% level	-3.959148	
5% level	-3.081002	
10% level	-2.681330	

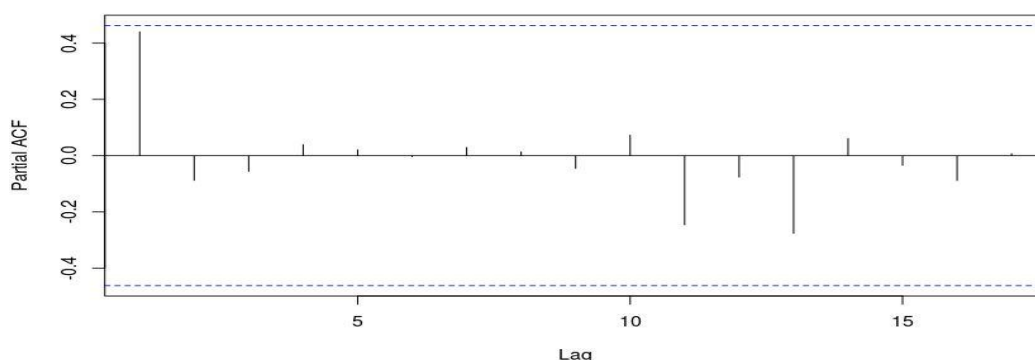
*Mackinnon (1996) one-sided p-values.

الرسم البياني لدالتي الارتباط الذاتي ACF والذاتي الجزئية PACF الشكل (8) يظهر أن جميع معاملات الارتباط تقع ضمن حدي الثقة $\pm \frac{1.96}{\sqrt{n}} = \mp \frac{1.96}{\sqrt{18}} = 0.462$ أي أنه لا يوجد تأثير للملاحظات السابقة على المشاهدة الحالية ولكن من داله الارتباط الذاتي نلاحظ اقتراب الفجوة الاولى من حد الثقة وبالتالي يمكن أن يكون لها تأثير على المشاهدة الحالية وبالاعتماد على معيار المعلومات AIC الرتبة التي تعطي أقل قيمة للمعيار AIC هي الأفضل .

Series ts_gradstud



Series ts_gradstud



الشكل (8) : دالتي الارتباط الذاتي والذاتي الجزئية لأعداد الخريجين .

من الجدول (13) نلاحظ أن أقل قيمة لمعيار AIC هي عندما $P=1, q=0$ أي أن النموذج هو $ARIMA(1,1,0)$

الجدول (13) : مقارنة النماذج باستخدام معيار المعلومات AIC .

p	q	AIC
0	1	194.48
1	1	196.15
2	1	196.37
1	0	194.35
2	0	195.05

تقدير معالم النموذج :

إن نتائج التقدير باستخدام برنامج SPSS هي :

		Estimate
الخريجين-Model	Constant	18.441
	AR Lag 1	.117
	Difference	1

فيكون النموذج :

$$\nabla Y_t = 18.441 + 0.117\nabla Y_{t-1} + \varepsilon_t$$

$$Y_t - Y_{t-1} = 18.441 + 0.117(Y_{t-1} - Y_{t-2}) + \varepsilon_t$$

$$Y_t = 18.441 + 1.117Y_{t-1} - 0.117Y_{t-2} + \varepsilon_t$$

- يتم اختبار مدى كفاءة النموذج من خلال تحليل الأخطاء ، يجب أن تسلك الأخطاء سلوك الضجة البيضاء

نطبق اختبار Ljung-Box للتأكد من كون الأخطاء مستقلة وتظهر نتائج الآتية الجدول (14) أن :

$0.05 < P.value = 0.9407$ ومنة نقبل فرضية العدم التي تقول أن البواقي غير مرتبطة وأيضا يمكن ملاحظة

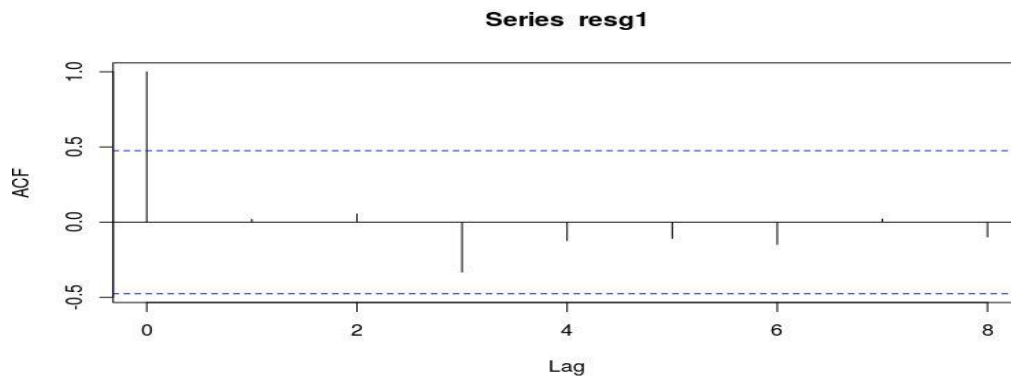
ذلك من الرسم البياني لدالتي الارتباط الذاتي والذاتي الجزئية للأخطاء حين أنه جميع الفجوات الزمنية تقع داخل حدي

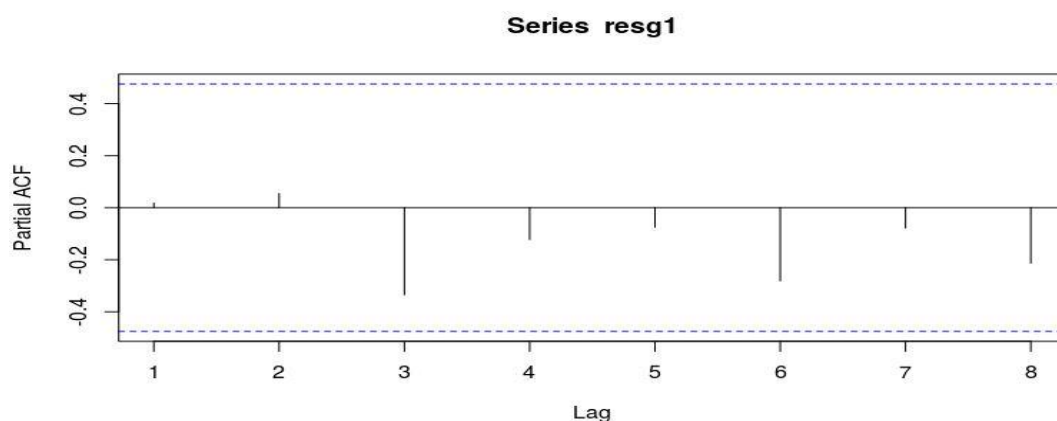
الثقة $\pm \frac{1.96}{\sqrt{n}}$ الشكل (9) .

Box-Pierce test

data: resg1
X-squared = 0.0055273, df = 1, p-value = 0.9407

الجدول (14) : نتائج اختبار Ljung-Box للأخطاء .





الشكل (9): دالتي الارتباط الذاتي والذاتي الجزئية للأخطاء .

أما اختبار التوزيع الطبيعي للأخطاء يظهر من خلال اختبار Shapiro-Wilk ، نتائج الاختبار الجدول (15) تظهر أن $P.value = 0.1717 < 0.05$ (مستوى المعنوية) ومنه نقبل فرضية العدم أي أن الأخطاء تتوزع توزيعاً طبيعياً .

Noise residual from -الخريجين_عدد	Statistic	Sig.
Data : resg1	0.9238	0.1717*

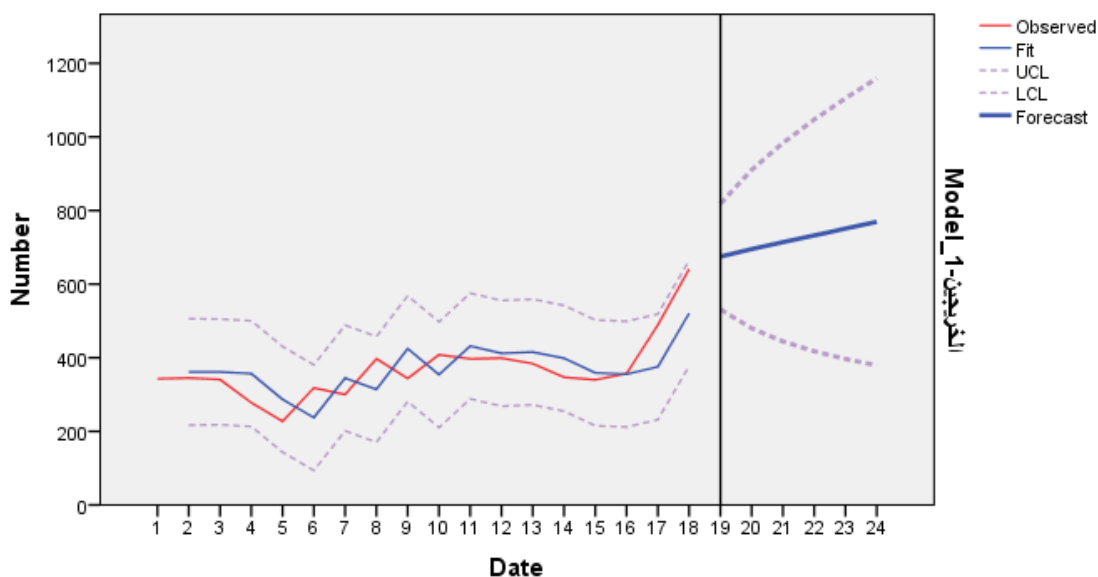
الجدول (15) : نتائج اختبار Shapiro-Wilk للأخطاء .

التنبؤ :

تم استخدام النموذج للتنبؤ بأعداد الخريجين خلال السنوات القادمة حيث يظهر الجدول (16) نتائج التنبؤ لخمس سنوات قادمة باستخدام النموذج الناتج ومقارنتها مع طريقة هولت الخطية :

الجدول(16) : التنبؤات للأعوام الخمسة القادمة .

السنوات	ARIMA(1,1,0)	التمهيد الأسي (هولت الخطية)
2018/2019	675	767
2019/2020	695	927
2020/2021	714	1080
2021/2022	751	1237
2022/2023	769	1394



الشكل (10): التمثيل البياني لأعداد الخريجين الحقيقية والمنتبأ بها

للمقارنة بين القدرة التنبؤية للنموذجين ونظراً لصغر حجم العينة سنستخدم أول 16 مشاهدة لتقدير النموذجين ومنتبأ بالأعداد للعامين 2017/2018 & 2018/2019 الجدول (17) .

الجدول (17) : التنبؤ بالمشاهدتين 2017 & 2018 باستخدام التمهيد الأسّي والنموذج $ARIMA(1.1.0)$

السنة	القيمة الفعلية	التنبؤات	
		$ARIMA(1.1.0)$	هولت الخطية
2017/2018	641	459	566
2018/2019	568	672	767

الجدول (17) يوضح القيم المنتبأ بها باستخدام الطريقتين كما أنه يمكن الاعتماد عليهما في التنبؤ بأعداد الخريجين لأن التنبؤات تقع ضمن 95% فترة ثقة التي تُعطى بالعلاقة (1) لكل من المشاهدين.

95% فترة ثقة للمشاهدة 2017/2018 هي [377 , 664] .

95% فترة ثقة للمشاهدة 2018/2019 هي [532 , 819] .

يمكن مقارنة القدرة التنبؤية باستخدام المعيارين جذر متوسط الخطأ التربيعي RMSE ومتوسط الأخطاء المطلق النسبي MAPE كما في الجدول (18) التالي :

الجدول (18): نتائج المقارنة بين طريقة هولت الخطية والنموذج $ARIMA(1.1.0)$.

	$ARIMA(1,1,0)$	التمهيد الأسّي
RMSE	67.368	66.005
MAPE	14.50	12.792

يتضح من الجدول (18) أن التمهيد الأسّي (طريقة هولت الخطية) أفضل من طريقة بوكس جنكيز لأن له أقل قيمة لكلا المعيارين RMSE & MAPE .

الاستنتاجات والتوصيات:

بينت نتائج الدراسة أنه لا يوجد استقرار في أعداد الطلاب المقبولين في كلية العلوم والخريجين منها ومن خلال الرسم البياني نلاحظ تزايد وتناقص في الأعداد على فترات زمنية أي أنها غير مستقرة وهذا ما أظهره اختبار جذر الوحدة لديكي- فولر، كذلك عندما تم تقسيم فترة الدراسة الى فترتين : فترة قبل الأزمة و خلال الأزمة نتائج اختبار ستودنت أظهرت وجود فرق معنوي بين متوسطي أعداد الطلاب وأعداد الخريجين وهذا يعود لتأثير الأزمة حيث أن أعداد المقبولين ازدادت بشكل كبير مع بداية الأزمة السورية .

إن النموذج الذي تم التوصل اليه لأعداد الطلاب المقبولين والخريجين هو $ARIMA(1,1,0)$ وتم مقارنته مع طريقة التمهيد الأسّي(هولت الخطية) وكانت طريقة بوكس جنكنز الأفضل في حاله الأعداد المقبولين بينما طريقة هولت الخطية كانت الأفضل في حالة الخريجين .

References:

- [1] Naqar, Othman ; Al-Awwad, Munther. The box-jenkins methodology in time series analysis and forecasting (applied study on the number of students in the first grade of basic education in Syria). Damascus University Journal of Economic Sciences and legal.2011. 27(3), 151-125 .
- [2] Al-Balkhi, Rateb; Kazma, Jean. The general trend of stock prices listed on the Damascus Securities Exchange And building a model to predict it .An applied study on the shares of the Bank of Commerce and International Finance using Arima models and regression models for time series. Tishreen University Journal - Economic and Legal Sciences Series. 2014 . 139 - 149 ،(5) 36 .
- [3] Al-Ahmad, Zina; Al-Saleh, Muhannad. Choosing a random walk model on company stock returns Listed on the Damascus Stock Exchange. Al-Baath University Journal.2016. 38(26), .34-11[9]
- [4] SHARAF , MEHDI. *Forecasting the number of zoonotic Cutaneous Irishmaniasis case in south Farts province , Iran using seasonal ARIMA time series method .* Asian pacific journal of tropical medicine.10(1) , 2016 ,79-86.
- [5] ABONAZEL , ABD-ELFTAH . *Forecasting egyption GDP using ARIMA model.* <https://doi.org/10.12988/ref.2019.81023> .2019.
- [6] NYONI . *ARIMA modeling and Forecasting of consumer Price Index (CPI) in Germany .* <https://mpira.ub.uni-muenchen.de/92442/> . 2019.
- [7] MATRIOUSHI. *Hybrid computational intelligence systems based on statistical and neural networks methods for time series forecasting : the case of gold price ,* Lincolu university, United Kingdom . 2011 .
- [8] HEINO BOHN, NIELSEN . *Non-stationary time series and unit Root testing.* http://web.econometrics2-06-L/slides/08_unitroottests_2pp.pdf .2007. Pp.7.
- [9] Bitar, Omar Marwan. A study of forecasting methods using seasonal time series.2017. Master Thesis. Pp.37-39.
- [10] DAMODAR N.GUJARATI .*Basic econometrics.* MCGraw-Hill. 2003 .Pp.816 .
- [11] KIRCHGASSNER , GEBHARD ; WOLTERS , JURGEN. *Introduction to modern time series analysis ,* springer-verlag Berlin Heidelberg.2013 , pp.165-175

- [12] LIU, LON-MU ; HUDAK , GREGORY ; BOX , GEORGE ; MULLER , MERVIN ; TIAO ,GEORGE. *Forecasting and time series analysis using the SCA statistical System* . Scientific computing associates. USA .1994.
- [13] SHUMWAY , ROBERT H ; STOFFER , DAVID S. *Time series analysis and its applications with R examples*. Springer, New York, .2011 PP. 83 – 143 .
- [14] CHRISTODOULOS , CHARISIOS ; MICHALAKELIS , CHRISTOS; VAROUTAS,DIMITRIS. *Forecasting with limited data : combining ARIMA and diffusion models*.Technological forecasting & social change.2010 .
- [15] SHAFER, DOUGLAS; ZHANG, ZHIYI.*beginning statistics*.
<http://lardbucket.org> .2012 . Pp. 466 .
- [16] DRITSAKI, CHAIDO . *FORECASTING Real GDP rate through econometric models : An empirical study from Greece* . International Business and economics . U. S. A . 2015. vol. 3, No. 1. Pp. 13-19 .
- [17] Vandal, Walter. *Time series from an applied point of view and Box_Genghis models*. Arabization: Abdul-Mardi Hamed Azam. Mars Publishing House. 1992. 190-194 .
- [18] Shaarawy, Samir Mustafa. *Introduction to modern analysis of time series*. King Abdulaziz University.2008 .PP .111-140