

# منهجية Box-Jenkins في تحليل السلاسل الزمنية والتنبؤ دراسة تطبيقية على أعداد تلاميذ الصف الأول من التعليم الأساسي في سورية

الدكتور منذر العواد

الدكتور عثمان نقار

كلية الاقتصاد

جامعة دمشق

## الملخص

تولى الجمهورية العربية السورية الاهتمام والرعاية اللازمة لمنظومة التعليم في الدولة، إذ يكفل القانون حق التعليم لكل مواطن، وهو إلزامي ومجاني في مرحلة التعليم الأساسي. هذا فضلاً عن أن معدل النمو السكاني المرتفع، وعوامل أخرى مختلفة تؤدي إلى ارتفاع أعداد التلاميذ المنتسبين إلى الصف الأول من التعليم الأساسي ويشكل متسارع جداً. لابداً لمواجهة هذا الواقع من وضع خطط سنوية من قبل المعنيين لمواجهة الالتزامات المتوقعة لكل عام جديد: المدارس، الشعب، المعلمين، المقررات....

هدفت هذه الدراسة إلى وضع نماذج قياسية للتنبؤ بأعداد التلاميذ المتوقع توافدهم إلى الصف الأول تعليم أساسي باستخدام منهجية (بوكس جنكينز) Box-Jenkins، وتوفيق أفضل نموذج من نماذج ARMA و ARIMA .

خلصت الدراسة إلى وضع نموذج يمكن استخدامه في التنبؤ بأعداد التلاميذ، وتم التنبؤ بأعدادهم حتى عام 2015 ، وهذا ما يشكل قاعدة علمية لوضع خطط التعليم والخطط المرتبطة بها .

## المقدمة:

إن مجانية التعليم وإلزاميته التي يضمنها القانون في الجمهورية العربية السورية، ومعدلات النمو السكاني المرتفعة نسبياً 2.4%<sup>1</sup>، وفتوة المجتمع، كل هذه العوامل مجتمعة فضلاً عن عوامل أخرى تؤدي إلى ظهور مشكلة حقيقية في أعداد المنتسبين سنوياً إلى الصف الأول من التعليم الأساسي، وما يرتبط بهذه الأعداد من التزامات يجب التخطيط لها ومواجهتها من خلال التنبؤ بأعداد التلاميذ. يمكن أن تتم عملية التنبؤ بأعداد التلاميذ بأساليب مختلفة: التحليل التقليدي للسلسلة الزمنية، أو النماذج السببية، أو التحليل الحديث<sup>2</sup> للسلسلة الزمنية طريقة (بوكس جنكينز) Box-Jenkins. ولكل أسلوب من هذه الأساليب ميزاته ومساوئه.

يصعب استخدام النماذج السببية في التحليل والتنبؤ بأعداد التلاميذ بسبب العدد الكبير للعوامل المؤثرة في أعداد التلاميذ، وصعوبة إدخال بعض المتغيرات الوصفية كالأمية في المجتمع والفقر. أما فيما يتعلق بالتحليل التقليدي والحديث للسلسلة الزمنية، فسنعقد في هذا البحث بإجراء مقارنة بين الأسلوبين، وقد أكدت الدراسة التي قمنا بها أن الأسلوب الحديث يعطي نتائج أقرب إلى القيم الواقعية.

## أولاً: الإطار العام للبحث:

### 1-1 - أهمية البحث:

تكمن أهمية البحث باستنتاج نموذج قياسي يستخدم للتنبؤ بأعداد التلاميذ المتوقع انتسابهم للصف الأول في كل عام دراسي، وذلك باستخدام منهج التحليل الحديث للسلاسل الزمنية المبني على منهجية (بوكس جنكينز) Box-Jenkins، ومن ثمّ التنبؤ بهذه الأعداد حتى عام 2015.

### 1-2 - أهداف البحث:

هدف البحث إلى:

- اختبار إمكانية تطبيق الأسلوب الحديث في تحليل السلاسل الزمنية في التنبؤ بأعداد التلاميذ.

1 حُصِبَ هذا المعدل باستخدام البيانات السكانية في المجموعة الإحصائية للعام 2009. للفترة الزمنية 2000-2009.

2 مازالت طريقة بوكس جنكينز تصنف مع التحليل الحديث للسلاسل الزمنية في الأدبيات رغم مرور نحو 35 سنة عليها، انظر كتاب:

Kirchgässner G. and Wolters J. (2007) "Introduction to Modern Time Series Analysis", SPRINGER-Verlag, Berlin Heidelberg, p. 3-5.

- وضع نموذج قياسي للتنبؤ بأعداد تلاميذ الصف الأول من التعليم الأساسي.
- التنبؤ بأعداد التلاميذ المتوقع انتسابهم إلى الصف الأول من التعليم الأساسي.

### 1-3- منهجية البحث:

استُخدم المنهج الوصفي التحليلي في إنجاز هذا البحث من خلال الاطلاع على عدد من المراجع باللغة العربية، والفرنسية، والانكليزية) تناولت منهجية (بوكس جنكينز) *Box-Jenkins* في تحليل السلاسل الزمنية، ومن ثمّ تم الحصول على بيانات السلسلة الزمنية لأعداد التلاميذ من المجموعة الإحصائية السورية، وتم التطبيق عليها. واستخدمت الحزمة البرمجية SPSS في تحليلها .

### 1-4- الدراسات السابقة:

يوجد كثير من الدراسات السابقة التي استخدمت طريقة (بوكس جنكينز) في التنبؤ بسلاسل زمنية ذات طبيعة اقتصادية، ومعظم المراجع المذكورة في نهاية البحث تحوي أمثلة تطبيقية على ذلك. إلا أنه لم يتمكن الباحثان من الوصول إلى أية دراسة سابقة تناولت موضوع استخدام منهجية (بوكس جنكينز) *Box-Jenkins* في التنبؤ بأعداد التلاميذ المتوقع انتسابهم إلى الصف الأول من التعليم الأساسي.

كما أن الأساليب المستخدمة من قبل المعنيين بهذا الأمر هي أساليب تقليدية، منها العد المباشر لأعداد التلاميذ المتوقع انتسابهم في العام الجديد من واقع البطاقات العائلة للأسر القريبة من كل مدرسة، وفي أحسن الأحوال تُستخدَم معادلة الاتجاه العام على مستويات أعلى.

### ثانياً: منهجية (بوكس جنكينز) *Box-Jenkins* :

سيعتمد الباحثان عند بناء نموذج التنبؤ بأعداد تلاميذ الصف الأول الابتدائي في الجمهورية العربية السورية على طريقة (بوكس جنكينز) *Box-Jenkins* المقدمة في كتابهم الشهير: *Time Series Analysis Forecasting and Control* الذي نشر عام 1976، والتي انتشرت بحيث أصبحت الطريقة الأكثر استخداماً في التحليل الحديث للسلاسل الزمنية، وهي تقوم على مجموعة من المراحل:

- المرحلة الأولى: فحص استقرار السلسلة الزمنية، وتطبيق التحولات اللازمة لجعلها مستقرة وإن لم تكن كذلك.

- المرحلة الثانية: تعرّف النموذج المناسب من عائلة نماذج ARIMA (Autoregressive integrated moving average).
- المرحلة الثالثة: تقدير النموذج.
- المرحلة الرابعة: فحص النموذج للتحقق من ملاءمته للسلسلة الزمنية - موضوع البحث - وعندما يكون غير ملائماً نعود إلى المرحلة الثانية، وإلا ننتقل إلى المرحلة التالية (الخامسة).
- المرحلة الخامسة: التنبؤ باستخدام النموذج المختار.

سنقوم فيما بعد بإجراء هذه المراحل على السلسلة الزمنية المختارة (أعداد تلاميذ الصف الأول الابتدائي في الجمهورية العربية السورية).

## 2-1 - اختبار (جذر الوحدة) Unit Root:

يعدّ Box-Jenkins السلسلة الزمنية عبارة عن تحقق لـ (سياق عشوائي) Stochastic Process، ومن أجل تطبيق طريقتيهما يجب أن يكون السياق العشوائي المولد للسلسلة الزمنية (مستقراً) Stationary. **تعريف:** نقول عن السياق العشوائي  $X_t$  إنه مستقر من المرتبة الثانية إذا تحققت الشروط الثلاثة الآتية<sup>3</sup>:

$$- \quad \forall t \in Z, EX_t^2 < \infty \quad (t: \text{الزمن}, Z: \text{مجموعة الأعداد الصحيحة})$$

$$- \quad \forall t \in Z, EX_t = m \quad (\mu: \text{التوقع الرياضي وهو مستقل عن الزمن})$$

$$- \quad \forall t \in Z, \forall h \in Z, COV(X_t, X_{t+h}) = g(h)$$

( $\gamma(h)$ : التغاير وهو مستقل عن الزمن، و  $h$ : الفجوة الزمنية بين اللحظتين المأخوذتين).

نادراً ما تكون السلاسل الزمنية التي نتعامل معها مستقرة من المرتبة الثانية حسب التعريف السابق.

3 GOURIEROUX C. et MONFORT A., (1990) "Séries Temporelles et Modèles Dynamiques " Ed. Economica-Paris. p.152

إن عدم الاستقرار الذي يمكن أن نواجهه في السلاسل الزمنية التي تمثل مشاهدات واقعية يأتي من أن هذه السلاسل إما أن تكون من نمط (Trend Stationary) TS، أو من نمط (Difference) DS (Stationary)<sup>4</sup>.

- النوع الأول TS: هي سلاسل غير مستقرة لها معادلة اتجاه عام محددة فضلاً عن سياق عشوائي مستقر توقعه الرياضي يساوي الصفر وتباينه ثابت.

- النوع الثاني DS: هي سلاسل غير مستقرة ذات اتجاه عام عشوائي وتتميز بوجود جذر الوحدة مرة واحدة على الأقل، ومن أجل جعلها مستقرة نقوم بتطبيق مرشح الفروق الأولى.

إن التمييز بين هذين النوعين من السلاسل يكون باستخدام اختبار جذر الوحدة الذي اقترحه (ديكي وفيللر) Dickey and Fuller عام 1979 ثم قاما بتحسينه عام 1981.

### 2-1-1- اختبار Dickey and Fuller البسيط (D.F)<sup>5</sup>:

يعتمد اختبار (D.F) البسيط على ثلاث معادلات بسيطة تفترض وجود سياق عشوائي من نمط اتحدار ذاتي من المرتبة (1) هذه المعادلات هي :

$$I) \quad \Delta X_t = a_1 X_{t-1} + e_t$$

$$II) \quad \Delta X_t = a_0 + a_1 X_{t-1} + e_t$$

$$III) \quad \Delta X_t = a_0 + a_1 X_{t-1} + B_t + e_t$$

إذ إن:

$$\Delta X_t = X_t - X_{t-1} \text{ : معامل الفروق الأولى، أي:}$$

$e_t$ : سياق الضجة البيضاء<sup>6</sup> White Noise Process.

4 HENIN P.Y. (1989), "Bilans et essais sur la non-Stationnarité des séries Macroéconomiques" revue d'économie politique - n5-p. 667,668.

5 Dickey D. and Fuller W.(1979), " Distribution of the estimators for Autoregressive Time Series With a unit Root ", Journal of the American Statistical Association, n74: pp .427-431.

الفرضية التي نختبرها  $H_0 : a_1 = 0$  (وجود جذر وحدة أي عدم استقرار). تُقارن إحصائية الاختبار  $t = \frac{a_1}{SE(a_1)}$  مع القيم النظرية التي وضعها Dickey and Fuller في جدول .

إن اختبار Dickey and Fuller البسيط يقتصر على نماذج انحدار ذاتي من المرتبة (1). وقد قام Dickey and Fuller بتوسيع الاختبار إلى سياقات الانحدار الذاتي من مرتبة أكبر من (1).

### 2-1-2- اختبار Dickey and Fuller الموسع (A.D.F):<sup>7</sup>

يعتمد الاختبار على المعادلات الثلاث الآتية:

$$I) \quad \Delta X_t = a_1 X_{t-1} + \sum_{j=1}^p B_j \Delta X_{t-j} + e_t$$

$$II) \quad \Delta X_t = a_0 + a_1 X_{t-1} + \sum_{j=1}^p B_j \Delta X_{t-j} + e_t$$

$$III) \quad \Delta X_t = a_0 + a_1 X_{t-1} + \sum_{j=1}^p B_j \Delta X_{t-j} + dt + e_t$$

حيث أن  $e_t$ : سياق الضجة البيضاء .

والاختبار الذي يتم هو نفسه في الفقرة السابقة : ( $H_0 : a_1 = 0$ ) وجود جذر وحدة .

### 2-2- نماذج ARIMA:

إن النماذج التي اقترحها Box و Jenkins لتمثيل السلاسل الزمنية المستقرة من المرتبة الثانية هي:

### 2-2-1- سياق الانحدار الذاتي AR(p):

تعريف: نسمي سياق انحدار ذاتي من المرتبة P السياق المستقر ( $X_t, t \in Z$ ) الذي يحقق العلاقة الآتية<sup>8</sup>:

6 سياق الضجة البيضاء: هو سلسلة من المتغيرات العشوائية ( $t \in Z, et$ ) توقعها الرياضي معدوم و غير مرتبطة فيما بينها (أي

إنّ تبايناتها المشتركة معدومة)، ولها التباين نفسه. انظر (GOURIEROUX et MONFORT, 1990, p.153)

7 Dickey D. and Fuller W.(1981) The likelihood Ratio Statistics for Autoregressive Time Series With a unit Root", Econometrica, n49: pp .1057-1072.

$$\forall t \in Z, X_t + \sum (-J_i X_{t-i}) = e_t$$

إذ:  $J_i$ : هي أعداد حقيقية قيمها المطلقة أصغر من الواحد .

$e_t$ : هو سياق الضجة البيضاء تباينه هو  $d^2$  .

يمكن كتابة العلاقات السابقة بالشكل الآتي:

$$(1 - J_1 B - J_2 B^2 - \dots - J_p B^p) X_t = e_t$$

$$\Phi(B) X_t = e_t$$

إذ إن  $B$ : هو معامل التباطؤ .

2-2-2 - سياق الوسط المتحرك (MA(q):

تعريف: نسمي سياق متوسط متحرك من المرتبة  $q$  السياق المستقر  $(X_t, t \in Z)$  الذي يحقق العلاقة الآتية<sup>9</sup>:

$$X_t = e_t - q_1 e_{t-1} - q_2 e_{t-2} - \dots - q_q e_{t-q} \quad \forall t \in Z$$

إذ إن  $q_i$ : أعداد حقيقية قيمها المطلقة أصغر من الواحد .

يمكن كتابة العلاقة السابقة بالشكل:

$$X_t = (1 - q_1 B - q_2 B^2 - \dots - q_q B^q) e_t$$

$$X_t = H(B) e_t \quad \text{أو :}$$

2-2-3 - سياق الانحدار الذاتي والمتوسط المتحرك (ARMA(p,q):

تعريف: نقول عن السياق المستقر  $(X_t, t \in Z)$  إنه يقبل تمثيل ARMA(p,q) إذا كان يحقق

$$\Phi(B) X_t = H(B) e_t \quad \text{العلاقة}^{10}$$

8 Kirchgässner G. and Wolters J. (2007) "Introduction to Modern Time Series Analysis", SPRINGER-Verlag, p.49

9 المرجع السابق p.64

10 SHUMWAY R.H. and STOFFER D.S. (2005) "Time Series Analysis and Its Applications". SPRINGER, p. 93.

إذ إنَّ :

$$J_p \neq 0, q_q \neq 0 -$$

- لكثيري الحدود  $\Phi(B), H(B)$  جذور جميعها أكبر من الواحد.

- ليس لكثيري الحدود  $\Phi(B), H(B)$  جذور مشتركة.

- السياق  $(X_t, t \in Z)$  هو سياق الضجة البيضاء تباينه هو  $d^2$ .

\* السياق العشوائي ARIMA:

تعريف: نقول عن السياق العشوائي غير المستقر  $(X_t, t \in Z)$ : إنه سياق انحدار ذاتي ومتوسط متحرك متكامل من الدرجة  $d$ ، ويكتب ARIMA(p,d,q) إذا كان يحقق المعادلة من الشكل<sup>11</sup>

$$f(B)(1-B)^d X_t = H(B)e_t :$$

إذ إنَّ :  $(1-B)^d = \Delta^d$  هو مرشح الفروق الأولى من الدرجة  $d$ .

بعبارة أخرى: إذا كانت  $X_t$  سياقاً عشوائياً من نمط ARIMA(p,d,q)، فهذا يعني أن السياق العشوائي المستقر  $\{\Delta^d X_t; t \geq 0\}$  هو عبارة عن سياق ARMA(p,q).

2-3- تعرف النموذج المولد للسلسلة الزمنية:

إن مرحلة تعرف السياق العشوائي المولد للسلسلة الزمنية تعدُّ من المراحل الحرجة، إذ نبحث في عائلة نماذج ARMA على النموذج الذي يلائم السلسلة الزمنية التي لدينا. وقد اقترح Jenkins و Box<sup>12</sup> الاعتماد على دالة الارتباط الذاتي، (A.C.F) ودالة الارتباط الذاتي الجزئي (P.A.C.F)

إذ إنَّ :

\* دالة الارتباط الذاتي الجزئي تحدد لنا رتبة السياق AR(p) إذا أصبحت هذه الدالة غير معنوية بعد عدد معين من التباطؤات، يكون عدد التباطؤات المعنوية هو رتبة AR.

11 PANKRATZ A. (1983) "Forecasting with Univariate Box-Jenkins Models". JOHN WILEY & SONS, p.99

12 Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis Forecasting and Control*, 2nd ed., Holden-Day, San Francisco.

\* بينما تحدد لنا دالة الارتباط الذاتي رتبة السياق  $MA(q)$  إذا أصبحت هذه الدالة غير معنوية بعدد عدد معين من التباطؤات، يكون عدد التباطؤات المعنوية هو  $q$  رتبة سياق المتوسط المتحرك.  
\* أمّا إذا كانت قيم كل من (A.C.F)، تتخامد ولا تتعدم بعد عدد معين من التباطؤات فنكون أمام ( )  
ARMA(p,q).

#### 2-4- تقدير النموذج:

إن تقدير معاملات النموذج إذا كان نموذج انحدار ذاتياً لا تطرح أية مشكلة ، حيث يمكن استخدام طريقة المربعات الصغرى، وفي هذه الحالة فإن أي برنامج إحصائي يعطي معاملات الانحدار الخطي المتعدد يفى بالغرض.

أمّا في حالة نموذج ARMA ، فإنّ تقدير المعاملات يصبح معقداً وتوجد عدة خوارزميات مقترحة لتقدير النموذج، فعلى سبيل المثال يمكن استخدام طريقة الإمكانية القصوى، أو طريقة المربعات الصغرى.

وتختلف البرامج الإحصائية فيما بينها بتقدير معاملات النموذج بحسب الطريقة المتبعة، لذلك قد تعطي نتائج متباينة للنموذج نفسه (BENSABER A., 1989).

أمّا بالنسبة إلينا فنستخدم برنامج SPSS، وهو يعتمد على طريقة الإمكانية القصوى في تقدير النموذج.

#### 2-5- التحقق من صحة النموذج:

من الضروري التحقق من أن النموذج المقدر يحقق الفرضيات التي بني عليها ولاسيما ما يتعلق بالحد العشوائي  $e_t$  من حيث استقلال المشاهدات وعدم ارتباطها ذاتياً، وبأنها تخضع للتوزيع الطبيعي.

نعدّ البواقي  $e_t$  إذ  $e_t = X_t - \hat{X}_t$  هي تقدير للحد العشوائي  $e_t$  ونخضعها للاختبارات الآتية:

### 2-5-1- اختبار الاستقلال السلسلي Ljung-Box<sup>13</sup>

الهدف من الاختبار هو التأكد من عدم وجود ارتباط ذاتي للبقاقي، وأن السياق المولد لها هو عشوائي تماماً.

$$\text{الفرضيات: } H_0 : r_1(e_t) = r_2(e_t) = \dots = r_k(e_t) = 0$$

$H_1$  : يوجد على الأقل معامل غير معدوم .

إذ إنَّ  $r_k(e_t)$  : معامل الارتباط الذاتي للبقاقي بمدة تباطؤ  $k$  .

تحسب إحصائية الاختبار من العلاقة :

$$Q = N(N+2) \sum_{k=1}^K \frac{r_k^2(e_t)}{N-k}$$

نقبل  $H_0$  إذا كانت  $Q < c_{0.05}^2(K - (p+q))$

### 2-5-2- اختبار التوزيع الطبيعي للبقاقي:

يمكن استخدام اختبار Jarque-Bera (BERA, A.K., 1981) وبحسب بالشكل الآتي:

$$S = \frac{T}{6} b_1 + \frac{T}{24} (b_2 - 3)^2$$

$$\text{إذ إنَّ: } b_1 = \frac{m_3^2}{m_2^3} \text{ معامل التفرطح}$$

$$\text{معامل التماثل } b_2 = \frac{m_4}{m_2^2}$$

$S$  يتبع توزيع كاي تربيع بدرجتي حرية.

وتكون الفرضيات  $H_0$  : إن توزيع  $e_t$  هو توزيع طبيعي .

<sup>13</sup> Ljung, G.M., and Box G.E.P. (1978) "on a measure of the lack of fit in time Series models". Biometrika, n65:PP.297-303.

$H_1$  إن توزيع  $e_i$  هو توزيع ليس طبيعياً .

يمكن أيضاً استخدام اختبار (كولموغروف) Kolmogorov الذي يعطيه برنامج SPSS ، أو بالاعتماد على شكل المدرج التكراري للبيانات.

ثالثاً: الحالة التطبيقية:

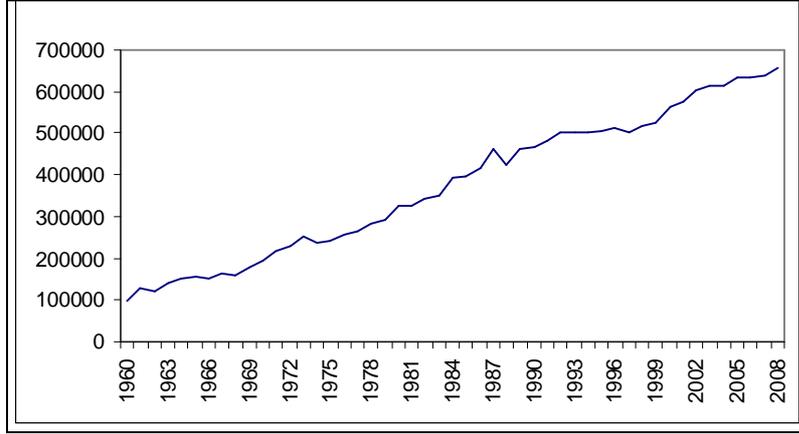
لدينا سلسلة زمنية تمثل أعداد التلاميذ المسجلين في الصف الأول الابتدائي تضم 49 مشاهدة تمتد من عام 1960 إلى عام 2008م. تم الحصول عليها من المجموعة الإحصائية في الجمهورية العربية السورية لسنوات مختلفة.

عدد التلاميذ المسجلين بالصف الأول -تعليم أساسي في الجمهورية العربية السورية لسنوات 1960-2008 (الجدول رقم 1)

السنوات	عدد التلاميذ								
1960	96239	1970	193624	1980	324895	1990	466656	2000	564708
1961	127251	1971	216447	1981	328367	1991	482022	2001	576739
1962	119851	1972	230710	1982	342209	1992	501347	2002	602160
1963	141183	1973	251246	1983	349037	1993	501654	2003	615038
1964	152952	1974	236120	1984	390881	1994	503606	2004	613600
1965	153681	1975	241640	1985	398568	1995	504374	2005	631950
1966	153609	1976	258328	1986	416665	1996	513441	2006	635195
1967	161389	1977	265460	1987	461523	1997	502114	2007	639705
1968	159589	1978	282012	1988	422812	1998	517707	2008	656477
1969	179998	1979	292347	1989	461092	1999	525595		

المصدر: المجموعة الإحصائية السورية للأعوام: 2009 و 1989 و 1976م.

إن الشكل البياني لهذه السلسلة هو كما يظهر في الشكل رقم (1):



الشكل (1) الشكل البياني لتطور أعداد التلاميذ من عام 1960 إلى 2008م

### 3-1- المرحلة الأولى: فحص استقرار السلسلة الزمنية:

في هذه المرحلة نطبق اختبار Dickey and Fuller (اختبار جذر الوحدة) على بيانات السلسلة الزمنية. إن نتيجة الاختبار تظهر وجود جذر الوحدة، إذ كانت إحصائيات الاختبار كما يأتي:

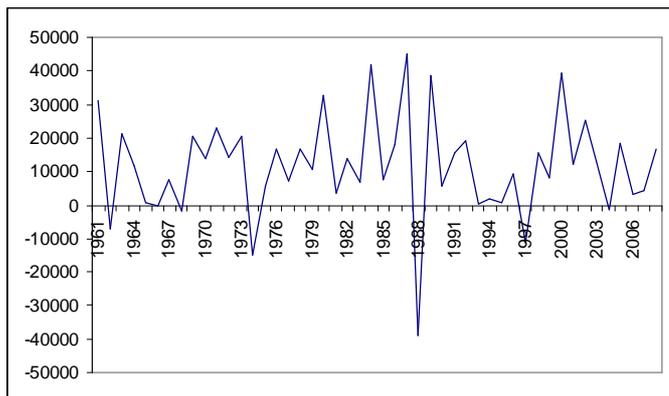
الجدول رقم (2) نتائج اختبار Dickey-Fuller

الاختبار	النموذج	قيمة t المحسوبة	القيم النظرية عند $\alpha=0.05$
D.F.	II	0.353	2.39
A.D.F.	II	0.729	2.39

بالمقارنة نجد أن :  $0.353 < 2.39$  ، وكذلك  $0.729 < 2.39$  ، إذاً تقبل وجود جذر الوحدة.

من أجل جعل السلسلة الزمنية مستقرة نطبق عليها مرشح الفروق الأولى، فيصبح الشكل البياني للسلسلة الناتجة كما يظهر في الشكل رقم (2) ، إذ يبدو من الشكل أنها أصبحت مستقرة<sup>14</sup>.

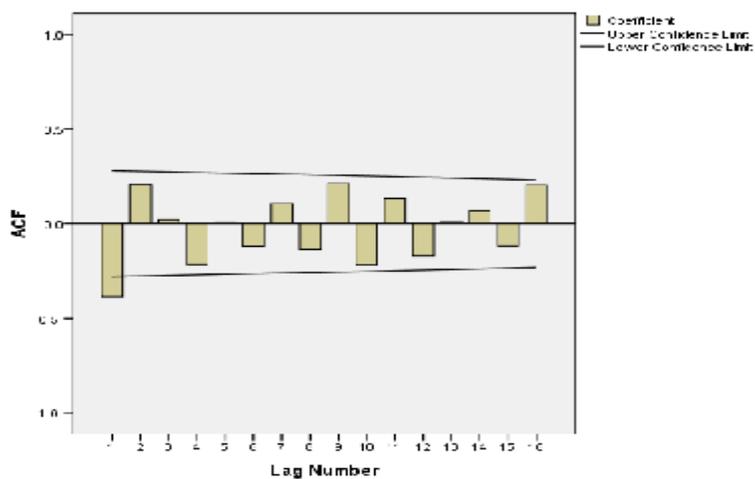
14 من الشكل البياني إذا كانت القيم تتذبذب حول قيمة ثابتة و لا يوجد اتجاه عام نحو الارتفاع أو الانخفاض تكون السلسلة مستقرة.



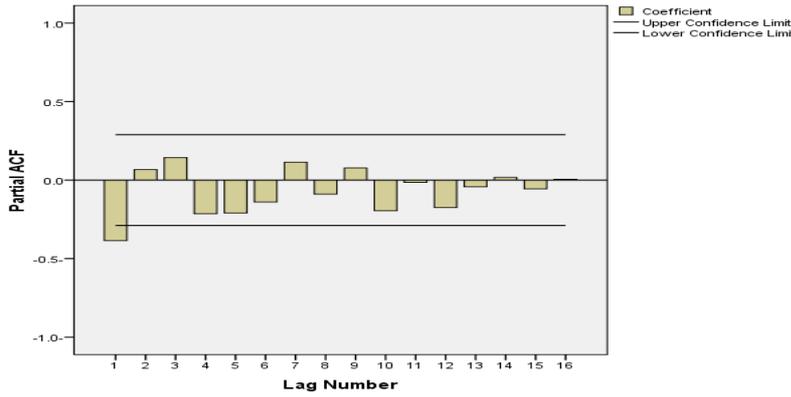
الشكل (2) الشكل البياني لأعداد التلاميذ بعد تطبيق مرشح الفروق الأولى

### 3-2- المرحلة الثانية: تعرّف النموذج المناسب الذي يمثل سلسلة الفروق الأولى:

إن فحص دالة الارتباط الذاتي ACF لسلسلة الفروق الأولى الشكل (3) يقود إلى اقتراح نموذج  $MA(1)$ . أمّا فحص دالة الارتباط الذاتي الجزئي PACF الشكل (4) فهو يقود إلى اقتراح نموذج  $AR(1)$ . وإذا نظرنا إلى الشكلين فإنه يمكن أن نقترح نموذج  $ARMA(1,1)$  :



الشكل (3) دالة الارتباط الذاتي ACF لسلسلة الفروق الأولى



الشكل (4) دالة الارتباط الذاتي الجزئي PACF لسلسلة الفروق الأولى

### 3-3- المرحلة الثالثة و الرابعة تقدير النماذج المقترحة و التحقق من ملائمتها للسلسلة الزمنية:

سنقوم بتطبيق هاتين المرحلتين على كل نموذج من النماذج الثلاثة المقترحة. و النماذج التي يتبين بنتيجة الفحص أنها صالحة لتمثيل السلسلة الزمنية، نقارن فيما بينها من حيث القدرة التنبؤية و نختار أفضلها.

#### 3-3-1- تقدير النموذج AR(1) لسلسلة الفروق الأولى:

وهذا يعني تقدير نموذج ARIMA(1,1,0) للسلسلة الأساسية. قمنا بتقدير النموذج باستخدام الـ 44 مشاهدة الأولى من السلسلة الزمنية من عام 1960 حتى 2003، وباستخدام النموذج المقدر قمنا بالتنبؤ للملاحظات الخمس المتبقية من عام 2004 حتى 2008 وذلك حتى نتأكد من تقييم القدرة التنبؤية للنموذج.

إن نتائج التقدير باستخدام برنامج SPSS كانت كما في الجدول رقم (3):

الجدول رقم (3) نتائج التقدير باستخدام برنامج SPSS

			Estimate	SE	t	Sig.
X-Model_1	X	Constant	11934.170	1599.993	7.459	.000
		AR Lag 1	-.391-	.143	-2.735-	.009
		Difference	1			

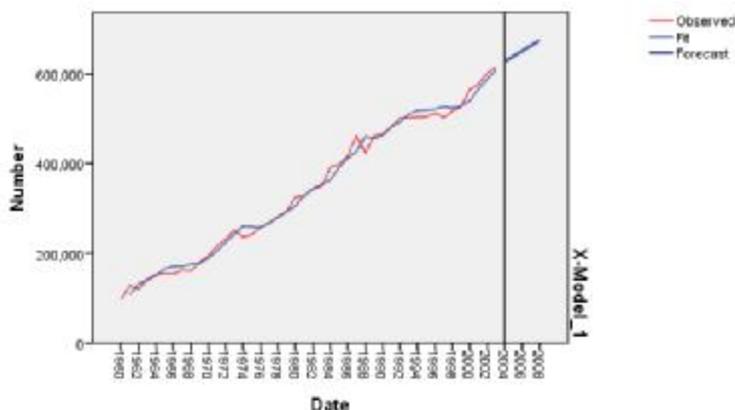
من الجدول (3) نلاحظ أن المعاملات المقدرة معنوية، وهي لا تساوي الصفر، من ذلك نرفض فرضية العدم. والنموذج المقدر هو:

$$(1 - B)(1 + 0.391B)X_t = 11934.17 + e_t$$

أو يكتب بالشكل الآتي:

$$X_t = 11934.17 + 0.609X_{t-1} + 0.391X_{t-2} + e_t$$

والشكل (5) يبين المنحنى البياني للسلسلة الأساسية، و للقيم المقدرة، و للقيم المتنبأ بها:



الشكل (5) الشكل البياني للسلسلة الأساسية وللقيم المقدرة وللقيم المتنبأ بها

3-3-2- التحقق من البواقي:

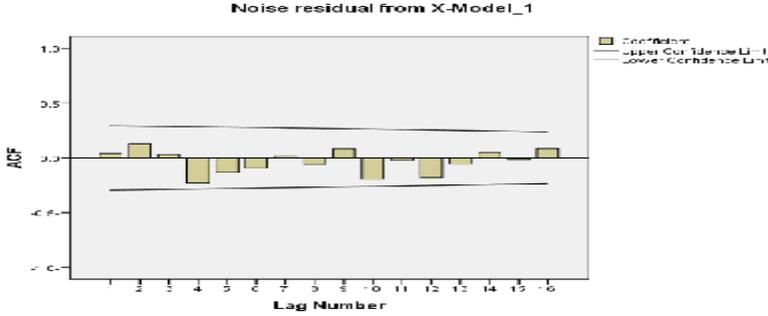
إن الاختبارات التي نطبقها على سلسلة البواقي هي:

أ- اختبار الارتباط السلسلي:

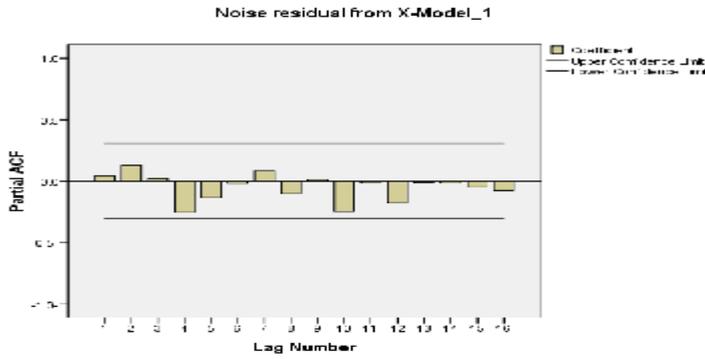
الجدول رقم (4) نتائج اختبار Ljung-Box

Ljung-Box Q(18)		
Statistics	DF	Sig.
11.550	17	.827

إن اختبار Ljung-Box كما يظهر في الجدول يشير إلى قبول فرضية العدم، أي عدم وجود ارتباط ذاتي بين الأخطاء، ويؤكد ذلك دالتا الارتباط الذاتي والارتباط الجزئي للبواقي كما في الشكل (6) و (7):



الشكل (6) دالة الارتباط الذاتي للبواقي



الشكل (7) دالة الارتباط الذاتي الجزئي للبواقي

ب - طبيعة التوزيع الاحتمالي للبواقي:

نستخدم الاختبار اللامعلمي K-S كلوموجوروف-سميرونوف، فنجد أن البواقي تخضع للتوزيع الطبيعي (المعتدل) كما يظهر في الجدول رقم (5) .

الجدول رقم (5) نتائج اختبار Kolmogorov-Smirnov

		Noise residual from X-Model_1
N		43
Normal Parameters <sup>a</sup>	Mean	173.5805
	Std. Deviation	14374.55222
Most Extreme Differences	Absolute	.075
	Positive	.075
	Negative	-.066-
Kolmogorov-Smirnov Z		.490
Asymp. Sig. (2-tailed)		.970
a. Test distribution is Normal.		

إن نتائج الاختبارات المطبقة على البواقي تؤكد صلاحية النموذج المقدر لتمثيل السلسلة الزمنية، ومن ثم إمكانية استخدامه في التنبؤ.

### 3-3-4- تقدير النموذج MA(1) لسلسلة الفروق الأولى:

وهذا يعني تقدير نموذج ARIMA(0,1,1) للسلسلة الأساسية.

إن نتائج التقدير باستخدام برنامج SPSS كانت كما في الجدول رقم (6).

الجدول رقم (6) نتائج التقدير باستخدام برنامج SPSS

			Estimate	SE	t	Sig.
X-Model_2	X	Constant	11926.006	1593.816	7.483	.000
		Difference	1			
	MA	Lag 1	.303	.151	2.009	.051

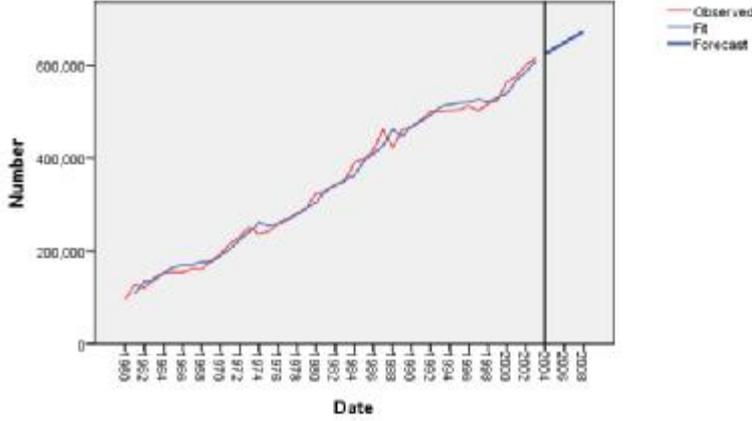
من الجدول نلاحظ أن المعاملات المقدرة معنوية، وهي لا تساوي الصفر، إذ أننا نرفض فرضية العدم والنموذج المقدر هو:

$$(1 - B)X_t = 11926.01 + e_t - 0.303e_{t-1}$$

أو يكتب بالشكل الآتي:

$$X_t = 11926.01 + X_{t-1} + e_t - 0.303e_{t-1}$$

والشكل (8) يبين المنحنى البياني للسلسلة الأساسية، وللقيم المقدرة، وللقيم المتنبأ بها:



الشكل (8) المنحنى البياني للسلسلة الأساسية وللقيم المقدرة وللقيم المتنبأ بها

3-3-5- التحقق من البواقي:

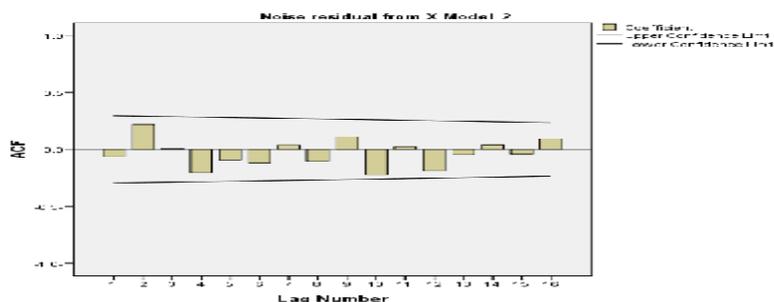
إن الاختبارات التي نطبقها على سلسلة البواقي هي:

أ- اختبار الارتباط السلسلي:

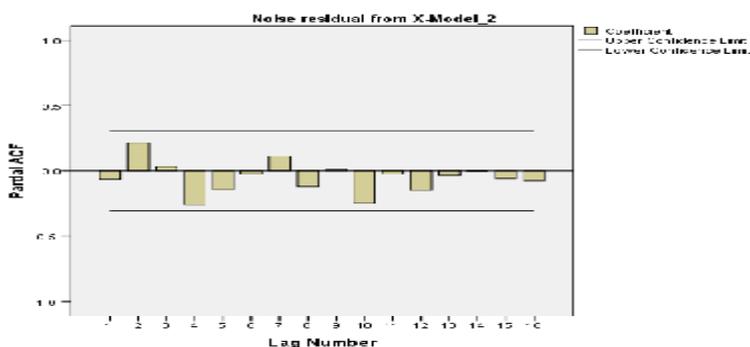
الجدول رقم (7) نتائج اختبار Ljung-Box

Ljung-Box Q(18)		
Statistics	DF	Sig.
13.888	17	.675

إن اختبار Ljung-Box كما يظهر في الجدول يشير إلى عدم وجود ارتباط ذاتي بين الأخطاء، ويؤكد ذلك دالتا الارتباط الذاتي والارتباط الذاتي الجزئي للبواقي، كما في الشكل (9) و (10):



الشكل (9) دالة الارتباط الذاتي للبواقي



الشكل (10) دالة الارتباط الذاتي الجزئي للبواقي

ب- طبيعة التوزيع الاحتمالي للبواقي:

الاختبار اللامعلمي K-S كلوموجوروف-سميرونوف يظهر أن البواقي تخضع للتوزيع الطبيعي (المعتدل) كما في الجدول رقم (8).

الجدول رقم (8) نتائج اختبار Kolmogorov-Smirnov

		Noise residual from X-Model_2
N		43
Normal Parameters <sup>a</sup>	Mean	126.1942
	Std. Deviation	14678.99773
Most Extreme Differences	Absolute	.081
	Positive	.081
	Negative	-.061-
Kolmogorov-Smirnov Z		.529
Asymp. Sig. (2-tailed)		.943
a. Test distribution is Normal.		

إن نتائج الاختبارات المطبقة على البواقي تؤكد صلاحية النموذج المقدر الثاني لتمثيل السلسلة الزمنية، ومن ثم إمكانية استخدامه في التنبؤ.

### 3-3-6- تقدير النموذج ARMA(1,1) لسلسلة الفروق الأولى:

إن تقدير النموذج ARMA(1,1) لسلسلة الفروق الأولى، يعني تقدير نموذج ARIMA(1,1,1) للسلسلة الأساسية. إن نتائج التقدير باستخدام برنامج SPSS كانت كما في الجدول رقم (9).

الجدول رقم (9) نتائج التقدير باستخدام برنامج SPSS

			Estimate	SE	t	Sig.
X-Model_1	X	Constant	11947.624	1712.457	6.977	.000
		AR Lag 1	-.542-	.330	-1.646-	.108
		4 Difference	1			
		MA Lag 1	-.177-	.388	-.457-	.650

من الجدول نلاحظ أن المعاملات المقدرة بعضها معنوي، وبعضها غير معنوي، وهي تشير إلى أن النموذج ARIMA(1,1,0) سيكون أفضل لتمثيل السلسلة، ومع ذلك سنحتفظ بهذا النموذج مبدئياً لنبحث في قدرته التنبؤية<sup>15</sup>. إن معادلتها هي:

15 لن يستخدم هذا النموذج والنموذج السابق في التنبؤ. بل سيتم استنتاج نموذج جديد معاملاته كلها معنوية.

$$(1 - B)(1 + 0.542B)X_t = 11947.624 + e_t + 0.177e_{t-1}$$

أو يكتب بالشكل الآتي:

$$X_t = 11947.624 + 0.458X_{t-1} + 0.542X_{t-2} + e_t + 0.177e_{t-1}$$

التحقق من البواقي:

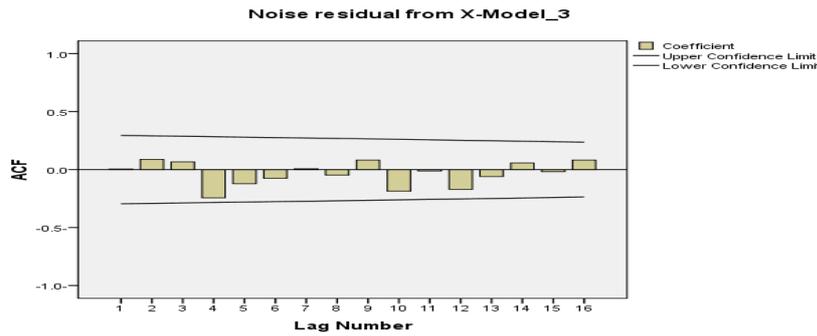
إن الاختبارات التي نطبقها على سلسلة البواقي هي:

أ - اختبار الارتباط السلسلي:

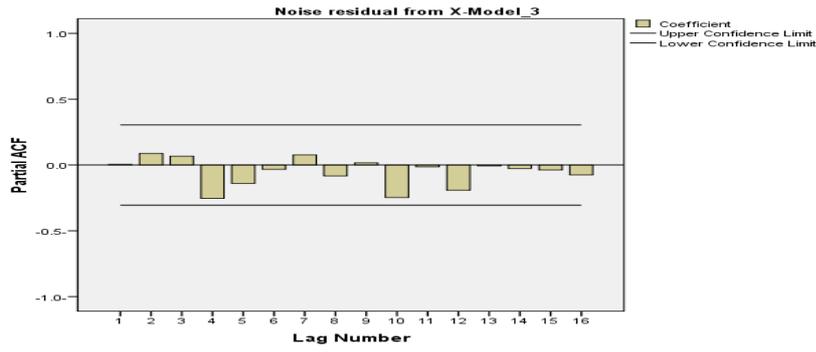
الجدول رقم (10) نتائج اختبار Ljung-Box

Ljung-Box Q(18)		
Statistics	DF	Sig.
10.751	16	.825

يشير اختبار Ljung-Box كما يظهر في الجدول إلى قبول فرضية العدم، أي عدم وجود ارتباط ذاتي بين الأخطاء، ويؤكد ذلك دالتا الارتباط الذاتي والارتباط الذاتي الجزئي للبواقي، كما في الشكل (11) و(12):



الشكل (11) دالة الارتباط الذاتي للبواقي



الشكل (12) دالة الارتباط الذاتي الجزئي للبواقي

ب- طبيعة التوزيع الاحتمالي للبواقي:

الاختبار اللامعلمي K-S كلوموجوروف-سميرونوف يظهر أن البواقي تخضع للتوزيع الطبيعي (المعتدل) كما في الجدول رقم (11) .

الجدول رقم (11) نتائج اختبار Kolmogorov-Smirnov

		Noise residual from X-Model_1
N		43
Normal Parameters <sup>a</sup>	Mean	175.9500
	Std. Deviation	14336.40698
Most Extreme Differences	Absolute	.095
	Positive	.095
	Negative	-.070-
Kolmogorov-Smirnov Z		.622
Asymp. Sig. (2-tailed)		.834
a. Test distribution is Normal.		

تؤكد نتائج الاختبارات المطبقة على البواقي صلاحية النموذج المقدر الثالث لتمثيل السلسلة الزمنية، ومن ثم إمكانية استخدامه في التنبؤ.

## 3-3-7 - مقارنة القدرة التنبؤية للنماذج الثلاثة:

أوجدنا باستخدام النماذج الثلاثة التنبؤات للقيم الخمس الأخيرة التي لم ندخلها في تقدير النماذج، وهي معطاة في الجدول رقم (12).

الجدول رقم (12) القيم المتنبأ بها باستخدام النماذج الثلاثة

السنوات	القيم الفعلية	التنبؤات		
		ARIMA(1,1,0)	ARIMA(0,1,1)	ARIMA(1,1,1)
2004	613600	626602.7078	625220.5637	627574.3211
2005	631950	638681.504	637146.5697	639202.5829
2006	635195	650559.0603	649072.5758	651323.4565
2007	639705	662515.3919	660998.5819	663177.0938
2008	656477	674440.8869	672924.588	675175.7037

من أجل مقارنة القدرة التنبؤية للنماذج الثلاثة نستخدم المعيارين الآتيين:

$$MAPE = (1/T) \sum (|X_t - X_t^*| / X_t) * 100$$

$$RMSE = [(1/T) \sum (X_t - X_t^*)^2]^{1/2}$$

إذ إن:

$x_t$ : القيمة المشاهدة.

$x_t^*$ : القيمة المتنبأ بها من خارج العينة.

نتائج المقارنة بين النماذج الثلاثة ملخصة في الجدول رقم (13).

الجدول رقم (13) نتائج المقارنة بين النماذج الثلاثة

	ARIMA(1,1,0)	ARIMA(0,1,1)	ARIMA(1,1,1)
RMSE	16083.862	14687.197	16784.454
MAPE	2.381	2.147	2.496

نلاحظ من الجدول السابق أن النموذج  $ARIMA(0,1,1)$  يعطي أفضل التنبؤات، إذ إن له أصغر قيمة لكل من المعيارين RMSE و MAPE. ثم يأتي في المرتبة الثانية النموذج  $ARIMA(1,1,0)$  ، وفي المرتبة الأخيرة يأتي النموذج  $ARIMA(1,1,1)$  .

### 3-3-8- مقارنة القدرة التنبؤية للنموذج $ARIMA(0,1,1)$ بمعادلة الاتجاه العام :

يتم إيجاد معادلة خط الاتجاه العام باستخدام الأسلوب التقليدي لتحليل السلاسل الزمنية الذي مازال الأسلوب الأكثر استخداماً في التطبيقات العملية والبحثية. إن معادلة خط الاتجاه العام Line للسلسلة (موضوع البحث) هي :

$$Y_t = 664430 + 12048t$$

أما نتائج التنبؤ لهذا النموذج مقارنة بالنموذج  $ARIMA(0,1,1)$  فتبين أن نتائج النموذج الأخير هي الأقرب إلى الفعلية، وهذا ما تؤكدته المعايير RMSE و MAPE الجدول رقم (14) والجدول رقم (15).

الجدول رقم (14) نتائج التنبؤ لنموذج ARIMA المختار مع نموذج الانحدار الخطي

line	ARIMA(0,1,1)	فعلية	السنوات
676477.6783	625220.5637	613600	2004
688525.7211	637146.5697	631950	2005
700573.764	649072.5758	635195	2006
712621.8068	660998.5819	639705	2007
724669.8497	672924.588	656477	2008

الجدول رقم (15) نتائج المقارنة بين النماذج الثلاثة بنموذج الانحدار الخطي

	line	ARIMA(1,1,1)	ARIMA(0,1,1)	ARIMA(1,1,0)
RMSE	65415.470	16784.454	14687.197	16083.862
MAPE	10.25576	2.496	2.147	2.381

### 3-4- المرحلة الخامسة: التنبؤ باستخدام النموذج المختار:

من أجل التنبؤ نختار النموذج  $ARIMA(0,1,1)$ ؛ لذلك نعيد تقدير هذا النموذج باستخدام مشاهدات السلسلة التي لدينا كلها، أي حتى عام 2008 فنجد أن المعاملات المقدره هي كما في الجدول رقم (16) .

الجدول رقم (16) نتائج تقدير معاملات النموذج ARIMA(0,1,1) باستخدام برنامج SPSS

			Estimate	SE	t	Sig.
X-Model_2	X	Constant	11563.188	1420.480	8.140	.000
		Difference	1			
		MA Lag 1	.315	.141	2.240	.030

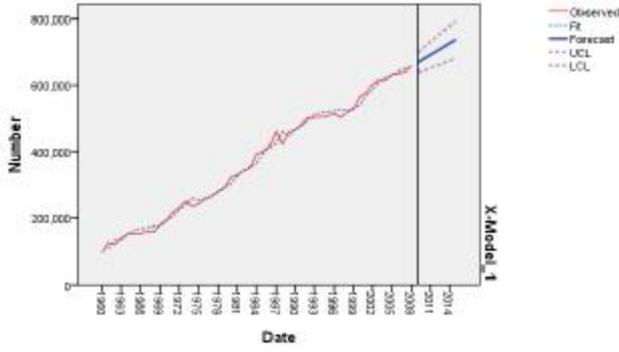
من الجدول نلاحظ أن المعاملات المقدرة معنوية، و هي لا تساوي الصفر، من هذا نرفض فرضية العدم.

إنّ التنبؤات للأعوام السبعة من خارج السلسلة - أي حتى عام 2015 - مع حدي الثقة هي كما يأتي في الجدول رقم (17).

الجدول رقم (17) التنبؤات مع حدي الثقة بدرجة 95%

العام	LCL_X	Predicted_X	UCL_X
2009	638695.8223	667326.0148	695956.2074
2010	644181.8498	678889.2031	713596.5564
2011	650583.6967	690452.3914	730321.0862
2012	657581.0743	702015.5797	746450.0851
2013	665005.7537	713578.768	762151.7823
2014	672756.3651	725141.9563	777527.5475
2015	680766.226	736705.1446	792644.0632

أما الشكل البياني للسلسلة المشاهدة مع التنبؤات فهي معروضة في الشكل (13):



الشكل (13) الشكل البياني للسلسلة المشاهدة مع التنبؤات

### النتائج والتوصيات:

- 1- تشكل سلسلة أعداد المنتسبين إلى الصف الأول من التعليم الأساسي سياقاً عشوائياً غير مستقر، وظهر اختبار Dickey and Fuller وجود جذر الوحدة، وقد أخذ مرشح الفروق الأولى لجعلها مستقرة .
- 2- تبيّن من مقارنة نماذج التحليل الحديث المبنية على منهجية Box-Jenkins مع النماذج المبنية على الأسلوب التقليدي أفضلية النماذج المبنية على منهجية (بوكس-جنكينز) .
- 3- إن النموذج الأفضل من بين النماذج التي وُضعت في هذا البحث للتنبؤ بأعداد المنتسبين إلى الصف الأول من التعليم الأساسي هو النموذج  $ARIMA(0,1,1)$  ، وصيغته هي :  

$$(1 - B)X_t = 11563.19 + e_t - 0.315e_{t-1}$$
 أو  

$$X_t = 11563.19 + X_{t-1} + e_t - 0.315e_{t-1}$$
- 4- نوصي باستخدام النموذج الذي تم التوصل إليه في التنبؤ بأعداد المنتسبين إلى الصف الأول من التعليم الأساسي، واعتماد التنبؤات التي يعطيها بوضع الخطط المستقبلية لقطاع التعليم، ومن ثمّ الخطط المبنية عليها خلال السنوات القريبة.
- 5- نوصي باستخدام هذا المنهج في استنتاج النموذج القياسي وتطويره للتنبؤ بأعداد التلاميذ المتوقع انتسابهم إلى الصف الأول من كل عام، وذلك حسب تطور السلسلة الفعلية لأعداد التلاميذ .

## المراجع

### المراجع باللغة الانكليزية :

- 1- BERA, A.K. and Jarque .C.M.(1981), "An efficient large Sample test for normality of observations and regression residuals ", Working paper in Econometrics No 40,Australian National university, Canberra.
- 2- Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis Forecasting and Control*, 2<sup>nd</sup> ed., Holden-Day, San Francisco.
- 3- Dickey D. and Fuller W.(1979), " Distribution of the estimators for Autoregressive Time Series With a unit Root ", Journal of the American Statistical Association, n74: pp .427-431.
- 4- Dickey D. and Fuller W.(1981) The likelihood Ratio Statistics for Autoregressive Time Series With a unit Root", *Econometrica* ,n49: pp .1057-1072.
- 5- FULLER A.W. (1996) "Introduction to Statistical Time Series". JOHN WILEY & SONS, INC, New York.
- 6- Kirchgässner G. and Wolters J. (2007) "Introduction to Modern Time Series Analysis", SPRINGER-Verlag, Berlin Heidelberg.
- 7- Ljung, G.M., and Box G.E.P. (1978) "on a measure of the lack of fit in time Series models". *Biometrika*, n65:PP.297-303.
- 8- PANKRATZ A. (1983) "Forecasting with Univariate Box-Jenkins Models". JOHN WILEY & SONS, New York.
- 9- SHUMWAY R.H. and STOFFER D.S. (2005) "Time Series Analysis and Its Applications". SPRINGER, New York.
- 10- Wei, W. W. S. (1990). *"Time Series Analysis Univariate and Multivariate Methods"*, Addison Wesley.
- 11- YAFFEE R. and McGee M. (1999) "Introduction to Time Series Analysis and Forecasting". ACADEMIC PRESS, INC, New York.

### المراجع باللغة الفرنسية :

- 1- BENSABER A. et BLEUSE-TRILLON B., (1989) "Pratique des chroniques et de le prévision à court terme". Edition MASSON-Paris.
- 2- GOURIEROUX C. et MONFORT A., (1990) "Séries Temporelles et Modèles Dynamiques " Ed. Economica-Paris.

- 3- HENIN P.Y. (1989), "Bilans et essais sur la non-Stationnarité des séries Macroéconomiques" revue d'économie politique – n5-pp 661-691.
- 4- MARCHAL,J.L ,SEYS ,B. , et autres (2005), "Agrégation de prévisions mensuelles de consommations medicamenteuses à l'aide d'un Modèle ARIMA "Rev. Statistique Appliquée ,LIII (2) , PP. 5-28.
- 5- MATHIS A., (1990), "Une Approche en terme de processus stochastiques vectoriels de la dette publique Française" Thèse de doctorat, EHESS-Paris.

المراجع باللغة العربية :

- 1- المجموعة الإحصائية، الصادرة عن المكتب المركزي للإحصاء في الجمهورية العربية السورية، للأعوام: 2009 و 1989 و 1976م .
- 2- بري، عدنان ماجد عبد الرحمن، طرق التنبؤ الإحصائي ( الجزء الأول).جامعة الملك سعود، 2002 م .
- 3- شعراوي، سمير مصطفى، مقدمة في التحليل الحديث للسلاسل الزمنية. مركز النشر العلمي، جامعة الملك عبد العزيز، 2005م .
- 4- فاندل، والتر، السلاسل الزمنية من الواجهة التطبيقية و نماذج بوكس و جنكينز، تعريب: عبد المرضي عزام و أحمد هارون. دار المريخ للنشر 1992م .