استخدام نماذج السلاسل الزمنية للتنبؤ باعداد الاطفال المصابين بالتهابات المجاري التنفسية في محافظة السليمانية

م.د. محمد محمود فقى حسين
كلية الإدارة والاقتصاد
جامعة السليمانية

م. د.اخترخان صابر حمد
كلية الإدارة والاقتصاد
جامعة صلاح الدين

Use the time-series models to predict the numbers of children with Respiratory Infections Streams(RIS) in the governorate of Sulaymaniyah

تاريخ استلام البحث ٢٠١٣/٦/١٨ تاريخ قبول النشر ٢٠١٣/٩/٢٢

المستخلص

يهدف البحث الى نمذجة التطور الزمني لأعداد الاطفال المصابين بمرض التهاب المجاري التنفسية العليا في محافظة السليمانية و ذلك بالاعتماد على البيانات الشهرية للفترة (٢٠٠٥-٢٠١١). وقد أظهرت نتائج التحليل أنَّ النموذج الملائم لها هو نموذج المتوسط المتحرك التكامل من الدرجة الثانية ARIMA(0,1,2) و بالاعتماد على هذا النموذج تم التنبؤ باعداد الاطفال المصابين بمرض التهاب المجاري التنفسية العليا شهرياً و لخمسة سنوات قادمة و قد كانت القيم التنبؤية متناسقة مع قيم السلسلة الأصلية مما يدل على كفاءة النموذج.

Abstract :

The main objective of this paper is modeling the time evolution of the number of children with upper respiratory tract infection (URT) in the governorates of Sulaymaniyah and relying on monthly data for the period (2005-2011). The results of the analysis showed that the appropriate model is integrated moving average model of second-class ARIMA (0,1,2), and based on this model we can predicted to prepare children with upper respiratory tract infection (URT) per month and for the next five years and has been a consistent predictive values with the values of the original series which shows the efficiency of the model.

هدف البحث:

١. نمذجة التطور الزمني لأعداد الاطفال المصابين بمرض التهاب المجاري التنفسية العليا في محافظة السليمانية خلال السنوات (٢٠٠٥-٢٠١١).
٢. التنبؤ بأعداد الأ طفال المصابين بمرض التهاب المجاري التنفسية العليا في السنوات القادمة وهذا الأمر مفيد لإدارة المستشفيات و المراكز الصحية ، وكذلك الحكومة في تخطيط برامج صحية ملائمة من حيث توفير الأدوية وتحديد الأطباء ذوي الاختصاصات والإمكانيات الملائمة لهذا العدد من المصابين.

فرضية البحث

إنَّ أعداد المصابين من الأطفال المصابين بمرض التهاب المجاري التنفسية العليا ينمو بشكل متزايد خلال الفترة (٢٠٠٥-٢٠١١) . و يعتبر التنبؤ با عداد المصابين في محافظة السليمانية مدخلاً للتنبؤ باعداد المصابين بهذا المرض لبقية المحافظات التي واجهت نفس الظروف.

منهجية البحث

تم تقسيم البحث الى جانبين هما الجانب النظري و الذي تم فيه التطرق بشكل مبسط الى الأسس النظرية الخاصة بنماذج السلاسل الزمنية من حيث الشكل العام و مراحل بناء النموذج و طرق التقدير و التنبؤ ، أمَّا الجانب التطبيقي فقد تم فيه إجراء دراسة تطبيقية على بيانات واقعية عن عدد الاطفال المصابين بمرض التهاب المجاري التنفسية العليا في محافظة السليمانية للوصول الى نموذج رياضي للتنبؤ بعدد الاطفال المصابين بهذا المرض لفترات اللاحقة و يتضمن الجزء الأ خير على أهم الاستنتاجات و التوصيات و المصادر.

الأدوات المستخدمة

الأدوات المستخدمة في البحث هي البرنامج الإحصائي Minitab21 ، Excel

أولاً: الجانب النظري

١ نماذج السلاسل الزمنية Time Series Model [1] [3]

١-١ المقدمة Introduction

يُعَّ موضوع تحليل السلاسل الزمنية من المواضيع الإحصائية المهمة التي تتناول سلوك الظواهر، وتفسرها عبر حقب محددة. ويمكن إجمال أهداف تحليل السلاسل الزمنية بالحصول على وصف دقيق للملامح الخاصة للعملية التي تتولد منها السلسلة الزمنية، وبناء نموذج لتفسير سلوك السلسلة الزمنية واستخدام النتائج للتكهن بسلوك السلسلة في المستقبل، إضافة إلى التحكم في العملية التي تتولد منها السلسلة الزمنية بفحص ما يمكن حدوثه عند تغيير بعض معلمات النموذج. ولتحقيق ذلك يتطلب الأمر دراسة تحليلية وافية لنماذج السلاسل الزمنية بالاعتماد ع لى الأساليب الإحصائية والرياضية.

١-٢ بعض المصطلحات الاساسية لتحليل السلاسل الزمنية [1][2]

١ -السلسلة الزمنية Time Series

عبارة عن مجموعة من المشاهدات التي تتولد خلال فترة زمنية معينة .

٢- الاستقرارية Stationary

تعتبر الاستقرارية أحد المفاهيم الأساسية الهامة في تحليل السلاسل الزمنية بالطريقة الحديثة، حيث تعتبر السلسلة الزمنية موضع الدراسة ساكنة إذا كانت الخصائص الإحصائية الأساسية لها

(التوقع ،التباين ، التغاير) ثابتة عبر الزمن أى إنَّها مستقلة عن الزمن، وذلك بالاضافة الي النموذج العام ARMA(p,q) وعلي ذلك فقد تكون السلسلة الزمنية موضع الدراسة غير ساكنة في المتوسط أو غير ساكنة في التباين أو في الإثنين معاً.

٢ - الإنعكاس Invertibility

هذا المفهوم يعتبر من خصائص النموذج ويقصد به إمكانية التعبير عن الأخطاء العشوائية للنموذج بدلالة المشاهدات السابقة للسلسلة $y_{t-1}, y_{t-2}, \ldots\ldots\ldots, y_{t-p}$

ويلاحظ أنَّ نماذج السلاسل الزمنية تنقسم إلي نوعين من حيث الإنعكاس:

- منعكس دائما(نماذج الانحدار الذاتي(P) AR)
- منعكس تحت شروط (نماذج المتوسطات المتحركة (q) MA والنماذج المختلطة (ARMA(p,q)

٣ - دالة الارتباط الذاتي Autocorrelation (ACF) function [1][3]

عبارة عن دالة بين ρ_k (قيمة معامل الارتباط بين مشاهدات السلسلة الزمنية التي تبعد عن بعضها البعض مسافة قدرها k وحدة زمنية) و الفجوة الزمنية K (تعرف دالة الارتباط الذاتي من الرتبة k)، وتقع قيمة معاملات دالة الارتباط الذاتي في الفترة [1+ ،1 -] ويمكن القول أنَّ دالة الارتباط الذاتي ρ_k للسلاسل الزمنية الساكنة قد تأخذ شكل دالتين آسيتين أو شكل موجات من دالة الجيب أو تنقطع تماما، أما في حالة السلسلة الزمنية غير الساكنة فإنّها تتناقص ببطء نحو الصفر.

٤ - دالة الارتباط الذاتي الجزئي Partial (PACF) autocorrelation [1][2]

تعرف دالة الارتباط الذاتي الجزئي ϕ_{kk} بأنَّها الدالة التي توضح العلاقة بين المتغيرين y_t , y_{t+k} بعد استبعاد تأثير المتغيرات التي تقع بينهما، وهذا يعنى إمكانية استخدام معامل الارتباط الذاتي الجزئي لقياس درجة العلاقة بين أى متغيرين بعد حذف أثر المتغيرات الوسيطة بينهما .

بعد هذا العرض السريع للمفاهيم الأساسية التي يعتمد عليها التحليل الحديث للسلا سل الزمنية سوف نقوم باستعراض النماذج السابق الإشارة إليها بشئ من التفصيل للتعرف على الخصائص الأساسية المميزة لها.

1-٣ نماذج بوكس وجينكنز [2][6]

افترض بوكس وجينكنز أنّ نماذج الانحدار الذاتي والمتوسطات المتحركة ARMA مناسبة لوصف سلوك السلسلة الزمنية والتنبؤ بها ، وقّ دما طريقتهما في نمذجة بيانات السلسلة الزمنية، وتشترط هذه الطريقة سكون السلسلة الزمنية سواء في المتوسط (السلسلة لها اتجاه عام) أو في التباين (أي تكون التذبذبات حول المتوسط الثابت مستقرة طول السلسلة الزمنية).

وعلى ذلك فإنَّ أولى خطوات التحليل الحديث للسلسلة الزمنية هي التحقق من توفر هذين الشرطين وذلك عن طريق رسم المنحنى الزمني للسلسلة الزمنية ، فإذا لاحظنا منه عدم تحقق أي من الشرطين فإنَّه يتم إجراء عدد من التحويلات على السلسلة الزمنية كالآتي:

- إذا كانت السلسلة الزمنية غير ساكنة في المتوسط فإنّ نا نقوم بأخذ الفروق المتتالية لتسكينها، ففي حالة كون معادلة الاتجاه العام من الرتبة الأولى فإّ ننا نحصل على السلسلة الساكنة من خلال أخذ الفرق الأول و في حالة كونها من الرتبة الثانية فإنن ا نأخذ الفرق الثاني للحصول على السلسلة الساكنة ونرمز عادة لعدد الفروق التي يتم أخذها لتسكين السلسلة بالرمز d.

إذا كانت السلسلة غير ساكنة في التباين نق وم بإجراء تحويله مناسبة مثل الج ذر التربيعى أو اللوغاريتم

إذا كانت السلسلة غير ساكنة في كل من التباين والمتوسط نقوم بإجراء تحويله لتسكين التباين أولاً ثم نختبر هل السلسلة ساكنة في المتوسط، فإذا كانت غير ساكنة نقوم بأخذ الفرق المناسب.

وبعد التحقق من توفر شرط سكون السلسلة يتم تطبيق خطوات هذه الطريقة والتي تتلخص في:

١ - تحديد النموذج Identification ٢- التقدير Estimation

٣ – التشخيص Diagnostic ٤- التنبؤ Prediction

1-٣-١ تحديد النموذج Identification:

في هذه الخطوة يتم تحديد النموذج الملائم لوصف السلسلة الزمنية محل الاهتمام ، ويتضمن ذلك تحديد رتب نموذج ARMA (p,q) المناسب للسلسلة حيث **p** ترمز إلى رتبة الانحدار الذاتي في النموذج **q**,ar(p) ترمز إلى رتبة المتوسطات المتحركة MA(q) ولتحديد هذه الرتب للسلسلة الزمنية محل الدراسة ندرس كل من دالتي الارتباط الذاتي ACF و الارتباط الذاتي الجزئي PACF ونقارنها بالسمات النظرية لكل نموذج من النماذج المتاحة من عائلة ARMA لاختيار النموذج المناسب.

1- نماذج الانحدار الذاتي AR(p)

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots\ldots + \varphi_p Y_{t-p} + \varepsilon_t \quad \ldots\ldots\ldots(1)$$

حيث تتميز كل من دالتي الارتباط الذاتي و الارتباط الذاتي الجزئي لهذا النموذج بالآتي :

- دالة الارتباط الذاتي (ACF) تتناقص هذه الدالة أسيا او بإشارات مترددة عند كل قيم الفجوات الزمنية k دون انقطاع.
- دالة الارتباط الجزئي (PACF) تنقطع هذه الدالة بعد فجوة زمنية مساوية لرتبة النموذج.

2- نماذج المتوسطات المتحركة MA(q)

$$Y_t = -\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \text{-------} - \theta_q \varepsilon_{t-q} + \varepsilon_t \ldots\ldots(2)$$

نجد أنّ دالة الارتباط الذاتي ACF تنقطع بعد فترة زمنية مساوية لرتبة النموذج q ،

أمَّا دالة الارتباط الذاتي الجزئي PACF فهي تتناقص ولا تنقطع عند كل الفجوات الزمنية k.

3- النماذج المختلطة ARMA(p,q)

$$Yt = \varphi 1 Yt-1 + \varphi 2 Yt-2 + \ldots\ldots + \varphi p Yt-p - \theta 1 \varepsilon t-1 - \theta 2 \varepsilon t-2 - \ldots\ldots - \theta q\, \varepsilon t-q + \varepsilon t \ldots\ldots(3)$$

وبناءً على أنّها مزيج من النموذجين السابقين فان دالتي الارتباط الذاتي (ACF) والارتباط الذاتي الجزئي (PACF) تأخذ النمط المتناقص دائما عند كل الفجوات دون انقطاع.

٢-٣- 1 التقدير Estimation :

بعد أن انتهينا من عرض أولى خطوات منهجية بوكس وجينكنز وهى التعرف ننتقل إلى الخطوة التالية لها وهى التقدير "اى تقدير معالم النموذج الذي تم اختياره " ،تتعدد الطرق التي يمكن استخدامها للتقدير في نظرية الإحصاء أهمها طريقتي المربعات الصغرى. وطريقة الإمكان الاعظم.

٣-٣-1 التشخيص Diagnostic :

تُعَدُّ هذه الخطوة ثالث خطوات منهجية بوكس و جينكنز بعد التعرف على النموذج الملائم و تقدير معالمه، وتعد أهم مراحل المنهجية وذلك لأنَّ ها توضح مستقبل النموذج الأ ساسي هل سيكون النموذج الملائم لتحليل السلسلة محل الاهتمام وبالتالي يتم استخدامه في التنبؤ أم سيتم تطويره بحيث يصبح أكثر ملائمة لتمثيل السلسلة محل الدراسة.

إذاً فجوهر عملية التشخيص هو التأكد من أنّ الافتراضات النظرية للنموذج الأساسي متحققة في الواقع التطبيقي ، فإذا كانت الافتراضات النظرية الاساسي المبدئي منطبقة على السلسلة محل الدراسة كان هو النموذج الملائم لوصف هذه السلسلة و إن لم تكن منطبقة على السلسلة محل الدراسة نقوم بتطويره من خلال عِدَّة فحوص أو اختبارات.

الفحوص المستخدمة في مرحلة التشخيص [2] [3][5]:

١) اختبار الاستقرارية Test of stationary ٢) اختبار الانعكاس Test of invertibility

٣) تحليل البواقي Residuals analysis ٤) توفيق النموذج الأدنى مباشرة Under fitting

٥) توفيق النموذج الأعلى مباشرة Over fitting

أولاً: اختبار الاستقرارية Test of stationary:

يقوم التحليل الحديث للسلاسل ا لزمنية على فرضية الاستقرارية السلسلة الزمنية ،وبالتالي لابد أن يصف النموذج الملائم سلسلة مستقرة ، وحتى نتأكد من أن النموذج المبدئي يصف سلسلة ساكنة نقوم بفحص معاملات دالة الانحدار الذاتي للنموذج المبدئي والتي حصلنا عليها في مرحلة التقدير لمعرفة ما إذا كانت هذه المعاملات تحقق شرط السكون أم لا؟ وهذا الشرط هو أن تكون القيم المطلقة لجذور المعادلة $\Phi(\beta) = 0$ أكبر من الواحد الصحيح ، فإذا كانت إحدي القيم المطلقة لجذور المعادلة السابقة = ١ فمعنى ذلك أنّ السلسلة غير ساكنة أي لها إتجاه عام وبالتالي لكي نتخلص من هذا ا لاتجاه العام نقوم بأخذ الفروق المتتالية حتى يمثل النموذج سلسلة مستقرة .

فمثلاً إذا كان النموذج المبدئي ARIMA (1,0,1) وكانت المعلمة $\boldsymbol{\phi=1}$ أي إنَّ هذا النموذج لا يصف سلسلة ساكنة إذا نقوم بأخذ الفروق حتى تصبح السلسلة ساكنة كالتالي:

إذا كان النموذج الأصلي ARIMA (1,0,1) علي الصورة :

$$(1- \phi\ \beta)\ y_t = (1- \theta\ \beta)\ \varepsilon t \quad \ldots\ldots\ldots(4)$$

فإنّنا نستبدله بالنموذج ARIMA(0.1.1)

$$(1- \beta)\ Y_t = (1- \theta\ \beta)\ \varepsilon t \quad \ldots\ldots\ldots\ldots(5)$$

ثانيا: اختبار الانعكاس Test of Invertibility

نجد أنَّ لخاصية الإ نعكاس أهمية في تقدير الأخطاء وبالتالي لابد للنموذج الملائم أن يكون منعكس لذلك نقوم باختبار تحقق شرط الانعكاس وهو أن تكون جذور المعادلة $\theta(\beta)=0$ اكبر من الواحد الصحيح، ولو كانت مساوية للواحد سيكون النموذج غير منعكس فنقوم بتعديله بحيث يصبح منعكس عن طريق إلغاء الفروق المتتالية والعودة إلي السلسلة التي كانت موجودة قبل أخذ الفروق، حيث يقال في هذه الحالة إنَّها وقعنا في مشكلة Over differencing .

ثالثا: تحليل البواقي [4] Residuals Analysis.

إذا كان النموذج المبدئي ملائم فلابد أن تحقق البواقي نفس خصائص الأخطاء أو على الأق ل لا تتعارض معها ومن هذه الخصائص عدم وجود ارتباط ذاتي بين الأخطاء ، ومتوسطها يساوي الصفر تباينها يساوي σ^2 ونقوم بتحليل البواقي من خلال عِدَّة خطوات :

١. رسم البواقي Residual:

نقوم برسم البواقي ضد الزمن فإذا كانت البواقي ليس لها نمط معين أي متذبذبة حول الصفر فهي تحقق خصائص الأخطاء وبالتالي فهذا مؤشر علي أنَّ النموذج ملائم .

٢. فحص دالة الارتباط الذاتي للبواقي.

يكون النموذج المبدئي ملائماً إذا كانت الأخطاء تمثل تغيرات عشوائية بحتة وبالتالي لابد أن تعكس البواقي هذه الخاصية لذلك نقوم بدراسة دالة الارتباط الذاتي لل بواقي وحتى تكون الأخطاء متغيرات عشوائية بحتة لابد أن تكون معاملات الارتباط الذاتي صغيرة بحيث يمكن قبول عدم اختلافها معنوياً عن الصفر أي تكون دالة الارتباط الذاتي خالية من النتوءات بمعني أنَّه إذا كان معامل الارتباط الذاتي للبواقي يقع داخل الفترة $(2+\sqrt{n}\,,\,-2-\sqrt{n})$. ولا يمكن الإ عتماد علي فحص دالة الارتباط الذاتي وحدها للتأكد من عشوائية الأخطاء وذلك لاحتمالية وجود معاملات ارتباط ذاتي داخل الفترة السابقة وبذلك فهي تختلف معنوياً عن الصفر حيث إنَّ التباين المستخدم في فترة الثقة السابقة أكبر من التباين الحقيقي.

رابعاُ: توفيق النموذج الأدنى مباشرة Under fitting.

قد يري الباحث بخبرته الشخصية أنَّ وجود أحد المعالم في النموذج غير ضروري ويرى أنَّ هذ ا مؤشر جيد لتبسيط النموذج لذا ف إنَّه يقوم باختبار معنوية هذه المعلمة ود راسة ارتباطها

بالمعاملات الأخرى، فإذا كانت المعلمة المتحدث عنها لا تختلف معنوياً عن الصفر أو إذا كان بينها وبين إحدى المعالم الأخرى ارتباط قوي فيمكن حذفها وبالتالي تبسيط النموذج.

خامساً: توفيق النموذج الأعلى مباشرة Over fitting .

- إذا اعتقد الباحث أنَّ هناك أحد المعالم التي لو أضيفت إلي النموذج جعلته أكثر ملائمة لتحليل السلسلة محل الدراسة ، فإنَّه يقوم بدمجها في النموذج المبدئي ثم يقوم باختبار معنويتها فان كانت معنوية تركها في النموذج والعكس.
- ثم يقوم بدراسة معنوية المعالم الأصلية التي قد تتأثر بالمعالم المضافة فإن كانت غير معنوية أو إذا كان بينها وبين المعلمة المضافة ارتباط قوي يتم حذفها.

٤-٣-1 التنبؤ Forecasting [3]

إنَّ إحدى الأهداف الأساسية لبناء نماذج السلاسل الزمنية هي امكانية التنبؤ بالقيم المستقب لية لهذه المتغيرات و كذلك يمكن أن نصف أهم إحدى الأهداف الأساسية لتحليل السلاسل الزمنية و هو التنبؤ بالقيم المستقبلية للظاهرة قيد الدراسة $(\tilde{Z}_{t+l})$ في السلسلة الزمنية للظاهرة ، حيث إنَّ (t) و هي عدد المشاهدات و إنَّ (L=0,1,2,...) تمثل الفترات الزمنية المستقبلية المطلوب التنبؤ بها . إنَّ عملية التنبؤ تستخدم في النماذج المطابقة للبيانات لأ ي ظاهرة اقتصادية ، اجتماعية، طبية، أو أي ظاهرة ذات متغير زمني مؤثر على أن يكون النموذج مطابق لجم يع الخصائص سواء لمتغير الظاهرة أو الخطأ العشوائي في هذه النماذج و يتم اختيار النموذج الملائم للسلسلة الزمنية عندما ينجح في الاختبارات الإحصائية المناسبة و يتمتع بالخواص التالية و هي:-

١) اذا كان للنموذج أقل تباين للخطأ (SE_{t+l}) و هنا يجب حساب الصيغة (SE_{t+l})

٢) اختيار مستوى الثقة و حسب طبيعة و سلوك الظاهرة حيث يمكن اختيار درجة الثقة حسب الصيغة التالية ١٠٠%(α-١) عندما(α) تمثل مستوى قبول الخطأ في النموذج أو مستوى المعنوية المطلوب.

٣) تحديد أو اختيار فترة التنبؤ للقيمة الجديدة (Z_{t+l})و حسب سلوك الظاهرة المطلوبة لنفرض لدينا السلسلة الزمنية $(Z_t, Z_{t-1}, .., Z_1)$ للفترة (t) و عليه يمكن وضع صيغة التنبؤ بقيمة الظاهرة (Z_t) للفترة القادمة أو المستقبلية (L) و كما يلي:-

$$\tilde{Z}_t(l) = E(Z_{t+l}/Z_t, Z_{t-1}, \dots, Z_1) \dots\dots\dots\dots\dots\dots (6)$$

حيث إنَّ l: تمثل الزمن المستقبلي t: يمثل الزمن الأصلي

ثانياً : الجانب التطبيقي

المقدمة Introduction

نتطرق في هذا الإطار الى اختيار النموذج الملائم للبيانات الموضوعة للبحث باستخدام نماذج السلاسل الزمنية والتنبؤ بعدد الاطفال المصابين بهذا النوع من المرض . وقد تم الحصول على البيانات من مستشفى الأطفال في محافظة السليمانية خلال السنوات ٢٠٠٥ – ٢٠١١. إنَّ التطور الحاصل في مجال الطب جعل بالإمكان الوقاية من الكثير من الإمراض ومكافحتها بل القضاء عليها من خلال تطبيق الوسائل العلمية الحديثة للبحوث التي تهدف إلى دراسات مفصلة لمختلف أنواع الامراض و مسبباتها وطرائق تكاثرها بين أفراد المجتمع لغرض التعرف على طرائق الوقاية منها والسيطرة عليها ومنع انتشارها.

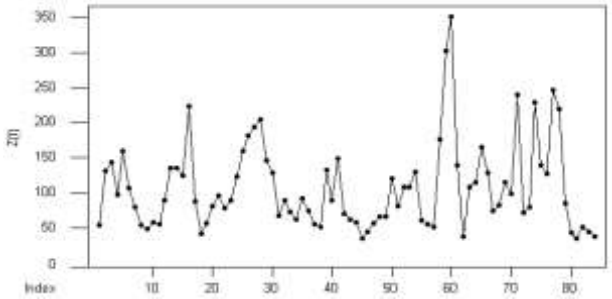
جدول رقم (١)

يوضح مجموع الاطفال المصابين بمرض وتصنيفهم حسب الجنس و الفئات العمرية خلال السنوات (٢٠٠٥-٢٠١١)

السنوات	مجموع الاطفال المصابين	النسبة المئوية	الجنس		الفئات العمرية		
			ذكور	اناث	اقل من سنة	من ١-٤	من ٥-١٤
٢٠٠٥	١٠٧٩	١١.٩٣٩٨	٦٣٨	٤٤١	٣٢١	٤٩٧	٢٦١
٢٠٠٦	١٢٧٦	١٤.١١٩٧٣	٧٤٣	٥٣٣	٣٧٨	٥٨١	٣١٧
٢٠٠٧	١٤٧٥	١٦.٣٢١٧٩	٨٢٣	٦٥٢	٤٥٩	٦٨٤	٣٣٢
٢٠٠٨	٨٧٥	٩.٦٨٢٤١٧	٤٧٥	٤٠٠	٣٣٨	٣٦٧	١٧٠
٢٠٠٩	١٦١٢	١٧.٨٣٧٧٨	٩٩٨	٦١٤	٦٠٠	٧١٨	٢٩٤
٢٠١٠	١٣٨٠	١٥.٢٧٠٥٥	٥٥١	٨٢٩	٤٩٤	٦١٨	٢٦٨
٢٠١١	١٣٤٠	١٤.٨٢٧٩٣	٧٤٤	٥٩٦	٤٨٠	٥٧٨	٢٨٢

إنَّ الخطوة الأولى لتصميم نظام التنبؤ هي تحديد النموذج الاحصائي للبيانات ومن ثم تقدير النموذج (ونقصد بها تقدير معالم النموذج) و اختبار النموذج .وفي هذا الجانب قمنا بتحديد النموذج الملائم لعدد المرضى المصابين بمرض الـتهاب المجاري التنفسية العليا للأ طفال و التنبؤ بأعداد المصابين للفترات القادمة و نتبع الخطوات التالية:-

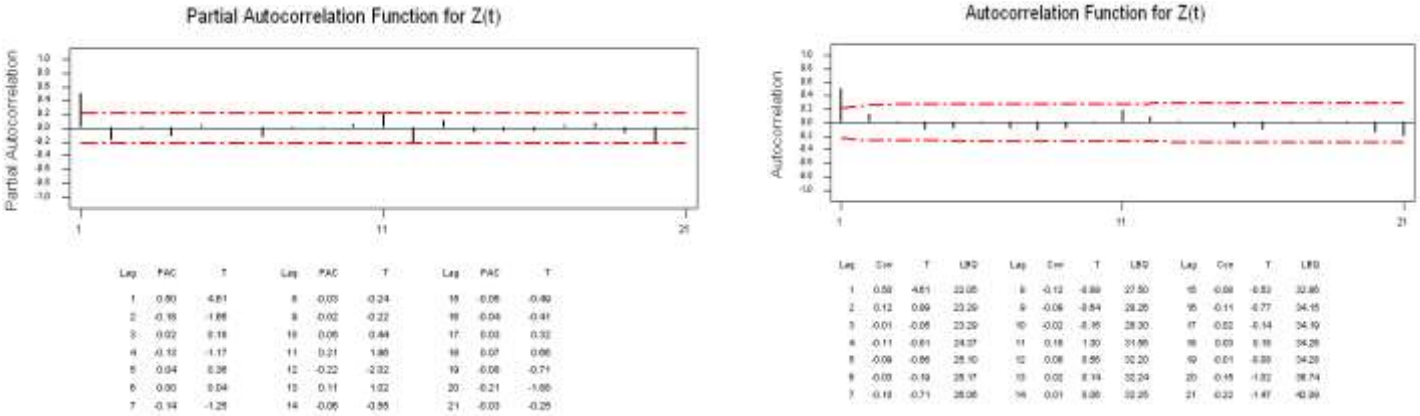
١) يتم رسم السلسلة الزمنية الأصلي (Time Plot) للتعرف على مدى استقراريتها و كمايلي:-



شكل رقم (١) يوضح الرسم البياني للسلسلة الأصليةِ لأعداد المرضى

من الرسم أعلاه نلاحظ بأنَّ السلسلة غير مستقرة الى حد ما ، يتم حساب و رسم دالة الارتباط الذاتي (ACF) و دالة الارتباط الذاتي الجزئي (PACF) كالآتي:-

2) نرسم الترابط الذاتي العيني (Autocorrelation) ٣) نرسم الترابط

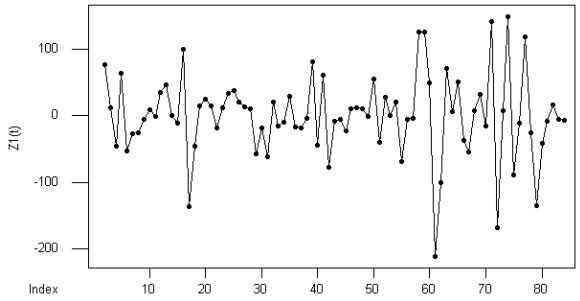


الذاتي الجزئي العيني (Partial Autocorrelation)

شكل رقم (٢) يوضح الرسم البياني للسلسلة الأصلية للمرضى

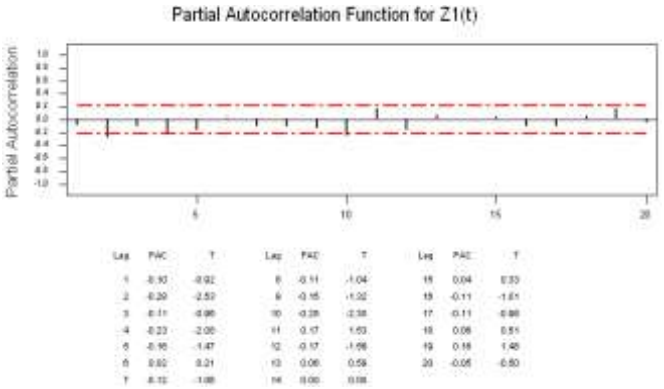
شكل رقم (٣) يوضح الرسم البياني للسلسلة الأصلية للمرضى

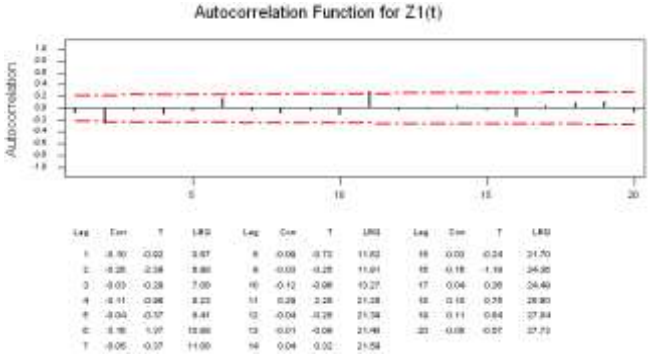
من خلال الشكلين أعلاه نلاحظ بأنَّ دالة الترابط الذاتي العيني(ACF) تدل على انخفاض بطيىء مما يدل على عدم استقرار في المتوسط . نقوم بأخذ الفرق الأول للسلسلة و نرسمها



شكل رقم (٤) الرسم البياني للسلسلة بعد أخذ الفرق الأول

الشكل اعلاه يوضح بان المتسلسلة مستقرة (Stationary) و دوال الترابط الذاتي (Autocorrelation) و الترابط الذاتي الجزئي (Partial Autocorrelation) للمتسلسلة المستقرة





شكل رقم (٥) يوضح رسم دالة (ACF) بعد أخذ الفرق الأول شكل رقم (٦)يوضح رسم دالة (PACF) بعد أخذ الفرق الأول

تقدير معالم النموذج

ARIMA Model: Z(t)

Final Estimates of Parameters

Type	Coef	SE Coef	T	P	MS	DF	ss
AR 1	-0.1028	0.1105	-0.93	0.355	3788	81	3068611
Constant	-0.300	6.756	-0.04	0.965			

Final Estimates of Parameters

Type	Coef	SE Coef	T	P	MS	DF	SS
AR 1	-0.1330	0.1072	-1.24	0.218			
AR 2	-0.2843	0.1071	-2.65	0.010	80	3526	282041
Constant	-0.653	6.517	-0.10	0.920			

Final Estimates of Parameters

	MS	DF	SS	Type		Coef	SE Coef	T	P
3723	81		301572	MA	1	0.2788	0.1066	2.61	0.011
				Constant		-0.431	4.832	-0.09	0.929

Final Estimates of Parameters

	MS	DF	SS	Type		Coef	SE Coef	T	P
2938	80		235058	MA	1	0.4272	0.1008	4.24	0.000
				MA	2	0.5385	0.1007	5.35	0.000
				Constant		0.1763	0.4596	0.38	0.702

Final Estimates of Parameters

	MS	DF	SS	Type		Coef	SE Coef	T	P
	80		240526	AR	1	0.5140	0.1022	5.03	0.000
									3007
				MA	1	0.9801	0.0393	24.93	0.000
				Constant		0.0327	0.3062	0.11	0.915

Final Estimates of Parameters

	MS	DF	SS	Type		Coef	SE Coef	T	P
	79		281970	AR	1	-1.0201	0.1108	-9.20	0.000
									3569
				AR	2	-0.1871	0.1108	-1.69	0.095
				MA	1	-0.9948	0.0110	-90.06	0.000
				Constant		-0.84	11.92	-0.07	0.944

Final Estimates of Parameters

	MS	DF	SS	Type		Coef	SE Coef	T	P
	79		231097	AR	1	0.2413	0.1997	1.21	0.230
									2925
				MA	1	0.5947	0.1979	3.00	0.004
				MA	2	0.3777	0.1752	2.16	0.034
				Constant		0.1185	0.4397	0.27	0.788

Final Estimates of Parameters

	MS	DF	SS	Type		Coef	SE Coef	T	P
2963	78		231083	AR	1	0.1585	0.4678	0.34	0.736
				AR	2	0.0531	0.3077	0.17	0.863
				MA	1	0.5137	0.4472	1.15	0.254
				MA	2	0.4577	0.4558	1.00	0.318
				Constant		0.1189	0.4437	0.27	0.789

لاختيار النموذج المناسب سوف نستخدم معيار الاعلام الذاتي و الذي يعطي بالعلاقة

$$\mathrm{AIC(M)} = nln\sigma^2 + 2\mathrm{m}$$

حيث m عدد المعالم المقدرة في النموذج و نختار النموذج الذي يعطي $\underbrace{min}_{m} AIC(m)$ سوف نطبق النماذج على التوالي

جدول(2)

يوضح المقارنة بين النماذج بالاعتماد على مقياس AIC

Models	σ^2	n	$nln\sigma^2$	2m	$\mathrm{AIC(M)} = nln\sigma^2 + 2\mathrm{m}$
ARIMA(1,1,0)	3788	84	692.1259	4	696.1259
ARIMA(2,1,0)	3526	84	686.1052	6	692.1052
ARIMA(0,1,1)	3723	84	690.6719	4	694.6719
ARIMA(0,1,2)	2938	84	670.7807	6	676.7807
ARIMA(1,1,1)	3007	84	672.7306	6	678.7306
ARIMA(2,1,1)	3569	84	687.1234	8	695.1234
ARIMA(1,1,2)	2925	84	670.4082	8	678.4082
ARIMA(2,1,2)	2963	84	671.4924	10	681.4924

من خلال الجدول أعلاه ومقارنة النماذج نلاحظ بأنَّ أفضل نموذج ((ARIMA(0,1,2) لأنَّ قيمة AIC أقل مقارنة مع بقية النماذج. ومن ثم نقوم بفحص البواقي و توليد التنبؤات.

اختبارات على البواقي (Residual)

بعد اختيار النموذج الملائم للبيانات بالاعتماد على مقياس (AIC) نقوم بفحص البواقي Residual

١) اختبار (T-Test of the Mean)

Null Hypothesis $H_0: \mu_{(residual)} = 0$

Alternate Hypothesis $H_1: \mu_{(residual)} \neq 0$

جدول (٣) يوضح اختبار (T-Student)

Test	n	Mean	Std.	T	P-value	95% CI
T	83	3.49	53.43	0.59	0.554	(-8.18 , 15.15)

نلاحظ من خلال الجدول أنَّ قيمة (P-Value) للاختبار يساوي (0.554) وعند مقارنة هذه القيمة مع المستوى المعنوي (α=0.05) نلاحظ بأنَّها أكبر من (α=0.05) أي إنَّ الاختبار غير معنوي لايمكن رفض أنَّ متوسط البواقي صفر بمعني نقبل فرضية عدم.

٢) اختبار (Runs-Test) للعشوائية البواقي

H_0: Sequence is i.i.d (random)

H_a: Sequence is not i.i.d (random)

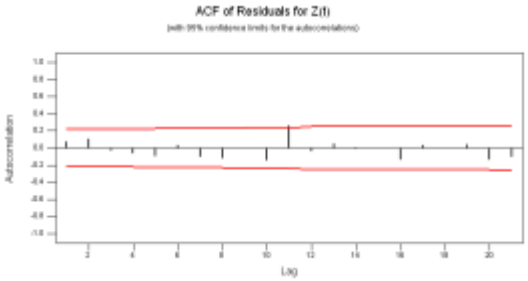
جدول (٤) يوضح اختبار (Run-Test)

K	Observed No.	Expected No.	Test is Significance at	Cannot reject at alpha
3.4860	34	41.1446	0.1027	0.05

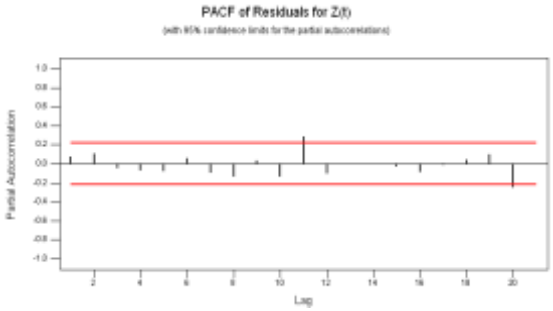
من خلال مقارنة قيمة الاخ تبار مع المستوى المعنوى نلاحظ بـأنَّ قيمة الاختبار معنوية و يساوي عند (0.1027) أي إنَّنا لانرفض فرضية عشوائية البواقي

٣) اختبار استقلالية البواقي Independence of residual

الاشكال الادناه يوضح دوال الترابط الذاتي (ACF) و الترابط الذاتي الجزئي (PACF)



شكل رقم (٧) يوضح الرسم البياني لدالة الارتباط الذاتي للبواقي



شكل رقم (٨) يوضح الرسم البياني لدالة الارتباط الذاتي الجزئي للبواقي

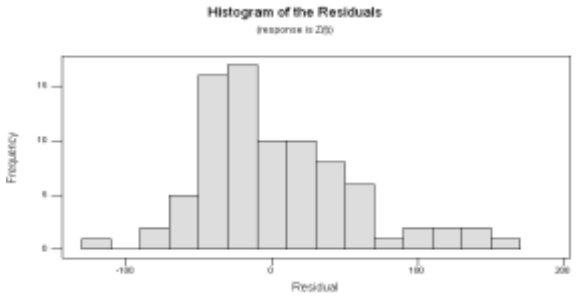
نلاحظ من خلال الشكلين أعلاه أنَّ أنماط الترابط الذاتي (Autocorrelation) و الترابط الذاتي الجزئي (Partial Autocorrelation) تتبع انماط متسلسلة الضجة البيضاء (White Noise)

4) اختبار طبيعة البواقي (Test for Normality Residual)

H_0: Sequence is Normally distributed

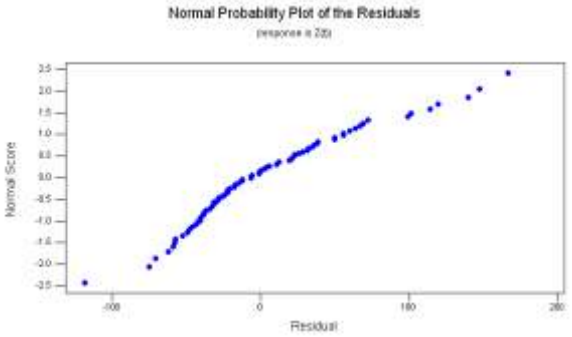
H_a: Sequence is Non-Normality distribution

1)نرسم المضلع التكراري للبواقي (Histogram of the residual)

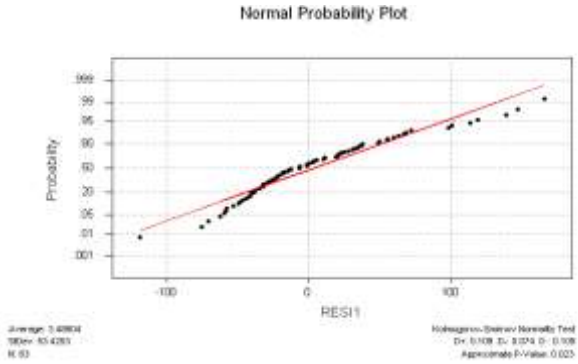


شكل(٩) يوضح رسم المضلع التكراري للبواقي (Residual)

2) نرسم الاحتمال الطبيعي للبواقي (Residual)



شكل(١٠) يوضح رسم احتمالية التوزيع الطبيعي للبواقي



شكل(١١)
يوضح رسم كولموجروف –سميرنوف للبواقي

نلاحظ من خلال شكل (١١) أنَّ قيمة (P-Value) لاختبار (K-S) يعطي (0.023) وهي أكبر من (٠.٠١) إذاً الاختبار معنوي عند ($\alpha = 0.01$) أي لانرفض فرضية طبيعية البواقي .

توليد التنبؤات Generating Forecasting

بعد مرحلة التشخيص و تـقدير معالم النموذج المقترح تاتي مرحلة التنبؤ ح يث تعد هذه المرحلة هي الخطوة الأخيرة في تحليل و تشخيص النموذج الملائم في السلاسل الزمنية و هي من

أهم أهداف بناء النموذج الإحصائي في السلاسل الزمنية، و على هذا الأساس تم التنبؤ بقيم عدد الأطفال المصابين بمرض التهاب المجاري التنفسية العليا من خلال النموذج المقترح و قوة النموذج المقدر من خلال حدود الثقة (٩٥%) المعتمد في بناء النماذج الاحصائية اضافة الى ذلك يتم اختبار دقة القيم التنبؤية و حدود الثقة لهذه القيم التنبؤية .

جدول (٥)

يوضح القيم التنبؤية لأعداد المرضى خلال الخمس سنوات القادمة

الفترة	٢٠١٢			٢٠١٣			٢٠١٤			٢٠١٥			٢٠١٦		
	عدد المرض Forecasting	حدود الثقة 95% Percent Limits		عدد المرض Forecasting	حدود الثقة 95% Percent Limits		عدد المرض Forecasting	حدود الثقة 95% Percent Limits		عدد المرض Forecasting	حدود الثقة 95% Percent Limits		عدد المرض Forecasting	حدود الثقة 95% Percent Limits	
		الحدود الدنيا Lower	الحدود العليا Upper		الحدود الدنيا Lower	الحدود العليا Upper		الحدود الدنيا Lower	الحدود العليا Upper		الحدود الدنيا Lower	الحدود العليا Upper		الحد ود الدنيا Lower	الحدود العليا Upper
1	117	40.7	253.5	121	-2.3	244	124	-0.01	247.7	128	0.9	254.7	130	٢.٤	٢٥٧.٤
2	118	-4.4	240.5	121	-2.1	244.3	124	0.2	248.1	128	1	254.9	130	٢.٥	٢٥٧.٦
3	118	-4.2	240.9	121	-1.9	244.6	124	0.4	248.4	128	1.2	255.1	130	٢.٦	٢٥٧.٩
4	119	-4	241.2	122	-1.7	245	125	0.6	248.7	128	1.3	255.4	130	٢.٧	٢٥٨.١
5	119	-3.8	241.5	122	-1.5	245.2	125	0.8	248.9	129	1.4	255.6	131	٢.٩	٢٥٨.٣
6	119	-3.7	241.8	122	-1.4	245.6	125	0.9	249.3	129	1.5	255.8	131	٣.٠٢	٢٥٨.٨
7	119	-3.5	242.1	122	-1.2	245.9	125	1.1	249.6	129	1.7	256	131	٣.١	٢٥٨.٩
8	120	-3.3	242.4	123	-0.9	246.2	126	1.3	249.9	129	1.8	256.3	131	٣.٣	٢٥٩.٠١
9	120	-3.1	242.7	123	-0.8	246.5	126	1.5	250.2	129	1.9	256.5	131	٣.٤	٢٥٩.٢
10	120	-2.9	243	123	-0.6	246.8	126	1.7	250.6	129	2	256.7	131	٣.٥	٢٥٩.٤
11	120	-2.7	243.4	123	-0.4	247.1	126	1.9	250.9	130	2.2	256.9	132	٣.٦	٢٥٩.٧
12	121	-2.5	243.7	124	-0.2	247.4	127	2.1	251.2	130	2.3	257.2	132	٣.٨	٢٥٩.٩

ثالثا : الاستنتاجات و التوصيات Conclusions & Recommendations

٣-١ الاستنتاجات Conclusions

من خلال هذه الدراسة استنتجنا مجموعة من النقاط هي:-

١) بالاعتماد على الجدول رقم (2) و من خلال المقارنة بين نماذج السلاسل الزمنية نلاحظ بأنَّ أفضل نموذج بين هذه النماذج هو نموذج (ARIMA(0,1,2)) لأنَّ قيمة مقياس (AIC) اقل مقارنة مع بقية النماذج وصيغة النموذج (ARIMA(0,1,2))هي:-

$$X_t = a_t + 0.4272\ a_{t-1} + 0.5385\ a_{t-2}$$

٤) من خلال رسم معاملات الأرتباط الذاتي للاخطاء الناتجة عن استخدام هذا النموذج بأنَّها تقع ضمن حدي الثقة وهذا ما يؤكد كفاءة اختيار نموذج السلسلة.

٥) من خلال الجدول رقم (5) نلاحظ بأنَّ عدد الأطفال المصابين حسب الجنس (ذكور و إناث) و الفئات العمرية في تزايد مستمر و هذا يعود الى عدم فاعلية الدواء المستخدم في معالجة مثل هذا النوع من المرض

٣-٢ التوصيات Recommendations

١) ضرورة إجراء بحوث أكثر للأطفال المصابين بهذا النوع من المرض لغرض الوصول الى أسباب حدوث المرض عندهم.

٢) نوصي المستشفيات الخاصة بالأطفال أن تعتمد على القيم التنبؤية المستخرجة من تطبيق النموذج برسم خططها المستقبلية.

٣) على الحكومة الاهتمام أكثر بالأطفال و توفير الدواء و فتح مراكز صحية أكثر لغرض تقليل الإصابة بهذا النوع من المرض.

رابعا : المصادر و المراجع References

١- المصادر العربية Arabic

[1] الخضيري ، محمد قدوري عبد، (١٩٩٦) " **دراسة مقارنة لط رائق و التنبؤ لبعض نماذج بوكس وجينكيز الموسمية** "، رسلة ماجستير في الاحصاء ، جامعة بغداد ، كلية الادارة و الاقتصاد

[2] بري ، د .عدنان ماجد عبدالرحمن ، **(2002)** ، "**طرق التنبؤ الإحصائي**" ، الجزء الاول ، جامعة الملك سعود، قسم الإحصاء و بحوث العمليات

[3] شعراوي ،سمير (٢٠٠٥)،"**مقدمة في التحليل الحديث للسلاسل الزمنية** "،الطبعة الأولى ،مركز النشر العلمي، جامعة الملك عبدالعزيز.

٢-المصادر الانكليزية English

[4] Abraham, B. and Ledoter, J. (1983). **Statistical Methods for Forecasting**, John Wiley, New York.

[5] Anderson, T. W. (1971). **The Statistical Analysis of Time Series**, John Wiley, NewYork.

[6] Wei, W. W. S. (1990). **Time Series Analysis Univariate and Multivariate Methods**, Addison Wesley.