



وزارة التعليم العالي والبحث العلمي

جامعة كربلاء

كلية الإدارة والاقتصاد

قسم الإحصاء

التنبؤ بأعداد المصابين بفيروس Covid-19 في محافظة كربلاء
المقدسة باستعمال أنموذجات السلاسل الزمنية اللاخطية

رسالة مقدمة الى مجلس كلية الإدارة والاقتصاد في جامعة كربلاء
وهي جزء من متطلبات نيل درجة الماجستير في علوم الإحصاء

تقدم بها

محمد صالح هاشم الأعرجي

بإشراف

أ. د جاسم ناصر حسين

2021م

1443 هـ

كربلاء المقدسة

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

﴿أَقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ﴿١﴾ خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ ﴿٢﴾﴾ أَقْرَأْ

وَرَبُّكَ الْأَكْرَمُ ﴿٣﴾ الَّذِي عَلَّمَ بِالْقَلَمِ ﴿٤﴾ عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمْ

﴿٥﴾ ﴿صَدَقَ اللَّهُ الْعَلِيِّ الْعَظِيمِ﴾

سورة العلق
الآية (1-5)

الإهداء

إلى صاحب السيرة العطرة، والفكر المستنير
فلقد كان له الفضل الأول في بلوغي التعليم العالي
(والدي الحبيب)، أطال الله في عمره
المنبع المحبة والحنان والعطاء
من انارت دربي بالدعاء
(أمي الغالية)، أطال الله في عمرها
إلى إخوتي من كان لهم بالغ الأثر في كثير من العقبات والصعاب
إلى زوجتي ورفيقة الكفاح في مسيرة الحياة
إلى أساتذتي الكرام ممن لم يتوانوا في مديد العون لي
أهدي إليكم هذا الجهد وأتمنى أن يحوز على رضاكم
داعياً المولى عز وجل أن يطيل في أعماركم ويرزقكم بالخيرات

الباحث

الشكر والتقدير

الحمد لله رب العالمين والصلاة والسلام على أشرف الأنبياء والمرسلين سيدنا محمد
وعلى آله وصحبه ومن تبعهم بإحسان إلى يوم الدين، وبعد . . .
فإنني أشكر الله تعالى على فضله إذ أتاح لي إنجاز هذا العمل بفضله، فله الحمد أولاً
وآخرًا.

ثم أشكر أولئك الأخيار الذين مدوا لي يد المساعدة في هذه المدة وفي مقدمتهم
أستاذي المشرف على الرسالة الأستاذ الدكتور / جاسم ناصر حسين على ما
أسداه لي من نصائح وارشادات كانت مثل النبراس المنير في كل خطواتي . كما
أتقدم بالشكر والتقدير والاحترام إلى الإساتيد الفضلاء في لجنة المناقشة على ما بذلوه
من جهد في قراءة الرسالة وعلى ما أبدوه من ملاحظات قيمة بما يثري
الرسالة . كما ان شكري موجه لعمادة كلية الإدارة والاقتصاد وبالخصوص اساتذتي
الفضلاء في قسم الإحصاء على ارشاداتهم القيمة طوال فترة الدراسة .

الباحث

قائمة المحتويات		
الصفحة	العنوان	رقم الفقرة
ذ	المستخلص Abstract	
الفصل الأول		
1	مقدمة عامة General Introduction	(1-1)
2	مشكلة الرسالة Thesis Problem	(2-1)
2	هدف الرسالة Thesis Goal	(3-1)
3	أهمية الرسالة The importance of thesis	(4-1)
3-6	الاستعراض المرجعي Literature Review	(5-1)
الفصل الثاني		
7	تمهيد	(1-2)
7-8	تعريف السلسلة الزمنية Define the time series	(2-2)
8	خصائص السلسلة الزمنية Properties the time series	(3-2)
9	الاستقرارية Stationary	(1-3-2)
9	طريقة الرسم البياني Graph method	(1-1-3-2)
11-13	أختبار ديكي-فولر الموسع Augmented Dickey Fuller test	(2-1-3-2)
13-14	الخطية Linearity	(2-3-2)
14-15	اختبار ماك لود ولي (ML) Mcleod and Li	(1-2-3-2)
15-17	اختبار مان كيندال Mann-Kendall	(2-2-3-2)
17	التجانس Homogeneity	(3-3-2)
17-18	اختبار التجانس الطبيعي القياسي Standard normal homogeneity	(1-3-3-2)
19	اختبار فون نيومان Von Neumann test	(2-3-3-2)
19-20	أنموذجات السلاسل الزمنية اللاخطية Nonlinear time series models	(4-2)
20-23	الأنموذج الأسّي Exponential Model	(1-4-2)
23-25	أنموذج فون بيرتلانفي Von Bertalanffy model	(2-4-2)
25-27	أنموذج جمبرتز Gompertz Model	(3-4-2)
27-30	الأنموذج اللوجستي Logistic Model	(4-4-2)
30-32	أنموذج چاپمان ريتشارد Chapman-Richard model	(5-4-2)
32	منهجية بوكس جينكز Box-Jenkis لتحليل السلاسل الزمنية	(5-2)
32	مرحلة التشخيص Diagnostic stage	(1-5-2)

33	مرحلة التقدير Appreciation stage	(2-5-2)
38	مرحلة اختيار الأنموذج الأفضل The stage of choosing the Best model	(3-5-2)
34-35	معيار معلومات أكاكي Akanke's Information Criterion	(1-3-5-2)
43-44	معيار حنان كوين Hannan-Quinn Criterion	(2-3-5-2)
44	معيار معلومات بيز Bayesian Information criterion	(3-3-5-2)
37	مرحلة اختبار صلاحية الأنموذج The model validity testing phase	(4-5-2)
37-38	اختبار معنوية معلمات الأنموذج The significance of the model parameters test	(1-4-5-2)
38	اختبار المعنوية الكلية للأنموذج Test the overall significance of the model	(2-4-5-2)
39-40	اختبار البواقي المتسلسلة Sequential residual test	(3-4-5-2)
40	التنبؤ Forecasting	(5-5-2)
40	مقاييس الدقة التنبؤية Predictive Accuracy Metrics	(1-5-5-2)
40-41	متوسط مربعات الأخطاء Mean Square Error	(1-1-5-5-2)
41	جذر متوسط مربعات الأخطاء Root Mean Square Error	(2-1-5-5-2)
الفصل الثالث - الجانب التطبيقي		
42	تمهيد	(1-3)
43-44	فيروس كورونا واعراضه وطرائق الوقاية والعلاج	(2-3)
45	خصائص السلسلة الزمنية Properties of the time series	(3-3)
45	استقرارية السلسلة الزمنية Stationary The time series	(1-3-3)
45	طريقة الرسم البياني Graph method	(1-1-3-3)
46-48	اختبار ديكي-فولر الموسع Augmented Dickey Fuller test	(2-1-3-3)
48	الخطية Linearity	(2-3-3)
48	اختبار ماك لود و لي Mcleod and Li	(1-2-3-3)
48	اختبار مان كيندال Mann-Kendall	(2-2-3-3)
49	التجانس Homogeneity	(3-3-3)
50	اختبار التجانس الطبيعي القياسي Standard normal homogeneity	(1-3-3-3)
51	اختبار فون نيومان Von Neumann test	(2-3-3-3)

51	تقدير معلمات الأنموذجات اللاخطية Estimating Parameters of Non Linear Models	(4-3)
51-54	Exponential Model الأنموذج الأسّي	(1-4-3)
55-58	Von Bertlanffy model أنموذج فون بينلانفي	(2-4-3)
58-61	Gompertz Model أنموذج جمبرتز	(3-4-3)
61-64	Logistic Model الأنموذج اللوجستي	(4-4-3)
64-68	Chapman-Richard model أنموذج جابمان ريتشارد	(5-4-3)
68-69	Choosing the Best model اختيار الأنموذج الأفضل	(5-3)
69-71	Forecasting التنبؤ	(6-3)
71	Predictive Accuracy Metrics مقاييس الدقة التنبؤية	(7-3)
	الفصل الرابع	
72	Conclusions الاستنتاجات	(1-4)
73	Recommendations التوصيات	(2-4)
74-78	References المصادر	
79-96	الملاحق	
	المستخلص باللغة الانجليزية	

قائمة الجداول		
الصفحة	عنوان الجدول	الجدول
44	اعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة للمدة من (2020-6-1) الى (2020-9-30)	1-3
46	ختبار Dickey-fuller Augmented الموسع في حالة انموذج لا يتضمن الحد الثابت ولا الاتجاه الزمني	2-3
47	ختبار Dickey-fuller Augmented الموسع في حالة انموذج يتضمن الحد الثابت فقط	3-3
47	اختبار Dickey-fuller Augmented الموسع في حالة انموذج يتضمن الحد الثابت والاتجاه الزمني	4-3
51	المعلومات المقدرة للأنموذج الأسّي Exponential model	5-3
53	القيم المقدرة للأنموذج الأسّي Exponential model	6-3
57	المعلومات المقدرة لأنموذج فون بيرتلانفي Von Bertalanffy	7-3
58	القيم المقدرة لأنموذج فون بيرتلانفي Von Bertalanffy model	8-3
58	المعلومات المقدرة لأنموذج جمبرتز Gompertz model	9-3
60	القيم المقدرة لأنموذج جمبرتز Gompertz model	10-3
61	المعلومات المقدرة للأنموذج اللوجستي Logistic model	11-3
63	القيم المقدرة للأنموذج اللوجستي Logistic model	12-3
64	المعلومات المقدرة لأنموذج جابمان ريتشارد Chapman Richard model	13-3
66	القيم المقدرة لأنموذج جابمان ريتشارد Chapman-Richard model	14-3
69	قيم معايير المفاضلة بين الأنموذجات اللاخطية	15-3
69	القيم التنبؤية للأنموذج الأسّي Exponential model	16-3
70	القيم التنبؤية لأنموذج فون بيرتلانفي Von Bertalanffy model	17-3
70	القيم التنبؤية لأنموذج جمبرتز Gompertz model	18-3
70	القيم التنبؤية للأنموذج اللوجستي Logistic model	19-3
70	القيم التنبؤية لأنموذج جابمان ريتشارد Chapman Richard model	20-3
71	مقاييس القوة التنبؤية للأنموذجات اللاخطية	21-3

قائمة الأشكال		
الصفحة	عنوان الشكل	الشكل
10	سلسلة زمنية مستقرة	1-2
10	سلسلة زمنية غير مستقرة	2-2
20	منحنى النموذج الأسي Exponential Growth model	3-2
23	منحنى نمو فون بيرتلانفي Von Bertalanffy Growth Curve	4-2
25	منحنى نمو جمبرتز Curve Model Gompertz	5-2
27	منحنى النمو اللوجستي Logistic Growth Curve	6-2
30	منحنى نمو جابمان ريتشارد Chapman-Richard Growth Curve	7-2
45	التمثيل البياني للسلسلة الزمنية لأعداد المصابين بفيروس كوفيد-19	1-3
49	التمثيل البياني لاختبار Mann-Kendal	2-3
50	التمثيل البياني لاختبار التجانس الطبيعي القياسي	3-3
54	رسم القيم المقدرة للنموذج الأسي Exponential model	4-3
58	رسم القيم المقدرة للنموذج Von Bertalanffy	5-3
61	رسم القيم المقدرة للنموذج جمبرتز Gompertz model	6-3
64	رسم القيم المقدرة للنموذج اللوجستي Logistic Model	7-3
67	رسم القيم المقدرة للنموذج جابمان-ريتشارد Chapman-Richard	8-3
68	الشكل البياني للنموذجات اللاخطية	9-3

Abstract

المستخلص

ان الزيادة الحاصلة في أعداد الإصابات اليومية بفيروس Covid-19 حول العالم ولا سيما العراق وبالخصوص في محافظة كربلاء المقدسة وما خلفه هذا الفيروس من خسائر بشرية ومادية واقتصادية وسبب كذلك اثاراً اجتماعية ونفسية لفئات المجتمع كافة، كان هذا الدافع الرئيس للقيام بهذه الدراسة التي تهدف الى اختيار أفضل أنموذج للتنبؤ بأعداد الإصابات اليومية بفيروس Covid-19 في محافظة كربلاء المقدسة من بين إنموذجات السلاسل الزمنية اللاخطية المستعملة في هذه الدراسة باستعمال المعايير الإحصائية (AIC, BIC, H-Q).

تضمنت هذه الرسالة جانبين (نظري وتطبيقي) يتضمن الجانب النظري خصائص السلسلة الزمنية (الاستقرارية، الخطية والتجانس) واهم الاختبارات الإحصائية التي تستعمل لاختبار هذه الخصائص. كذلك يتضمن بعض الأنموذجات اللاخطية للسلاسل الزمنية وهي (Exponential, Von Bertalanffy, Gompertz, Logistic and Chapman Richard) ويتضمن أيضا طريقة تقدير معالم هذه الأنموذجات وكذلك بعض الاختبارات والمعايير الإحصائية لقياس معنوية المعالم وجودة الأنموذجات وكذلك معايير المفاضلة بين الأنموذجات ومقاييس القوة التنبؤية. اما الجانب التطبيقي فقد تم تطبيق جميع ما ذكر في الجانب النظري على عينة من الأشخاص المصابين بفيروس Covid-19 وبحجم 122 مشاهدة وللمدة من (2020-6-1) الى (2020-9-30). وبينت النتائج انها سلسلة غير مستقرة، غير خطية وغير متجانسة. وان الأنموذج الأفضل هو الأنموذج اللوجستي Logistic model وذلك باستعمال معايير المعلومات الإحصائية (AIC, BIC, H-Q). والذي يعطي اقل قيمة لهذه المعايير. كما واطهرت النتائج ان هنالك تقارباً بين القيم التنبؤية للأنموذجات الخمسة المستعملة في هذه الرسالة. كذلك مقاييس القوة التنبؤية (MSE, RMSE) إذ أظهرت النتائج ان الأنموذج اللوجستي Logistic model يعطي اقل قيمة لهذه المقاييس وهذا يدل على كفاءة الأنموذج المقترح للتنبؤ بأعداد المصابين بفيروس Covid-19.

الكلمات المفتاحية: السلاسل الزمنية اللاخطية، الانموذج اللوجستي، التنبؤ، خصائص السلسلة الزمنية، كوفيد-19.

General Introduction

(1-1) مقدمة عامة

يعد التنبؤ من الموضوعات الأساسية والمهمة في مساعدة الجهات المعنية وأصحاب القرار باتخاذ القرارات الصحيحة والإجراءات اللازمة في العديد من المجالات الصحية والاقتصادية والاجتماعية وان أحد اهم الأساليب المستعملة للتنبؤ هو أسلوب السلاسل الزمنية. ونظرا لأهمية الجانب الصحي الذي يهتم بحياة الانسان ويقيه من الامراض بصورة عامة ولاسيما الامراض او الفيروسات المعدية ومنها فيروس Covid-19 إذ يعد هذا الفيروس من الفيروسات المعدية والخطرة التي تهدد حياة الانسان وصحته. ولهذا فقد اهتمت هذه الرسالة بدراسة فيروس Covid-19 بالاعتماد على اعداد الإصابات اليومية في محافظة كربلاء المقدسة وللمدة من (2020-6-1) الى (2020-9-30) بهدف وضع أنموذجات للتنبؤ بأعداد المصابين بفيروس Covid-19 في المستقبل وارسالها الى الجهات المختصة لوضع الخطط المستقبلية لمواجهة هذه الجائحة.

شهدت السنوات السابقة اهتماما كبيرا في موضوع التنبؤ وظهور العديد من الأساليب الحديثة للتنبؤ. ومنها أنموذجات السلاسل الزمنية وهي على نوعين خطية ولا خطية وفي هذه الدراسة سيتم استعمال الأنموذجات اللاخطية للتنبؤ بأعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة. وهذه الأنموذجات هي (Exponential, Von Bertalanffy, Gompertz, Logistic and Chapman Richard).

اما هيكلية الرسالة فقد تضمنت أربعة فصول، يحتوي الاول على مقدمة عامة، مشكلة الرسالة، هدف الرسالة، اهمية الرسالة والاستعراض المرجعي. اما الثاني فقد تضمن خصائص السلسلة الزمنية وبعض الاختبارات الإحصائية الخاصة بها. وكذلك بعض الأنموذجات اللاخطية وطريقة تقدير معالمها وبعض المقاييس والمعايير الإحصائية الخاصة بهذه الأنموذجات. اما الثالث فقد تضمن الجانب التطبيقي ويحتوي على تمهيد للفصل وتعريف فيروس Covid-19 كذلك تطبيق جميع الاختبارات الإحصائية للسلسلة الزمنية قيد الدراسة، وكذلك تقدير معالم هذه الأنموذجات واختبار معنوية المعالم والمعنوية الكلية للأنموذجات والتنبؤ بها. وتضمن الرابع أهم الاستنتاجات والتوصيات التي توصل اليها الباحث.

Thesis Problem**(2-1) مشكلة الرسالة**

ان فيروس Covid-19 هو مشكلة صحية وجائحة عالمية شملت جميع انحاء العالم دون استثناء. ولاسيما في العراق وبالخصوص محافظة كربلاء المقدسة إذ بلغت اعداد المصابين من بداية الجائحة ولغاية يوم (2021/9/1) ما يقارب (1895459) مصاباً وما يقارب (20878) حالة وفاة. أي ان أعداد الوفيات عالية جدا وهي بكل تأكيد تمثل مشكلة اقتصادية واجتماعية وصحية كبيرة. وهذا ما اثار اهتمامنا في البحث والتحري عن هذه الظاهرة وكذلك التنبؤ بأعداد المصابين في المستقبل.

Thesis Goal**(3-1) هدف الرسالة**

تهدف الرسالة الى التنبؤ بأعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة باستعمال أنموذجات السلاسل الزمنية اللاخطية، واختيار أفضل أنموذج للتنبؤ باستعمال بعض المعايير الإحصائية.

The importance of thesis**(4-1) اهمية الرسالة**

ان الزيادة الحاصلة في اعداد المصابين بفيروس Covid-19 بشكل ملحوظ والخسائر البشرية والمادية الذي تسبب بها هذا الفيروس وكذلك المشاكل الاقتصادية والصحية والاجتماعية والتي تشمل فئات المجتمع كافة. من هنا جاءت فكرة كتابة هذه الرسالة عن التنبؤ بأعداد الإصابات اليومية بفيروس Covid-19 وكذلك التحليل الاحصائي الشامل، إذ تقدم هذه الرسالة اهم ما تم التوصل اليه من استنتاجات امام انظار المسؤولين وصانعي القرار في المجالين الاقتصادي والصحي والجهات المعنية الأخرى لغرض القيام بالإجراءات اللازمة ووضع التدابير والخطط المستقبلية، وتقديم الحل المناسب للسيطرة على هذه الجائحة في الوقت الحاضر وحلها في المستقبل.

Literature Review

(5-1) الاستعراض المرجعي

في هذا المبحث نقدم مراجعة لبعض الاعمال المهمة المتعلقة بأنموذجات السلاسل الزمنية اللاخطية التي تم استعمالها في هذه الرسالة وسنراجع أيضا بعض البحوث التي تم توظيفها مع أنموذجات السلاسل الزمنية اللاخطية الخاصة بإصابات فيروس Covid-19. ومن الجدير بالذكر ان هنالك العديد من الدراسات والبحوث التي تم استعمال أنموذجات السلاسل الزمنية اللاخطية ولكننا قمنا بسردها أهمها واقربها الى موضوع بحثنا تلافيا للإسهاب.

• في عام 2008 قدم الباحثان (Wu, Trappy) [38] بحثهما الذي يهتم بمقارنة دقة التنبؤ بين نوعين من أنموذجات النمو اللاخطية وهما Gompertz and Logistic model باستعمال مجموعة من بيانات دورة حياة المنتج القصيرة. إذ تم تقييم أداء الأنموذجين باستعمال المقاييس المناسبة. وقد أظهرت النتائج ان الأنموذج اللوجستي Logistic model هو الأنموذج الأنسب للتنبؤ.

• وفي العام نفسه قدم الباحثان (Topal, Bolukbasi) [37] بحثهما الذي يهتم بدراسة تقدير منحنيات نمو تسمين الدجاج باستعمال بيانات وزن الجسم الأسبوعية. إذ تمت مقارنة أنموذجات منحنى النمو وتحديد الأنموذج الأنسب لبيانات عمر تسمين الدجاج. إذ أظهرت النتائج ان أنموذج Weibull وأنموذج Morgan-Mercer-Flodin هما أفضل النماذج التي تلائم البيانات. وان الأنموذج اللوجستي Logistic Model هو الأنموذج الأسوأ.

• وفي ذات العام قدم (Gbangboche وآخرون) [24] بحثهم الذي يهتم بمقارنة جودة ملائمة أربع أنموذجات نمو لدراسة نمو الأغنام عن طريق عدة معايير إحصائية ومنها معيار معلومات أكايكي (AIC). إذ اشارت النتائج الى ان انموذج Brody هو أفضل انموذج يناسب نمو الأغنام.

• وفي عام 2010 قدم الباحث (Tjorve) [36] بحثه الذي يهدف الى تركيب أنموذج Chapman-Richards وأنموذجات خمسة أخرى لمجموعات متعددة من البيانات الصناعية. وتم اجراء مقارنة بين هذه الأنموذجات بناء على معيار معلومات أكايكي المصحح (AIC_c). إذ كان أداء انموذج Chapman-Richards جيدا لجميع البيانات.

- وفي عام 2012 قدم (Mansano وآخرون) [29] بحثهم الذي يهتم بدراسة معدل نمو الضفادع باستعمال أنموذجات النمو اللاخطية. إذ تم استعمال أربع أنموذجات لتناسب زيادة الوزن وبيانات الطول للضفادع. وتم استعمال عدة معايير إحصائية لاختيار الأنموذج الأفضل من بين هذه الأنموذجات. فقد توصلت النتائج الى ان انموذج Gompertz وانموذج Logistic هما الأنسب لوصف معدل النمو.
- وفي عام 2015 قدم الباحث Hossin [25] بحثه الذي يهدف الى وصف نمط النمو في اغنام الشمال الإيرانية باستعمال العديد من الأنموذجات اللاخطية. والمقارنة بين هذه الأنموذجات واختيار الأنموذج الأفضل لملائمة منحنى النمو وذلك باستعمال عدة معايير إحصائية ومنها معيار أكايكي (AIC). إذ أظهرت النتائج ان أنموذج Richard هو أفضل انموذج لملائمة منحنى النمو والأنموذج الأسّي هو الأنموذج الأسوأ.
- وفي عام 2018 قدم الباحث (حموي) [3] بحثه الذي يهتم بالمقارنة بين نماذج النمو اللاخطية للأنواع السمكية المختلفة باستعمال معياري معلومات أكايكي (AIC) وشوارترز (SC). إذ أظهرت النتائج ان الأنموذج الأسّي وانموذج Von Bertalanffy هما أفضل الأنموذجات من بين الأنموذجات المستعملة في البحث.
- وفي عام 2019 قدم (Mohammadi وآخرون) [31] بحثهم الذي يهدف الى مقارنة عدة أنموذجات نمو لا خطية لاختيار الأنموذج الأفضل الذي يصف منحنى النمو لمجموعة من سلالات الأغنام الإيرانية. وتوصلت النتائج الى ان انموذج Von Bertalanffy هو الأنموذج الأفضل من بين الأنموذجات المستعملة في هذا البحث وذلك باستعمال عدة مقاييس او معايير للمقارنة ومنها معيار معلومات أكايكي (AIC).
- وفي عام 2020 قدم (Uba وآخرون) [39] بحثهم الذي يهتم في تحليل الاتجاه الوبائي والتنبؤ بأعداد المصابين بفيروس Covid-19 في البرازيل باستعمال مجموعة من الأنموذجات اللاخطية. وتوصلت النتائج الى ان الأنموذج الأفضل هو انموذج Morgan-Mercer-Flodin بناء على عدة معايير إحصائية ومنها معيار معلومات أكايكي (AIC).

- وفي العام نفسه قدم (Celik وآخرون) [20] بحثهم الذي يهدف الى شرح التغيرات في مؤشرات تفشي فيروس Covid-19 في الصين مع مجموعة من الأنموذجات الخطية واللاخطية للسلاسل الزمنية. وكذلك تحديد أفضل انموذج رياضي من بين الأنموذجات المستعملة. إذ أظهرت النتائج ان انموذج Sloboda هو الأنموذج الأفضل من بين الأنموذجات اللاخطية.
- وفي العام نفسه قدم (Ahmadi وآخرون) [18] بحثهم في هذا البحث تم استعمال أنموذجين لا خطيين للسلاسل الزمنية وهما (Gompertz model and Von Bertalanffy model) للتنبؤ بأعداد المصابين بفيروس Covid-19 في إيران. وتوصلت النتائج الى أن إنموذج Gompertz أفضل من أنموذج Von Bertalanffy للتنبؤ.
- وفي العام نفسه قدم (Jia وآخرون) [27] بحثهم الذي تم فيه استعمال ثلاث أنموذجات لا خطية وهي (Von Bertalanffy, Gompertz and Logistic) تم أولاً تحليل الاتجاهات الوبائية لفيروس Covid-19 وكذلك التنبؤ بهذين الأنموذجات. إذ توصلت نتائج هذه الدراسة الى ان نتائج التنبؤ تختلف باختلاف المعلمات والمناطق المختلفة وبشكل عام فإن الأنموذج اللوجستي Logistic model هو الأنموذج الأفضل من بين الأنموذجات الثلاثة التي تمت دراستها في هذا البحث. ويأتي بعده أنموذج Gompertz وبعده انموذج Von Bertalanffy.
- وفي العام نفسه قدم (Medina وآخرون) [30] بحثهم الذي يهدف الى اختبار صلاحية أنموذج Gompertz وأنموذج Logistic للتنبؤ بالحالات المؤكدة والوفيات بفيروس Covid-19 في كوبا. إذ أظهرت النتائج ان كلا الأنموذجين ذو ملائمة جيدة.
- في عام 2021 قدم (Celik وآخرون) [20] بحثهم الذي يهدف الى المقارنة بين مجموعة من الأنموذجات للسلاسل الزمنية الخطية وغير الخطية للتنبؤ بالعدد الإجمالي بالإصابات والوفيات بفيروس كوفيد-19 في الصين. وتم استعمال عدة معايير لاختيار الأنموذج الأفضل ومنها معيار معلومات أكايكي (AIC). إذ ظهرت نتائج أنموذجات النمو اللاخطية للعدد الإجمالي للحالات ان انموذج Sloboda هو الأفضل.

نلحظ من استعراضنا للدراسات السابقة مدى حداثة موضوع التنبؤ بأعداد الإصابات بفيروس Covid-19 باستعمال أنموذجات السلاسل الزمنية اللاخطية، وعلى حد علم الباحث فقد لوحظ ندرة الدراسات العربية التي تناولت هذا الموضوع، ولأن الإصابات بفيروس Covid-19 تعد المشكلة الأهم في عصرنا الحالي، لذلك تم تطبيق أنموذجات السلاسل الزمنية اللاخطية للتنبؤ بأعداد الإصابات بفيروس Covid-19 في محافظة كربلاء المقدسة

(1-2) تمهيد

خلف فيروس كورونا أزمة صحية تهدد حياة الانسان وكذلك أزمة اقتصادية عالمية اثرت على جميع دول العالم ولاسيما بلدنا الحبيب، ولأجل الاعتماد على استراتيجيات ودراسات دقيقة لمواجهة هذا الوباء من الخدمات الطبية والإجراءات الوقائية فان من المهم والضروري الاهتمام بموضوع التنبؤ إذ ظهرت العديد من البحوث والدراسات التي تهتم بهذا المجال ومن هذه الدراسات: التنبؤ باستعمال أنموذجات السلاسل الزمنية (الخطية واللاخطية) إذ اسهمت هذا الأنموذجات وبدرجة كبيرة في نمذجة العديد من الظواهر ومنها الصحية والاجتماعية والاقتصادية.

يعد موضوع تحليل السلاسل الزمنية من الموضوعات الإحصائية المهمة التي تتناول سلوك الظاهرة وتفسيرها في مدد زمنية معينة، ويمكن اجمال فوائد تحليل السلاسل الزمنية بالحصول على وصف دقيق للصفات التي تتولد منها السلسلة الزمنية، وبناء انموذج احصائي لتفسير سلوك السلسلة الزمنية واستعمال النتائج لغرض التنبؤ بالقيم المستقبلية للظاهرة او السلسلة الزمنية [10]، إذ سيتناول هذا الفصل بعض التعاريف والمفاهيم الأساسية للسلاسل الزمنية وخصائصها كذلك سنتطرق الى بعض أنموذجات السلاسل الزمنية غير الخطية ومراحل تحليل هذه الأنموذجات وطرائق تقدير معالم هذه الأنموذجات ومعايير تقويم جودة هذه الأنموذجات والتنبؤ بها. ومنها الأنموذج الأسّي Exponential وانموذج جابمان-ريتشارد Chapman-Richard وأنموذج جمبرتز Gompertz والأنموذج اللوجستي Logistic وانموذج فون بيرتلانفي Von Bertalanffy و..... الخ من الأنموذجات اللاخطية. وسيتم دراسة الأنموذجات الخمسة المذكورة آنفاً فقط.

Define the time series

(2-2) تعريف السلسلة الزمنية

السلسلة الزمنية هي مجموعة من المشاهدات والبيانات لقيم ظاهرة ما تكون مأخوذة في أوقات زمنية محددة، وغالبا ما تكون متساوية مثل (يوم، شهر، سنة الخ) حسب طبيعة الدراسة [15]. تتكون السلسلة الزمنية من متغير واحد او أكثر من المتغيرات التوضيحية (متغيرات مستقلة) مثل متغير (الزمن، العمر، اعداد الإصابات بفيروس كوفيد-19، درجة الحرارة، الخ) حسب طبيعة

الظاهرة المدروسة والآخر متغير الاستجابة (متغير تابع) وهو قيمة الظاهرة المدروسة، تكون السلسلة الزمنية على نوعين اعتماداً على طبيعة المشاهدات الموجودة في السلسلة، أما إن تكون السلسلة الزمنية متقطعة مثل السلاسل الزمنية السكانية والاقتصادية والفيزيائية وغيرها أو تكون السلسلة الزمنية مستمرة (متصلة) مثل درجات الحرارة والرطوبة.

إن الخطوة الأولى في تحليل السلسلة الزمنية هي رسم بيانات السلسلة الزمنية لغرض الحصول على بعض من ملامح السلسلة التي تتمثل بخصائص المكونات الرئيسية للسلسلة الزمنية مثل (الاتجاه العام والتأثيرات الموسمية والدورية).

إن دراسة وتحليل السلاسل الزمنية هو لتحقيق عدد من الأهداف منها وصف السمات أو الملامح الرئيسية للسلسلة الزمنية وبناء أنموذج احصائي وقد يكون إما أنموذج سببي أو أنموذج ذو المتغير الواحد أو المتعدد. ويعد التنبؤ بقيم المشاهدات المستقبلية للظاهرة المدروسة من أهم أهداف دراسة وتحليل السلاسل الزمنية. ويعد من أكثر الأهداف الذي لاقى شعبية هائلة بالنسبة لدراسة الإحصاء والذي من أجله صدرت العديد من الكتب العربية والأجنبية وكذلك الآلاف من الأبحاث المتخصصة في هذا المجال. وهو يمثل الهدف النهائي في مجال تحليل السلاسل الزمنية [6]. وكما هو معروف قبل إجراء التحليل للسلسلة الزمنية لا بد من معرفة خصائصها والتي تتمثل في (الاستقرارية والخطية والتجانس).

(2-3) خصائص السلسلة الزمنية Properties the time series

تم تخصيص هذا الفرع لعرض أدوات تحليل السلاسل الزمنية. ولاختبار خصائص السلسلة الزمنية، ومنها الكشف عن استقرارية السلسلة الزمنية من عدمها وذلك عن طريق الرسم البياني للسلسلة الزمنية واختبار جذر الوحدة Dickey Fuller. وكذلك اختبار خطية السلسلة الزمنية عن طريق مجموعة من الاختبارات الإحصائية ومنها اختبار Mann-Kendal واختبار Mcleod and Li. فضلاً عن اختبار تجانس السلسلة الزمنية عن طريق اختبار التجانس الطبيعي القياسي Standard Normal Homogeneity test واختبار Von Neumann.

(1-3-2) الأستقرارية

Stationary

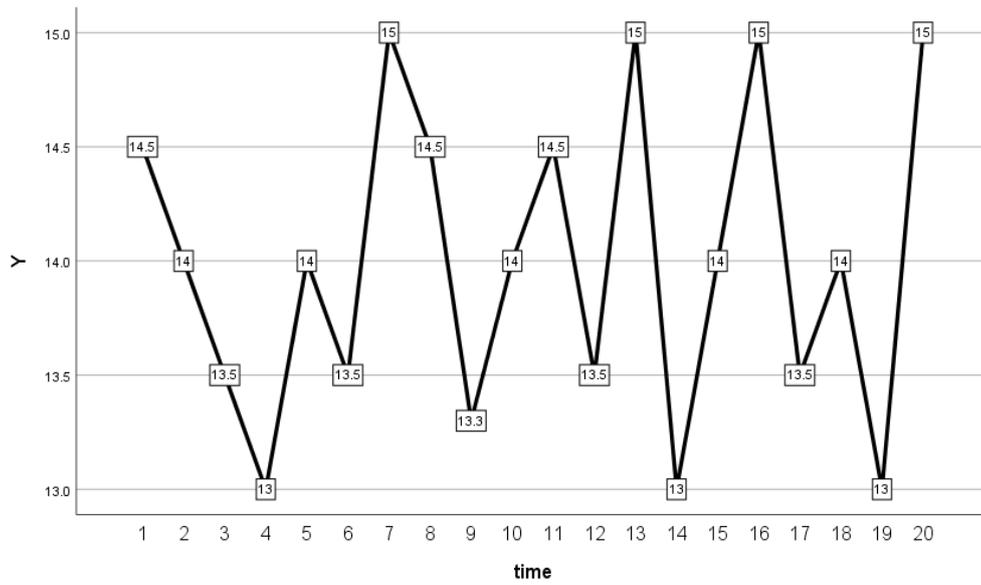
السلسلة الزمنية المستقرة تكون في حالة خاصة من الموازنة الإحصائية أي ان خصائصها لا تتأثر بالزمن أي انها تمتلك وسطاً حسابياً ثابتاً وتبايناً ثابتاً مع الزمن، أي تكون السلسلة الزمنية مستقرة إذا كانت الخصائص الاحتمالية لا تتأثر بتغير الزمن. بمعنى اخر إذا لم يكن هنالك اتجاه الى الأعلى او الى الأسفل في المعدل عبر الزمن. وكذلك يقال ان السلسلة الزمنية (Y_t) تامة الاستقرارية اذا كانت عزومها لا تتأثر بتغير الزمن [6]. "وان التوزيع الاحتمالي المشترك للمتغيرات $(X_{t1} \cdot X_{t2} \dots \dots X_{tn})$ هو نفسه التوزيع الاحتمالي المشترك للمتغيرات $(X_{t1+k} \cdot X_{t2+k} \dots \dots X_{tn+k})$ ". وتعد دراسة الاستقرارية من المراحل الأولى في دراسة وتحليل السلاسل الزمنية [2].

وهناك العديد من الطرائق والاختبارات للكشف عن استقراريه او عدم استقرارية السلسلة الزمنية. وهنا سنكتفي بطريقتين فقط وهما.

(1-1-3-2) طريقة الرسم البياني

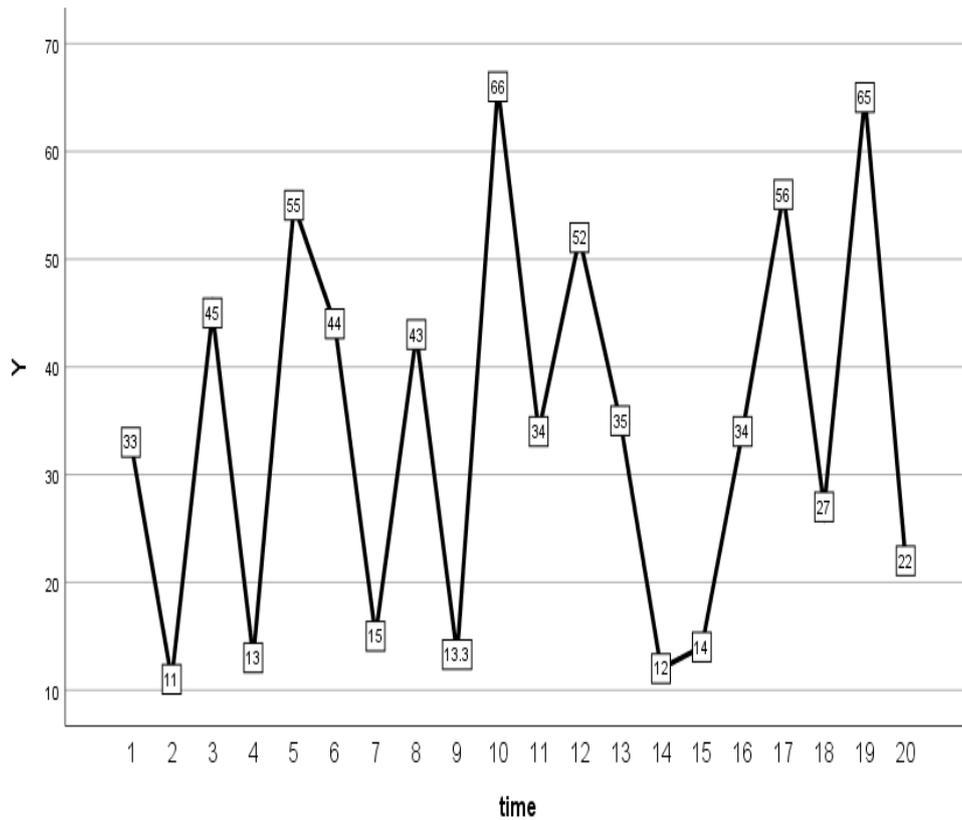
Graph method

يمكن الكشف عن استقراريه السلسلة الزمنية من عدمها وذلك عن طريق رسم المتغير (Y_t) على المحور العمودي ضد الزمن فاذا كانت المشاهدات في المتغير (Y_t) تزداد او تنخفض مع الزمن فهذا يعني ان السلسلة غير مستقرة. اما إذا كان رسم مشاهدات متغير معين عبر الزمن تتوزع حول وسطها الحسابي إذ يمكننا حصرها بين قيمتين (حد اعلى وحد أدنى) فعند ذلك تكون السلسلة الزمنية مستقرة. ويمكن تمثيل السلسلة الزمنية المستقرة وغير المستقرة كما في الشكل (1-2) و الشكل (2-2) [2].



الشكل (1-2) مثال على سلسلة زمنية مستقرة.

(المصدر من إعداد الباحث باستخدام برنامج SPSS)



الشكل (2-2) يمثل سلسلة زمنية غير مستقرة

(المصدر من إعداد الباحث باستخدام برنامج SPSS)

Augmented Dickey and Fuller test (2-1-3-2) اختبار ديكي-فولر الموسع

يحتاج تحديد الاستقرارية عن طريق الرسم الى خبرة وأحيانا الاشكال البيانية يصعب قراءتها وتحديد هل أن السلسلة مستقرة أو غير مستقرة وخصوصا إذا كان حجم البيانات كبيراً، لذلك يفضل استعمال الاختبارات الرقمية مثل اختبار Dickey Fuller للكشف عن استقرارية أو عدم استقرارية السلسلة الزمنية. وهو أحد اختبارات جذر الوحدة الأحادي. وعلى الرغم من أن هنالك الكثير من اختبارات جذر الوحدة الأحادي الا ان هذا الاختبار يعد من اهم الاختبارات وأكثرها شيوعا واستعمالا لأنه يكشف عن استقرارية السلسلة الزمنية لكل متغير من المتغيرات على حدة حتى لا يظهر لنا الانحدار الزائف. ويستند هذا الاختبار على المعادلة الآتية^[2].

$$\Delta Y_t = b_0 + b_1 T + \delta Y_{t-1} + e_t \quad (1-2)$$

إذ ان:

Y_t : متغير الاستجابة

Δ : تمثل الفرق الأول للسلسلة الزمنية Y_t .

δ : تمثل معلمة المتغير المتباطئ.

b_0 : المحور او المقطع الصادي.

T : تمثل الاتجاه الزمني.

وان معادلة الاتجاه الزمني تكتب بالشكل الآتي.

$$T = t - 1 - \frac{n}{2} \quad (2-2)$$

اذ ان:

n : تمثل حجم العينة

t : تمثل الزمن وان ($t=1,2,3,\dots,n$)

ولتجنب حصول الارتباط الذاتي في حد الخطأ العشوائي (e_t) وذلك بإضافة حداً ثابتاً وحسب طبيعة السلسلة الزمنية المدروسة. فتصبح المعادلة بالشكل الآتي.

$$\Delta Y_t = b_0 + b_1 T + \delta Y_{t-1} + a_i \sum_{i=1}^n \Delta Y_{t-i} + e_t \quad (3-2)$$

فيطلق على المعادلة (3-2) باختبار ديكي فولر الموسع (Augmented Dickey-Fuller) إذ تصبح (e_t) غير مرتبطة ذاتياً وتتميز بامتلاكها الخصائص المرغوبة أي تمتلك وسطاً حسابياً صفراً وتبايناً ثابتاً والمتغيرات العشوائية مستقلة وان فرضيات الاختبار تكتب بالشكل الآتي.

وجود جذر الوحدة في السلسلة الزمنية أي ان السلسلة غير مستقرة $H_0: \delta = 0$

لا تمتلك السلسلة الزمنية جذر وحدة وبالآتي فهي مستقرة $H_1: \delta \neq 0$

نستعمل احصاءه τ (tau) والتي تعرف باختبار DF وذلك عن طريق مقارنة قيمة احصاءه (τ) للمعلمة المقدرة (δ) مع القيم الجدولية لـ Dickey-Fuller. فاذا كانت القيمة المطلقة لاحصاءه (τ) المقدره اكبر من القيمة المطلقة لجدول اختبار (Dickey-Fuller) فعليه نرفض فرضية العدم H_0 وان السلسلة مستقرة. اما إذا كانت القيمة المطلقة لأحصاءه (τ) المقدره أصغر من القيمة المطلقة لجدول اختبار (Dickey-Fuller). فعليه لا نرفض فرضية العدم وان السلسلة غير مستقرة. او عن طريق قيمة (P-value) فاذا كانت قيمة P-value اصغر من مستوى المعنوية (0.05) فلا يوجد جذر وحدة ومن ثم نرفض فرضية العدم H_0 وان السلسلة الزمنية مستقرة.

اما اذا كانت قيمة P-value اكبر من مستوى المعنوية 0.05 فأنا لا نرفض فرضية العدم H_0 وان السلسلة تمتلك جذر وحدة وهي غير مستقرة.

وبعد تطبيق هذا الاختبار على السلسلة الزمنية فاذا كانت السلسلة الزمنية غير مستقرة وتحتوي على جذر الوحدة فنقوم بتحويلها الى سلسلة مستقرة.

وللتخلص من عدم الاستقرار في المتوسط يتم استعمال طريقة الفروق (method of differences) وهي أفضل طريقة لجعل السلسلة الزمنية مستقرة وذلك عن طريق التفاضل أي اخذ الفروق للسلسلة الزمنية الاصلية عن طريق طرح قيم (y_{t-1}) من قيم (y_t) وكذلك طرح قيم (y_{t-2}) من قيم (y_{t-1}) وهكذا الى ان نصل الى سلسلة زمنية جديدة ونقم باختبارها فاذا أصبحت مستقرة فان درجة التفاضل (d) او الفرق هو واحد أي تم اخذ فرق واحد للسلسلة الزمنية اما اذا كانت السلسلة غير مستقرة عند الفرق الأول (d) فنقوم بأخذ الفرق الثاني لها الى ان نحصل على سلسلة زمنية مستقرة وعادة ما تكون الفروق المأخوذة للسلسلة إما واحد او اثنان ويرمز لعامل الفروق بالرمز Δ .

$$\Delta y_t = y_t - y_{t-1} \quad (4-2)$$

$$\Delta = 1 - B$$

فيمكن كتابة المعادلة (4-2) بالشكل الآتي.

$$\Delta y_t = (1 - B)y_t \quad (5-2)$$

إذ ان:

B : تمثل عامل الارتداد الخلفي.

اما إذا كانت السلسلة الزمنية غير مستقرة في التباين فأننا نقوم بأخذ التحويل المناسب حتى تستقر السلسلة الزمنية فقد يكون التحويل هو بأخذ اللوغاريتم او مقلوب البيانات [2].

Linearity

(2-3-2) الخطية

من الضروري جدا التعرف الى خطية او لا خطية السلسلة الزمنية. حتى يسهل لنا التعامل معها والوصول الى نتائج صحيحة. فان التعامل مع السلسلة الزمنية دون المعرفة المسبقة لخطيتها يؤدي الى ان تقدير معالمها تكون خاطئة.

ويقصد بالخطية هو ان العلاقة بين المتغير المعتمد والمتغيرات المستقلة تكون خطية او غير خطية [12]. كما في المعادلات الآتية.

$$Y_t = B_0 + B_1 X_t + e_t \quad (6-2)$$

معادلة (6-2) مثال على أنموذج سلسلة زمنية خطية.

$$Y_t = \alpha e^{-bt} + e_t \quad (7-2)$$

معادلة (7-2) مثال على أنموذج سلسلة زمنية غير خطية.

وهناك العديد من الاختبارات التي تكشف عن خطية او لا خطية السلسلة الزمنية. ومنها اختبار (Mcleod and Li) واختبار Mann-Kendall واختبار (Keenan and Tsay) واختبار (Engle). وفي هذا البحث سنكتفي باختبارين فقط وكالاتي.

(1-2-3-2) اختبار (ML) Mcleod and Li

تم اقتراح إحصائية من نوع portmanteau (تستعمل احصاءه portmanteau لأختبار التأثيرات اللاخطية في بيانات السلاسل الزمنية) من قبل العالمان (Mcleod and Li) التي تستند الى دالة الارتباط الذاتي لمربعات البواقي إذ يتم الحصول عليها في انموذج (ARMA) الملائم. يعتمد هذا الاختبار على تطبيق إحصائية (Ljung-Box) على مربعات البواقي لأنموذج ARMA(p,q) للتأكد من عدم صلاحية الأنموذج. لذا فان إحصاء الاختبار هي [12]

$$Q(m) = n(n + 2) \sum_{k=1}^m \frac{(r_e^2(k))}{(n-k)} \quad (8-2)$$

وان حساب $r_e^2(k)$ يعتمد على البواقي المربعة كما في الصيغة الآتية.

$$r_e^2(k) = \frac{\sum_{t=k+1}^n e_t^2 e_{t-k}^2}{\sum_{t=1}^n e_t^2} \quad ; k=1, 2, \dots, n \quad (9-2)$$

إذ ان:

m : تمثل عدد الارتباطات الذاتية المستعملة في الاختبار.

n : تمثل حجم العينة.

r : الارتباط الذاتي.

إذ يتم حساب المعاملات للارتباطات الذاتية للبواقي المربعة (e_t^2) الناتج من الأنموذج المتعدد للبيانات. فإذا توزعت البواقي (e_t) بصورة مستقلة ومتطابقة (Identically independent distribution) فالتوزيع الذي يحاذي إحصاء الاختبار $Q(m)$ هو مربع كاي بدرجة حرية (m) [14]. أي ان.

$$Q(m) \sim \chi^2(m)$$

وان فرضيات الاختبار تكتب بالشكل الآتي.

H_0 : السلسلة الزمنية خطية

H_1 : السلسلة الزمنية غير خطية

فإذا كانت قيمة $Q(m)$ أكبر من القيمة الحرجة لجدول اختبار Mcleod and Li نرفض فرضية العدم وأن السلسلة غير خطية. أما إذا كانت قيمة $Q(m)$ أقل من القيمة الحرجة لجدول اختبار Mcleod and Li فلا نرفض الفرضية الصفرية وأن السلسلة الزمنية خطية.

كذلك يمكن تحديد القبول والرفض عن طريق قيمة P-value لاختبار Mcleod and Li فإذا كانت قيمة P-value أقل من مستوى المعنوية (0.05) فنرفض فرضية العدم وان السلسلة غير خطية. أما إذا كانت قيمة P-value أكبر من مستوى المعنوية (0.05) فلا نرفض فرضية العدم وان السلسلة الزمنية خطية [12]، [14].

Mann-Kendall Test

(2-2-3-2) اختبار مان كيندال

اختبار Mann-Kendall هو اختبار احصائي يتم تطبيقه بشكل واسع في تحليل الاتجاه للسلاسل الزمنية. ويستعمل في الكثير من الدراسات الوبائية والبيئية والمناخية. ويعبر عنه بالعلاقة الآتية.

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{var}(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sqrt{\text{var}(S)}} & \text{if } S < 0 \end{cases} \quad (10-2)$$

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n a_{ij} \quad (11-2)$$

$$a_{ij} = \text{sign}(x_j - x_i) = \text{sign}(R_j - R_i)$$

$$x_i < x_j \quad 1$$

$$x_i = x_j \quad 0$$

$$x_i > x_j \quad -1$$

$$\text{var}(S) = \frac{1}{18} [n(n-1)(2n+5) - \sum_{p=1}^q t_p(t_p-1)(2t_p+5)] \quad (12-2)$$

إذ ان:

n: حجم العينة

q: عدد المحاولات في المجموعة

t: عدد قيم البيانات في المجموعة

var(S): هو التباين المستعمل لاحصاءه الاختبار Z

(R_j - R_i): تمثل رتب المتغيرين

وان فرضيات الاختبار هي كالاتي.

H₀: There is no trend in the series. لا يوجد اتجاه للسلسلة الزمنية.

H₁: There is a trend in the series. هنالك اتجاه في السلسلة الزمنية.

فاذا كانت احصاءه Z المحسوبة أكبر من Z الجدولية ($Z_{\frac{\alpha}{2}}$) وبهذا ترفض الفرضية

الصفيرية H₀. يعني وجود اتجاه في السلسلة الزمنية وان السلسلة لا خطية، اما اذا كانت Z

المحسوبة أصغر من Z الجدولية ($Z_{\frac{\alpha}{2}}$) فعندئذ لا نرفض الفرضية الصفيرية H₀، وهذا يدل

على عدم وجود اتجاه في السلسلة الزمنية وأن السلسلة خطية، كذلك يمكن تحديد القبول

والرفض عن طريق قيمة P-value لجدول اختبار Mann-Kendall فاذا كانت قيمة P-

value أقل من مستوى المعنوية (0.05) فنرفض فرضية العدم وان السلسلة غير خطية، اما إذا كانت قيمة P-value أكبر من مستوى المعنوية (0.05) فلا نرفض فرضية العدم وان السلسلة الزمنية خطية [44]، [35].

Homogeneity

(3-3-2) التجانس

يستخدم اختبار التجانس لمعرفة ما إذا كانت مجموعة البيانات تتوزع التوزيع نفسه ام لا، أي ان الخصائص الإحصائية لأي جزء من مجموعة البيانات هي الخصائص نفسها في جزء آخر من البيانات نفسها، وتكون البيانات مستقلة وتتوزع بشكل عشوائي. هنالك أربع أنواع من اختبارات التجانس في السلاسل الزمنية، وهي اختبار بيتيت (Pettit test) واختبار بوشند (Buishand) واختبار التجانس الطبيعي القياسي (standard normal homogeneity) وكذلك اختبار فون نيومان (Von Neumann). وهنا سنكتفي باختبارين فقط [11].

standard normal

(1-3-3-2) اختبار التجانس الطبيعي القياسي

homogeneity

تم اقتراح هذا الاختبار من قبل العالم (Alexandersson 1986). وهو من الاختبارات المهمة والشائعة التي تستعمل بشكل واسع لاختبار تجانس البيانات. أن الفرضية الصفرية H_0 تفترض أن البيانات تتوزع توزيعاً طبيعياً ومتطابقاً وكذلك هي عبارة عن كميات مستقلة عشوائية، اما الفرضية البديلة H_1 فهي تفترض وجود تحولا وانكسارا يحصل في قيم معدل البيانات (Break in the mean). وان صيغة الاختبار تكتب بالشكل الآتي.

$$T(u) = u\bar{z}_1 + (n - u)\bar{z}_2 \quad (13-2)$$

$$\bar{z}_1 = \frac{1}{u} \sum_{i=1}^n \frac{u_i - \bar{u}}{s} \quad (14-2)$$

$$\bar{z}_2 = \frac{1}{n-u} \sum_{i=1}^n \frac{u_i - \bar{u}}{s} \quad (15-2)$$

إذ ان:

n: تمثل عدد السنوات الكلية.

u : تمثل عدد السنوات في الجزء الأول من البيانات.

$(n-u)$: تمثل عدد السنوات في الجزء الثاني من البيانات.

i : تمثل تسلسل السنوات.

s : الانحراف المعياري.

\bar{u} : معدل البيانات الكلية.

(z_1, z_2) : معدل القيم القياسية للجزء الأول والثاني على الترتيب.

في حالة وجود نقطة انكسار في سلسلة البيانات تحت الاختبار فإن قيمة T_0 يمكن حسابها كالآتي.

$$T_0 = \max T(u) \quad (16-2)$$

إذ يتم مقارنة قيمة T_0 مع القيمة الحرجة u_c وهي التي تعتمد على المدة الزمنية لمجموعة البيانات تحت الاختبار، فإذا كانت قيمة $T_0 > u_c$ فأنا نرفض الفرضية الصفرية H_0 أي ان البيانات غير متجانسة، اما اذا كانت قيمة $T_0 < u_c$ فأنا لا نرفض الفرضية الصفرية وان البيانات متجانسة. كذلك يمكن تحديد القبول والرفض عن طريق قيمة P-value لاختبار standard normal homogeneity فإذا كانت قيمة P-value اقل من مستوى المعنوية (0.05) فنرفض فرضية العدم وان السلسلة غير متجانسة، اما اذا كانت قيمة P-value أكبر من مستوى المعنوية (0.05) فلا نرفض فرضية العدم وان السلسلة الزمنية متجانسة [9].

Von Neumann test

(2-3-3-2) اختبار فون نيومان

هو نسبة مربع متوسط السلسلة (من سنة الى السنة اخرى) مقسوما على التباين للسلسلة نفسها. ولا تدل هذه النسبة الى السنة التي حدث فيها التغيير وانما تدل على وجود تغيير واضح وانكسار في السلسلة الزمنية. وهو من أكثر الاختبارات استعمالا في فحص تجانس السلسلة الزمنية [11]. ويمكن التعبير عن صيغة الاختبار بالمعادلة الآتية.

$$W = \sum_{i=1}^{n-1} (X_i - X_{i+1} + 1)^2 / \sum_{i=1}^n (X_i - \bar{X})^2 \quad (17-2)$$

إذ يتم مقارنة قيمة (W) مع القيمة الحرجة (Critical Value). فعندما تكون قيمة W أكبر من القيمة الحرجة فهذا يدل على ان النموذج تحت الاختبار يكون متجانساً، وإذا كانت قيمة W أصغر من القيمة الحرجة فهذا يدل على ان النموذج غير متجانس ويعاني من نقطة انكسار ولا يعطي هذا النموذج أي معلومات حول نقطة الانكسار. هنالك جداول خاصة للقيم الحرجة في المصادر المذكورة. كذلك يمكن تحديد القبول والرفض عن طريق قيمة P-value لجدول اختبار Von Neumann فاذا كانت قيمة P-value اقل من مستوى المعنوية (0.05) فنرفض فرضية العدم وان السلسلة غير متجانسة، اما اذا كانت قيمة P-value أكبر من مستوى المعنوية (0.05) فلا نرفض فرضية العدم وان السلسلة الزمنية متجانسة [17].

(4-2) أنموذجات السلاسل الزمنية اللاخطية

Nonlinear time series

models

سيتم في هذا الفرع التطرق الى أنموذجات الاتجاه العام اللاخطية إذ نلاحظ ان بعض البيانات او الحالات التي تتبع الاتجاه اللاخطي إذ لا يمكن استعمال الطرائق او الانموذجات الخطية في وصف تغيرات السلسلة الزمنية وكذلك لا يمكن معها إيجاد الاتجاه العام للسلسلة وهذا يتطلب أن نبحث عن أنموذجات أخرى تلائم هذه البيانات. هنالك الكثير من الانموذجات الخطية وتطبيقاتها الإحصائية والاقتصادية التي لاقت اهتماماً كبيراً في الدراسة والتطبيق لدى الكثير من الباحثين، بالمقابل لم تلقَ الانموذجات اللاخطية مكانتها العلمية كما في الانموذجات الخطية، أن العلاقات الخطية بين الظواهر المختلفة من العلاقات الشائعة الاستعمال، الا أن هنالك الكثير من الظواهر والعلاقات التي تأخذ الطابع اللاخطي وذلك ينبع من

طبيعة العلاقة نفسها وتأثير المتغيرات فيما بينها. وهنا قد يقع الباحث في خطأ عند مرحلة التقدير وتجاوز العلاقات اللاخطية الى العلاقات الخطية وينتج عن هذا أخطاء كبيرة في عملية التقدير ودرجة المعنوية. وان توقعاته تصبح متحيزة بشكل كبير الى الأعلى او الى الأسفل. فعند حصولنا على تقديرات لا تناسب متغير الاستجابة (Y) قياسا الى المتغير او المتغيرات المستقلة (X) بالمقارنة مع القيم الحقيقية. كذلك في حال وجدنا ان معامل الارتباط منخفضاً وكذلك الخطأ المعياري للتقدير كبير. فهذا دليل على ان جودة توفيقنا للأنموذج غير صحيحة. وهذا ناتج عن الزيادة الكبيرة في تشتت القيم حول وسطها الحسابي. وكذلك ناتج عن انخفاض القدرة التفسيرية للمتغيرات المستقلة (X) وذلك بسبب عدم صحة توفيق الأنموذج.

هنالك الكثير من أنموذجات السلاسل الزمنية اللاخطية ومنها الأسية والتربيعية واللوغاريتمية و..... الخ. ومنها نماذج لا خطية غير قابلة للتحويل أي لا يمكن تحويلها الى صيغ خطية ومنها أيضا قابلة للتحويل أي بالإمكان تحويلها الى صيغ خطية وهي الأكثر شيوعا واستعمالا لأنها تعطي نتائج دقيقة [5]. وفي هذه الرسالة تم استعمال خمس أنموذجات لا خطية قابلة للتحويل الى خطية وهي: (Exponential model, Von Bertalanffy model, Gompertz model, Logistic model and Chapman-Richard model)

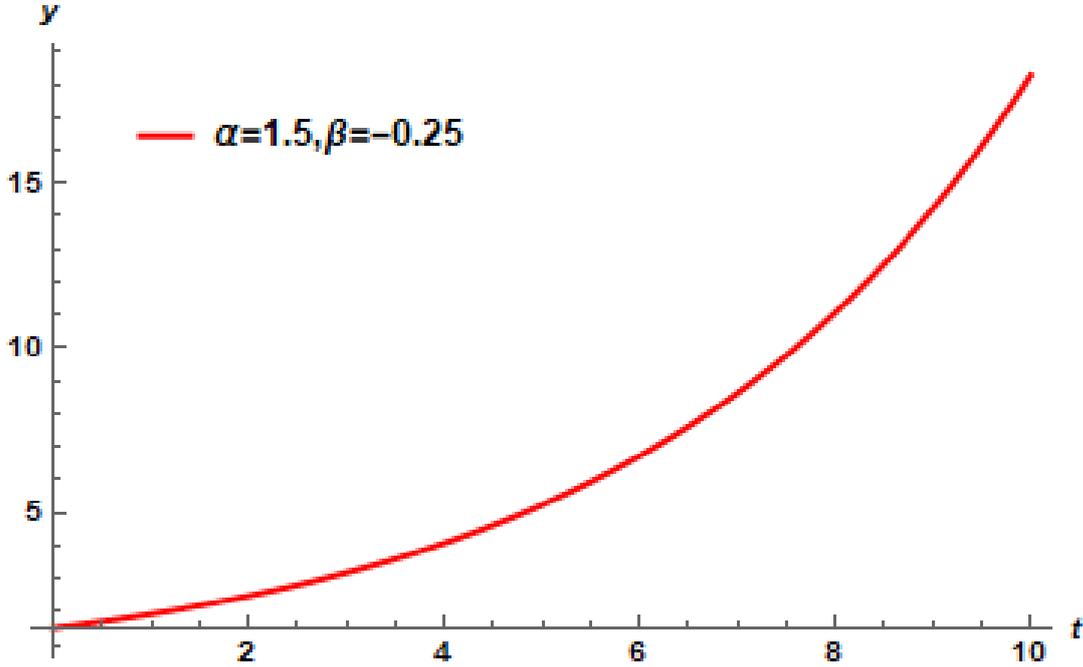
Exponential Model

(1-5-2) الأنموذج الأسّي

هو أحد اهم الأنموذجات اللاخطية في السلاسل الزمنية. وهو أنموذج او تعبير رياضي يصف عملية التزايد او التناقص (مثل تزايد أعداد الإصابات بفيروس كوفيد-19). ويطلق عليه أيضا أسم أنموذج النمو الطبيعي. أن معدل النمو الأسّي في حد ذاته مقياسا مهما لشدة الوباء [31][29]. وبشكل عام يحتوي هذه الأنموذج على معلمتين فقط وهما (α, β) . والشكل (2-3) يمثل منحنى النمو الأسّي

.Exponential Growth Curve

Exponential model



الشكل (3-2) يمثل منحنى النمو الأسي Exponential Growth Curve

ويعبر عن الأنموذج الأسي بالمعادلة الآتية.

$$Y_t = \alpha e^{-\beta t} + e_t \quad (18-2)$$

إذ ان:

Y_t : تمثل عدد الافراد المصابين في الزمن t

(α, β) : تمثل معاملات الأنموذج

t : الزمن

e_t : الخطأ العشوائي

إذ أن قيمة متغير الاستجابة Y_t تتزايد بفترات زمنية متساوية وبمعدل الزيادة نفسه. أن قيمة متغير الاستجابة Y_t (الإصابات) تزداد مع مرور الزمن (بزيادة أسية) أو في حال النقصان فأنها تتغير مع مرور الزمن (بنقصان أسي). أن معظم الأوبئة ومنها وباء فيروس كوفيد-19 تنمو أضعافا مضاعفة خلال الفترات الأولى من الوباء [36].

تشير نظريات النمو السكاني على أن السكان ذوي الموارد الطبيعية غير المحدودة ينمون بسرعة كبيرة وبعد ذلك يتناقص النمو السكاني وذلك بسبب ازدياد عدد السكان وقلة الموارد الطبيعية، ويسمى هذا النمو المتسارع لزيادة حجم السكان بالنمو الأسي. أي أن النمو السكاني الأسي يظهر عندما تكون الموارد كثيرة وغير محدودة وأن المنحنى الأسي يظهر على شكل حرف J، أما في حال كانت الموارد محدودة وقليلة فأن النمو السكاني ينخفض لان الموارد نادرة فيظهر هنا المنحنى اللوجستي ويكون على شكل حرف S (وهذا هو الفرق بين الأنموذج الأسي واللوجستي).

وأن هذا ينطبق على أعداد الإصابات بفيروس كوفيد-19 إذ أن الارتفاع في أعداد الإصابات بسبب العدوى يمثل منحنى نمو أسي، أما في حال وصول الإصابات الى حد معين وفجأة تقل العدوى ومن ثم تقل أعداد الإصابات بالفيروس هنا يدعى بالمنحنى اللوجستي [29].

وهناك عدة طرائق لتقدير الأنموذجات اللاخطية واحد هذه الطرائق هي طريقة الإمكان الأعظم (Maximum Likelihood) وسيتم الاعتماد عليها في هذا البحث لتقدير معالم الأنموذجات اللاخطية قيد الدراسة. وتتمثل هذه الطريقة بالخطوات الآتية.

1- نقوم بتعظيم دالة الأنموذج فيصبح الأنموذج بالشكل الآتي.

$$L(\hat{Y}_t) = \alpha^n e^{-\beta \sum t} \quad (19-2)$$

2- نأخذ Ln للمعادلة (19-2) فتصبح بالشكل الآتي.

$$\ln[L(\hat{Y}_t)] = n \ln(\alpha) - \beta \sum t \quad (20-2)$$

$$\text{Let: } \ln[L(Y_t)] = \hat{Y}_t^* , \quad \ln(\alpha) = \alpha^*$$

نعوض الفرضيات بالمعادلة (20-2) فتصبح بالشكل الآتي.

$$Y_t^* = \alpha^* - \beta \sum t \quad (21-2)$$

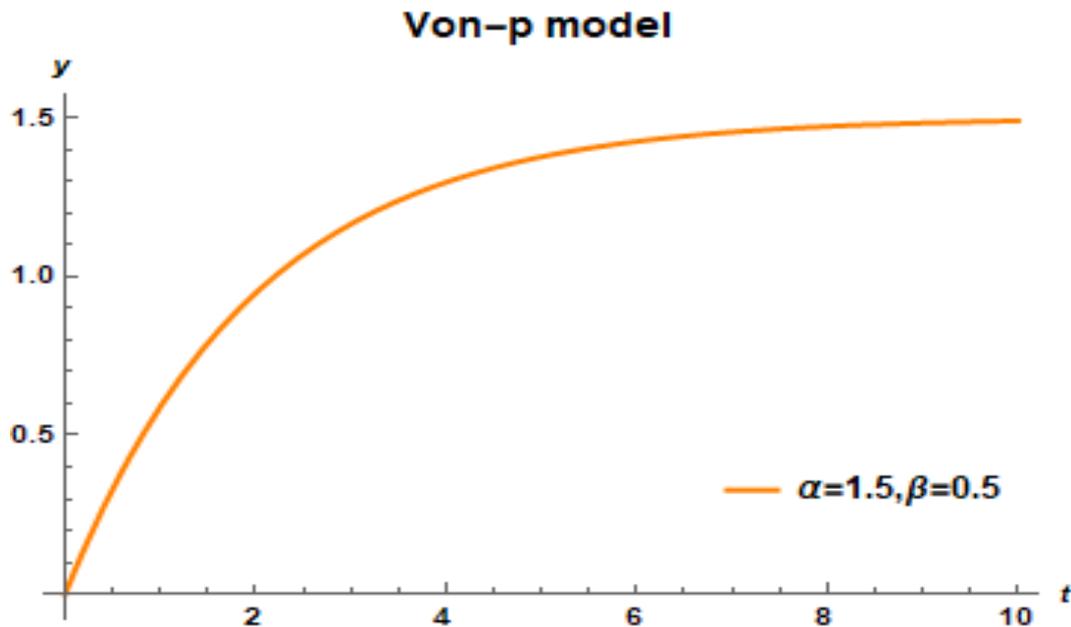
3- نشق دالة الأنموذج في المعادلة (2-21) بالنسبة للمعلمت (α, β) ونساوي المشتقة بالصفر لنحصل على مجموعة من المعادلات.

وبما أن هذه المعادلات هي معادلات لا خطية يصعب تقديرها بالطرائق الاعتيادية لذلك نلجأ الى استعمال أحد الطرائق التكرارية لتقدير هذه المعلمت كطريقة نيوتن رافسون (Newton-Raphson) التكرارية والتي تم الاعتماد عليها في هذه الرسالة. وهي أحد الطرائق العددية التي تستعمل لتقدير معلمت دالة الإمكان الاعظم. وهي تعد من أهم وأكفأ الطرائق التكرارية وأكثرها شيوعا واستعمالا.

(2-5-2) أنموذج فون بيرتلانفي Von Bertalanffy model

هو أحد أنموذجات السلاسل الزمنية اللاخطية. يتكون انموذج Von Bertalanffy من معلمتين فقط هما (α, β) وغالبا ما يستعمل أنموذج Von Bertalanffy كأنموذج للنمو. ويتم استعماله بشكل أساس لدراسة العوامل التي تتحكم في النمو وتؤثر عليه. تم اختيار هذا الأنموذج لوصف قانون انتشار الامراض المعدية ودراسة العوامل التي تتحكم في انتشار فيروس Covid-19 وتؤثر عليه [28]، [34].

والشكل (4-2) يمثل منحنى نمو Von Bertalanffy Growth Curve.



الشكل (4-2) يمثل منحنى نمو فون بيرتلانفي Von Bertalanffy Growth Curve.

ويعبر عن الأنموذج بالمعادلة الآتية.

$$Y_t = \alpha(1 - e^{-\beta t}) + e_t \quad (22-2)$$

إذ ان:

Y_t : متغير الاستجابة

(α, β) : تمثل معاملات الأنموذج

t : الزمن

e_t : الخطأ العشوائي

ولتقدير معاملات انموذج فون بيرتلانفي Von Bertalanffy نتبع نفس الخطوات في الأنموذج السابق وكالاتي.

1- نقوم بتعظيم دالة الأنموذج فيصبح بالشكل الآتي.

$$L(\hat{Y}_t) = \alpha^n \pi (1 - e^{-\beta t}) \quad (23-2)$$

2- نأخذ Ln للمعادلة (23-2) فتصبح بالشكل الآتي.

$$\ln[L(\hat{Y}_t)] = n \ln(\alpha) + \sum \ln(1 - e^{-\beta t}) \quad (24-2)$$

$$\text{Let: } \ln[L(\hat{Y}_t)] = \hat{Y}_t^* \quad , \quad \ln(\alpha) = \alpha^*$$

نعوض الفرضيات بالمعادلة (24-2) فتصبح بالشكل الآتي.

$$\hat{Y}_t^* = n\alpha^* + \sum \ln(1 - e^{-\beta t}) \quad (25-2)$$

3- نشق دالة الأنموذج في المعادلة (25-2) بالنسبة للمعاملات (α, β)

ونساوي المشتقة بالصفر لنحصل على مجموعة من المعادلات.

وبما أن هذه المعادلات هي معادلات لا خطية يصعب تقديرها بالطرائق الاعتيادية لذلك نلجأ الى استعمال احدى الطرائق التكرارية لتقدير هذه المعاملات كطريقة نيوتن رافسون (Newton-Raphson) التكرارية والتي تم الاعتماد عليها في هذه الرسالة. وهي احدى الطرائق العددية التي تستعمل لتقدير معاملات

دالة الإمكان الأعظم. وهي تعد من أهم وأكفأ الطرائق التكرارية وأكثرها شيوعاً واستعمالاً.

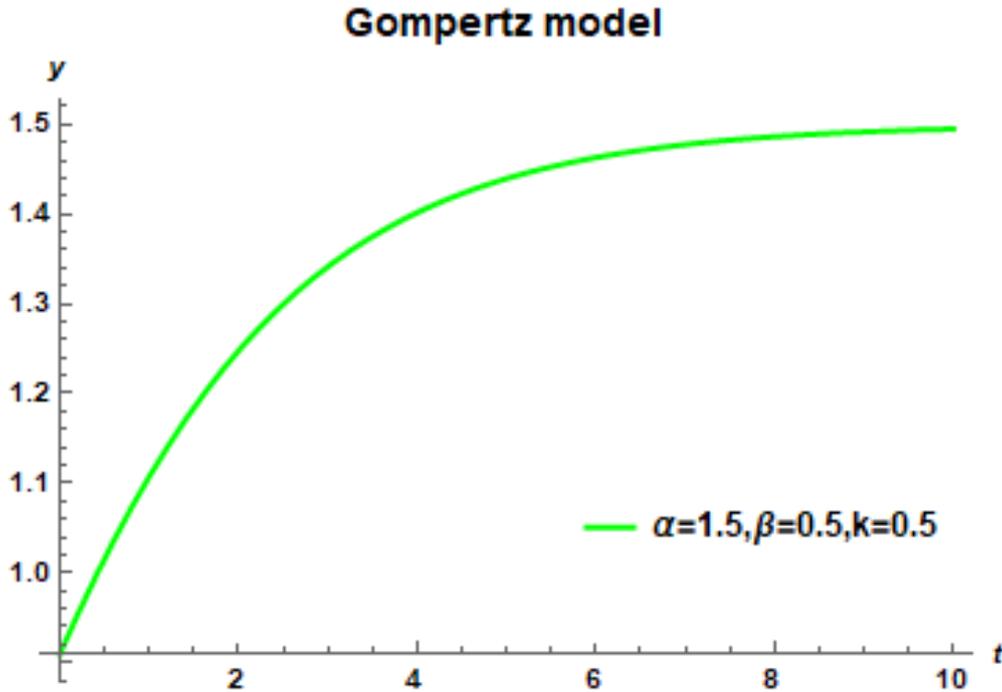
(3-5-2) أنموذج جمبرتز Gompertz Model

هو أحد أنموذجات النمو اللاخطية في السلاسل الزمنية. إذ تم صياغة هذا الأنموذج من قبل العالم (Benjamin Gompertz) عام 1825 ليتناسب مع جداول الوفيات، إذ تم استعماله كأنموذج للنمو ولا سيما في الدراسات الوبائية والطبية والحيوية [27]، وفي الجانب الاقتصادي يستخدم للتنبؤ بتطور السوق والأوراق المالية، وكذلك يستخدم على نطاق واسع في علم الأحياء، ويستعمل لوصف نمو الحيوانات والنباتات وكذلك يصف أعداد واحجام الخلايا البكتيرية والسرطانية [37]، كذلك أستعمل هذا الأنموذج لوصف قانون انتشار الامراض المعدية ودراسة العوامل التي تتحكم في انتشار فيروس كوفيد-19. ويعد منحنى جمبرتز مفيداً جداً للعديد من الحالات التجريبية لأنه يفترض وجود نمط نمو غير متماثل عكس منحنى النمو اللوجستي الذي يفترض وجود نمط نمو متماثل. يمتاز إنموذج نمو جمبرتز القياسي بثلاث معلمات وهي أحد الأسباب الواضحة في تطبيقه في الكثير من الممارسات والدراسات، وفي بعض الأحيان يحتوي على أكثر من ثلاث معلمات وحسب عدد المتغيرات المستقلة [24]، تم تطوير الانموذجات التي حاولت تسليط الضوء على الزيادة في معدلات الوفيات مع التقدم في العمر وان أحد هذه الانموذجات هو انموذج جمبرتز Gompertz model الذي يفترض ان معدلات الوفيات يزداد اضعافاً مضاعفة مع التقدم في العمر، لقد اثبت انموذج منحنى جمبرتز شعبيته إذ يقدم وصفاً عاماً ودقيقاً للعديد من أنماط او ظواهر الوفيات المرصودة. ان منحنى او دالة جمبرتز هي دالة سينية (Sigmoid Function) وهي أحد أنواع الانموذجات الرياضية للسلاسل الزمنية. إذ يكون انموذج جمبرتز بطيئاً في بداية ونهاية المدة الزمنية. وان المنحنى الأيمن للدالة يقترب بشكل أكبر من المنحنى الأيسر او السفلي للدالة. على عكس الدالة اللوجستية تماماً إذ ان المتغيرين يقتربان من المنحنى بشكل متناظر. ان انموذج جمبرتز هو حالة خاصة من انموذج Richards ذو أربع معلمات وهو ينتمي الى عائلة Richards من انموذجات نمو Sigmoida من ثلاث معلمات. جنباً الى جنب مع الانموذجات الأسية والأنموذج اللوجستي وانموذج فون بيرتلانفي. وان الصيغة العامة لأنموذج جمبرتز تتمثل بالمعادلة الآتية [33].

$$Y_t = \alpha e^{-\beta e^{-kt}} + e_t \quad (26-2)$$

إذ ان (α, β, k) تمثل معاملات النموذج وان قيمة متغير الاستجابة (Y) تتغير مع الزمن (t).

والشكل الآتي يمثل منحنى نمو جمبرتز Gompertz Growth Curve.



الشكل (5-2) يمثل منحنى نمو جمبرتز Gompertz Curve Model

ولتقدير معاملات النموذج جمبرتز Gompertz نتبع نفس الخطوات في النموذج السابق وكالاتي.

1- نقوم بتعظيم دالة النموذج فيصبح النموذج بالشكل الآتي.

$$L(\hat{Y}_t) = \alpha^n e^{-\beta^n e^{-k \sum t}} \quad (27-2)$$

2- نأخذ Ln للمعادلة (27-2) فتصبح بالشكل الآتي.

$$\ln[L(\hat{Y}_t)] = n \ln(\alpha) - \beta^n e^{-k \sum t} \quad (28-2)$$

نلاحظ أن معادلة النموذج ما زالت غير خطية فنأخذ Ln مرة ثانية فتصبح بالشكل الآتي.

$$\ln[\ln[L(\hat{Y}_t)]] = \ln[n \ln(\alpha)] - n \ln(\beta) - k \sum t \quad (29-2)$$

$$\text{Let: } \ln \left[\ln \left[L(\hat{Y}_t) \right] \right] = \hat{Y}_t^* \quad , \quad \ln(\alpha) = \alpha^* \quad , \quad \ln(\beta) = \beta^*$$

نعوض الفرضيات بالمعادلة (29-2) فتصبح بالشكل الآتي.

$$\hat{Y}_t^* = n \ln(\alpha^*) - n\beta^* - k \sum t \quad (30-2)$$

3- نشق دالة الأنموذج في المعادلة (30-2) بالنسبة للمعلمت (α, β, k) ونساوي المشتقة بالصفر لنحصل على مجموعة من المعادلات.

وبما أن هذه المعادلات هي معادلات لا خطية والتي يصعب تقديرها بالطرائق الاعتيادية لذلك نلجأ الى استعمال احدى الطرائق التكرارية لتقدير هذه المعلمت كطريقة نيوتن رافسون (Newton-Raphson) التكرارية والتي تم الاعتماد عليها في هذه الرسالة. وهي احدى الطرائق العددية التي تستعمل لتقدير معلمت دالة الإمكان الاعظم. وهي تعد من أهم وأكفأ الطرائق التكرارية وأكثرها شيوعا واستعمالا.

Logistic Model

(4-5-2) الأنموذج اللوجستي

هو أحد الأنموذجات اللاخطية في السلاسل الزمنية. يتكون هذا الأنموذج من ثلاث معلمت هي (α, β, k) . يستعمل الأنموذج اللوجستي بشكل رئيس في علم الأوبئة [28]. تم استعمال هذا الأنموذج لأول مره من قبل العالم البلجيكي Verhulst عام 1845. واستخدم منحنيات النمو اللوجستي لنمذجة نمو السكان. استند العالم Verhulst الى اعمال العالم Malthus الذي اعتقد ان سكان البلد او منطقة معينه سيكون قادراً على الوصول الى حد او مستوى معين من النمو. واخذ العالم Verhulst هذه الفكرة وأضاف عوامل أخرى الى المعادلة الأسية ليكون الأنموذج اللوجستي. ويطلق على المعادلة اللوجستية بمصطلح القدرة الاستيعابية. وهي الحد الأقصى للحجم الذي يمكن ينمو اليه السكان. وعند هذه النقطة يستقر حجم السكان ويزداد معدل النمو [23]. على الرغم من ان منحنى نمو جمبرتز ومنحنى النمو اللوجستي يصفان سلوكا مشابها في بعض مراحل هذا التطور الا ان أحد اهم الاختلافات هو ان عملية نمو جمبرتز تكون غير متماثلة في حين ان المنحنى اللوجستي هو عملية متماثلة. لذلك فإن استعمال منحنى نمو غير مناسب يمكن ان يكون له تأثير كبير على التنبؤ [39]. ويعبر عن الأنموذج اللوجستي بالصيغة الآتية.

$$Y_t = \frac{\alpha}{1+\beta e^{-kt}} + e_t \quad (31-2)$$

إذ ان:

Y_t : متغير الاستجابة

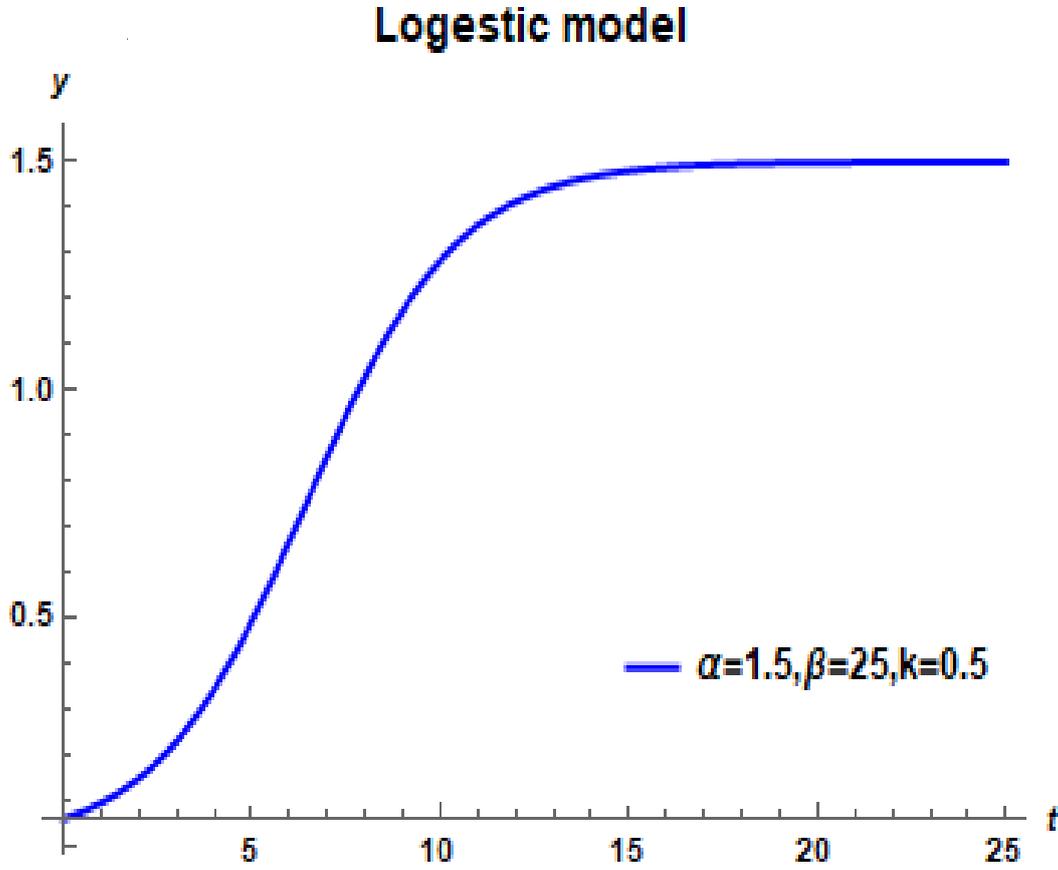
t : الزمن

(α, β, k) تمثل معلمات الأنموذج.

e_t : الخطأ العشوائي

ان قيمة متغير الاستجابة Y_t يتغير مع الزمن t (او متغيرا أخرى مستقلة). ان اهم ما يميز الأنموذج اللوجستي البسيط هو انه متمائل حول نقطة الانعطاف. وتشير هذه الميزة الى ان العملية التي ستحدث بعد نقطة الانعطاف هي صورة طبق الأصل للعملية التي حدثت بعدها [38]. وفي عام 1920 أعاد العالمان (Lowell Reed and Raymond Pearl) اكتشاف المعادلة اللوجستية. وقد لوحظ ان منحنى النمو اللوجستي يقدم وصفا لنمو المجتمعات السكانية التي لديها طفرات نمو مفاجئة. من مستوى ثابت ثم تصل الى حد معين ثم تتخفف ببطء للوصول الى معدل نمو منخفض مرة أخرى. ان الأنموذج اللوجستي Logistic model والانموذجات الأخرى تعد من الانموذجات الشائعة والكثيرة الاستعمال بين العلماء والتطبيقيين والباحثين. لما تمتلكه هذه الانموذجات من خصائص وصفات مناسبة ومهمة جدا بالنسبة الى نجاح التحليل للسلاسل الزمنية بناء على منحنيات النمو. كذلك طرائق اختبار جودة الانموذجات تكون ملائمة لنماذج منحنى النمو [24].

والشكل الآتي يمثل منحنى النمو اللوجستي Logistic Growth Model.



الشكل (6-2) يبين منحنى النمو اللوجستي Logistic Growth Curve

ولتقدير معالم النموذج اللوجستي Logistic model نتبع نفس الخطوات في النموذج السابق وكالاتي.

1- نقوم بتعظيم دالة النموذج فيصبح النموذج بالشكل الآتي.

$$L(\hat{Y}_t) = \frac{\alpha^n}{1 + \beta^n e^{-k \Sigma t}} \quad (32-2)$$

2- نأخذ Ln للمعادلة (32-2) فتصبح بالشكل الآتي.

$$\ln[L(\hat{Y}_t)] = n \ln(\alpha) - \ln[1 + \beta^n e^{-k \Sigma t}] \quad (33-2)$$

Let: $\hat{Y}_t^* = \ln[L(Y_t)]$, $\ln(\alpha) = \alpha^*$

نعوض الفرضيات بالمعادلة (33-2) فتصبح بالشكل الآتي.

$$\hat{Y}_t^* = n\alpha^* - Ln[1 + \beta^n e^{-k\Sigma t}] \quad (34-2)$$

3- نشق دالة النموذج في المعادلة (34-2) بالنسبة للمعلمات (α, β, k) ونساوي المشتقة بالصفر لنحصل على مجموعة من المعادلات.

وبما أن هذه المعادلات هي معادلات لا خطية والتي يصعب تقديرها بالطرائق الاعتيادية لذلك نلجأ الى استعمال أحد الطرائق التكرارية لتقدير هذه المعلمات كطريقة نيوتن رافسون (Newton-Raphson) التكرارية والتي تم الاعتماد عليها في هذه الرسالة. وهي أحد الطرائق العددية التي تستعمل لتقدير معلمات دالة الإمكان الاعظم. وهي تعد من أهم وأكفأ الطرائق التكرارية وأكثرها شيوعاً واستعمالاً.

(5-5-2) نموذج جابمان-ريتشارد Chapman-Richard model

هو أحد أنموذجات السلاسل الزمنية اللاخطية. يحتوي أنموذج Chapman-Richard على أربع معلمات هي $(\alpha, \beta, k, \theta)$. أن أنموذج Chapman-Richard هو تعميم لبقية الانموذجات اللاخطية مثل انموذج جمبرتز والأنموذج اللوجستي وانموذج فون بيرتلانفي. وان وضيفة هذه الانموذجات هي تقديم وصف دقيق وأكثر واقعية لمختلف الظواهر [19]. تم استعمال انموذج Chapman-Richard في العديد من الدراسات مثل دراسة العوامل التي تتحكم في نمو الحيوانات مثل نمو الأسماك والاعنام والابقار والخيول وغيرها. تم اختيار هذا الأنموذج لوصف قانون انتشار الامراض المعدية ودراسة العوامل التي تتحكم في انتشار فيروس covid-19 وتوثر فيه [41]. ويعبر عن هذا الأنموذج بالصيغة الآتية.

$$Y_t = \alpha(1 - \beta e^{-kt})^\theta + e_t \quad (35-2)$$

إذ ان:

Y_t : متغير الاستجابة

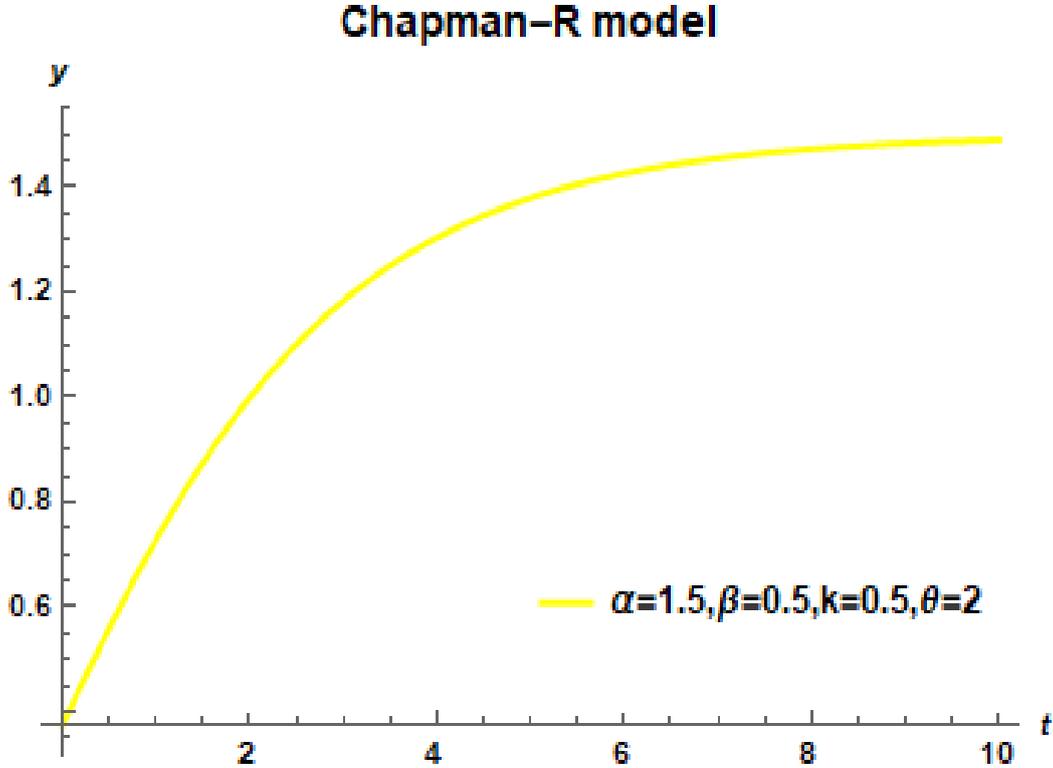
$(\alpha, \beta, k, \theta)$: تمثل معلمات الأنموذج.

t : الزمن

e_t : الخطأ العشوائي

والشكل (7-2) يمثل منحنى نمو جابمان ريتشارد Chapman-Ricard Growth

Curve



الشكل (7-2) يمثل منحنى نمو جابمان ريتشارد Chapman-Richard Growth Curve

ولتقدير معلمات أنموذج جابمان ريتشارد Chapman Richrad نتبع نفس الخطوات في الأنموذج السابق وكالاتي.

1- نقوم بتعظيم دالة الأنموذج فيصبح الأنموذج بالشكل الآتي.

$$L(\hat{Y}_t) = \alpha^n \pi (1 - \beta e^{-kt})^\theta \quad (36-2)$$

2- نأخذ Ln للمعادلة (36-2) فتصبح بالشكل الآتي.

$$\ln[L(\hat{Y}_t)] = n \ln(\alpha) + \theta \sum [\ln(1 - \beta e^{-kt})] \quad (37-2)$$

Let: $\ln[L(Y_t)] = \hat{Y}_t^*$, $\ln(\alpha) = \alpha^*$

نعوض الفرضيات بالمعادلة (37-2) فتصبح بالشكل الآتي

$$\hat{Y}_t^* = n\alpha^* + \theta \sum [Ln(1 - \beta e^{-kt})] \quad (38-2)$$

3- نشق دالة الأنموذج في المعادلة (58-2) بالنسبة للمعاملات $(\alpha, \beta, k, \theta)$ ونساوي المشتقة بالصفر لنحصل على مجموعة من المعادلات.

وبما أن هذه المعادلات هي معادلات لا خطية والتي يصعب تقديرها بالطرائق الاعتيادية لذلك نلجأ الى استعمال أحد الطرائق التكرارية لتقدير هذه المعاملات كطريقة نيوتن رافسون (Newton-Raphson) التكرارية والتي تم الاعتماد عليها في هذه الرسالة. وهي أحد الطرائق العددية التي تستعمل لتقدير معاملات دالة الإمكان الاكظم. وهي تعد من أهم وأكفأ الطرائق التكرارية وأكثرها شيوعاً واستعمالاً.

(5-2) منهجية بوكس جينكز (Box-Jenkins) لتحليل السلاسل الزمنية

ان عملية بناء نماذج السلاسل الزمنية وتحليلها لتمثيل بيانات تلك السلسلة تسمى بمنهجية بوكس-جينكز (Box-Jenkins). وهي واحدة من اهم الطرائق وأكثرها استعمالاً للتنبؤ بمختلف أنواع السلاسل (مستقرة وغير مستقرة، خطية ولا خطية) وتتمثل بالخطوات الآتية.

Diagnostic stage

(1-5-2) مرحلة التشخيص

ان من اهم مراحل تحليل السلاسل الزمنية هي مرحلة التشخيص. وتستند هذه المرحلة الى مجموعة البيانات قيد الدراسة. إذ نقوم برسم بيانات السلسلة الزمنية. لاكتشاف فيما إذا كانت هذه السلسلة تمتلك اتجاه عام او تحتوي على التأثيرات الموسمية او الدورية. وبيان صفات او خصائص السلسلة الزمنية (الاستقرارية والخطية والتجانس). وتم شرح هذه الخصائص واختباراتها بشكل مفصل في هذا الفصل في المبحث الفرعي (3-2).

Appreciation stage

(2-5-2) مرحلة التقدير

في هذه المرحلة يتم تقدير معلمات الانموذجات اللاخطية. وهناك عدة طرائق لتقديرها ومنها المربعات الصغرى الجزئية (Partial least squares) والمربعات الصغرى الموزونة (Weighted Least Squares) والامكان الأعظم (Maximum Likelihood) وغيرها من الطرائق الاخرى. وفي هذا البحث سنكتفي بطريقة الإمكان الأعظم فقط. إذ تم التقدير بهذه الطريقة بشكل مفصل في فرع الانموذجات اللاخطية.

The stage of choosing the أفضل (3-5-2) مرحلة اختيار الأنموذج الأفضل

Best model

بعد المرحلة السابقة والتعرف على خصائص السلسلة الزمنية والنموذجات التي تلائم بياناتها وكذلك تقدير معلمات هذه الانموذجات. تأتي مرحلة اختيار الأنموذج الأفضل من بين الانموذجات اللاخطية قيد الدراسة. وهناك عدة معايير لاختيار الأنموذج الأفضل ومنها معيار معلومات أكايكي (Akanke's Information Criterion) ومعيار حنان - كوين (Hannan-Quinn Criterion) ومعيار معلومات بيز (Bayesian Information Criterion). ان هذه المعايير تستعمل باتجاهين الأول هو للمقارنة بين الأنموذج نفسه ولكن برتب مختلفة لتحديد الرتبة المناسبة للأنموذج إذ تبرز المشكلة في تقدير رتبة الأنموذج فاذا اختيرت الرتبة إذ تكون اقل من الرتبة الفعلية للأنموذج فهذا يؤدي الى ان معلمات الأنموذج تكون غير متنسقة (Non-consistent) وإذا اختيرت الرتبة حيث تكون اعلى من رتبة الأنموذج فأن التباين للأنموذج سوف يزداد وبهذا فان الأنموذج يفقد صحته والسبب هو الزيادة في عدد معلمات الأنموذج. اما الاتجاه الثاني هو للمقارنة بين الانموذجات اللاخطية واختيار أفضل انموذج يعطي أصغر قيمة لهذه المعايير، وتستند هذه المعايير على الإحصاءات التي تخص البواقي Residuals والتي تنتج من مطابقة الأنموذج (Fitted model) ويكون الأنموذج بصورة عامة غير متحيز (Unbiased) [15].

Akanke's Information Criterion (1-2-5-2) معيار معلومات اكاكي

يعد معيار معلومات اكاكي (Akanke's Information Criterion) أحد الأدوات القياسية لنمذجة بيانات السلاسل الزمنية. لقياس ملائمة الأنموذج الاحصائي ب M من المعلمات. ويرمز له بالرمز (AIC). ويعبر عنه بالمعادلة الآتية.

$$AIC(M) = -2\ln[\text{Maximum Likelihood}] + 2M \quad (39-2)$$

OR

$$AIC(M) = -2\ln \frac{RSS}{n} + 2M \quad (40-2)$$

إذ ان:

RSS: تمثل مجموع مربعات الانحدار

M: تمثل عدد معلمات الأنموذج.

n: تمثل عدد المشاهدات.

وان دالة Log Likelihood هي كالآتي.

$$\ln L = -\frac{n}{2} \ln(2\pi\hat{\sigma}_a^2) - \frac{1}{2\hat{\sigma}_a^2} S(\underline{\phi}) \quad (41-2)$$

إذ ان:

$$S(\underline{\phi}) = \sum_{t=-P}^n [E(a_t/\underline{\phi}. Z)]^2 \quad (42-2)$$

وعند تعظيم المعادلة (42-2) نحصل على الصيغة التقريبية لمعيار أكاكي (AIC).

$$AIC \approx n \ln \hat{\sigma}_a^2 + n(1 + \ln(2\pi)) + 2P \quad (43-2)$$

ان الحد الثاني للمعادلة (43-2) يكون ثابتا (Constant) فيصبح معيار أكاكي AIC بالشكل الآتي.

$$AIC(M) = n \ln \hat{\sigma}_a^2 + 2M \quad (44-2)$$

فأن الرتبة المثالية للأنموذج يتم اختيارها عن طريق القيمة M والتي تقابل أقل قيمة للمعيار. وان الأنموذج الذي يعطي أقل قيمة لمعيار AIC هو الأنموذج الأفضل. يمثل الحد الأول للصيغة (2-44) مدى المطابقة وذلك بأخذ أقل تباين له. ويمثل الحد الثاني عدد المعلومات للأنموذج المطابق.

ويمكن كتابة معيار AIC بالصيغة المعيارية وذلك عن طريق قسمة معيار AIC على حجم العينة n ويرمز له بالرمز NAIC. ويعبر عنه بالصيغة الآتية.

$$NAIC(P) = AIC(p) / n \quad (45-2)$$

وان الأنموذج الذي يقابل أقل قيمة لمعيار AIC هو الأنموذج الأفضل.

في عام 1989 توصل الباحثان Hurvich و Tsai الى معياراً جديداً يدعى بمعيار معلومات اكاكي المصحح (Corrected Akaike Information Criterion) ويرمز له بالرمز (AIC_c) ويعبر عنه بالصيغة الآتية.

$$AIC_c = n \ln \hat{\sigma}_a^2 + n \frac{1+P/n}{1-(P+2)/n} \quad (46-2)$$

وان الفائدة من استعمال معيار AIC_c هو لتصحيح التحيز الذي يظهره المعيار السابق [8]، [15].

Hannan-Quinn Criterion (2-2-5-2) معيار حنان – كوين

اقترح الباحثان (Hannan-Quinn) عام 1979 معياراً جديداً يدعى معيار حنان – كوين (Hannan-Quinn) ويرمز له بالرمز H-Q ويستعمل لتحديد الأنموذج الأفضل. ويعبر عنه بالصيغة الآتية [15].

$$H_Q(m) = \ln(\hat{\sigma}_a^2) + \frac{2mc \ln(Lnn)}{n} ; c > 2 \quad (47-2)$$

OR

$$H_Q(m) = n \ln \frac{RSS}{n} + 2m \ln[\ln(n)] \quad (48-2)$$

إذ ان:

RSS: تمثل مجموع مربعات الانحدار

m: تمثل عدد المعلمات في الأنموذج.

n: تمثل حجم العينة.

وان الأنموذج الذي يقابل اقل قيمة لمعيار (H-Q) هو الأنموذج الأفضل.

Bayesian Information criterion معيار معلومات بيز (3-2-5-2)

وهو أحد معايير تحديد رتبة الأنموذج ويرمز له بالرمز (BIC) ان معيار معلومات بيز يعطي مقدر متسق (Consistence Estimate) للرتبة الحقيقية على عكس معيار اكاكي AIC الذي يعطي غالبا مقدر ذو رتبة اعلى من الرتبة الحقيقية. ويأخذ الصيغة الآتية [15][8].

$$BIC(m) = n \ln(\hat{\sigma}_a^2) - (n - m) \ln \left(1 - \frac{m}{n}\right) m \ln n + m \ln \left[\frac{1}{m} \left(\frac{\sigma_z^2}{\sigma_a^2} - 1\right)\right]$$

(49-2)

OR

$$BIC(m) = n \ln \frac{RSS}{n} + m \ln(n) \quad (50-2)$$

إذ ان:

RSS: تمثل مجموع مربعات الانحدار

$\hat{\sigma}_a^2$: تمثل مقدر الإمكان الأعظم ل σ_a^2 .

m: تمثل عدد المعلمات في الأنموذج.

σ_z^2 : تمثل تباين العينة للسلسلة الزمنية.

وان الأنموذج الذي يقابل اقل قيمة لمعيار (BIC) هو الأنموذج الأفضل.

The model validity (4-5-2) مرحلة اختبار صلاحية الأنموذج testing phase

في هذه المرحلة يتم فحص جودة الأنموذج ومدى صلاحيته عن طريق مجموعة من الاختبارات الإحصائية. فيكون الأنموذج ملائم وصالح للاستعمال عند اجتياز هذه الاختبارات. وفي حالة لم يجتاز الأنموذج هذه الاختبارات فيصبح الأنموذج غير ملائم ونعود الى الخطوة الثانية. ويتم اختبار صلاحية الأنموذج ومدى مقبوليته عن طريق اختبار معنوية معالم الأنموذج وكذلك اختبار معنوية الأنموذج بصورة عامة واختبارات أخرى سيتم شرحها بالتفصيل وكالاتي.

The significance of the (1-4-5-2) اختبار معنوية معاملات الأنموذج model parameters test

يستعمل اختبار t لاختبار معنوية المعلمات المقدرة للأنموذج المختار، ويستعمل كذلك اختبار t لمعرفة مدى استجابة التغيرات التي تحصل في متغير الاستجابة Y التي تسببه التغيرات التي تحصل في المتغير او المتغيرات المستقلة t في الأنموذج. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{المعلمات غير معنوية}$$

$$H_1: B_j \neq 0 \quad \text{المعلمات معنوية}$$

يتم استعمال إحصاء الاختبار $t_{\frac{\alpha}{2}}$ وبمستوى معنوية 0.05 وبدرجة حرية n- (2). ويعبر عن الاختبار بالمعادلة الآتية.

$$t = \frac{\hat{b}_1 - b_1}{s\hat{b}_1} \quad , \quad t = \frac{\hat{b}_0 - b_0}{s\hat{b}_0} \quad (51-2)$$

إذ ان:

(\hat{b}_0, \hat{b}_1) يمثلان القيمتين المقدرتين لكل من الحد الثابت والميل.

(b_0, b_1) يمثلان القيمتين الحقيقيتين لكل من الحد الثابت والميل.

ثم نقارن t المحسوبة مع t الجدولية، فإذا كانت t المحسوبة أكبر من t الجدولية فنقوم برفض فرضية العدم H_0 ونقبل الفرضية البديلة، أي ان المعلمات معنوية، أما إذا كانت t المحسوبة أصغر من t الجدولية فلا نرفض الفرضية الصفرية H_0 أي ان المعلمات غير معنوية، أو عن طريق قيمة P-Value لإحصائه t فإذا كانت قيمة P-Value أصغر من مستوى المعنوية 0.05 فأنا نرفض فرضية العدم وان المعلمة معنوية أما إذا كانت قيمة P-Value أكبر من مستوى المعنوية 0.05 فلا نرفض فرضية العدم وان المعلمة غير معنوية احصائيا [7].

Test the overall (2-4-5-2) اختبار المعنوية الكلية الأنموذج significance of the model

لاختبار معنوية الأنموذج بشكل عام نستعمل اختبار F وان فرضيات هذا الاختبار هي كالآتي.

$$H_0: B_j = 0 \quad \text{عدم وجود علاقة معنوية بين المتغيرات المستقلة ومتغير الاستجابة}$$

$$H_1: B_j \neq 0 \quad \text{وجود علاقة معنوية بين المتغيرات المستقلة ومتغير الاستجابة}$$

وان صيغة الاختبار هي كالآتي.

$$F = \frac{RSS/k-1}{SSE/n-k} = \frac{\sum y_i^2/k-1}{\sum e_i^2/n-1} \quad (52-2)$$

ثم تتم المقارنة بين قيمة F الحسابية وقيمة F الجدولية وبدرجات حرية $n-k$ للمقام و $k-1$ للبسط ومستوى المعنوية (0.05). فإذا كانت F الحسابية أكبر من F الجدولية فأنا نرفض فرضية العدم H_0 . أي ان هنالك علاقة معنوية بين متغير الاستجابة Y والمتغيرات المستقلة. أي ان الأنموذج معنوي احصائيا. أما إذا كانت قيمة F الحسابية أصغر من F الجدولية فأنا لا نرفض فرضية العدم H_0 . أي لا توجد علاقة معنوية بين متغير الاستجابة Y والمتغيرات المستقلة X . وبالتالي فان الأنموذج غير معنوي احصائيا [7].

Sequential residual test (3-4-5-2) اختبار البواقي المتسلسلة

في بداية دراسة السلوك الدوري للسلسلة الزمنية المستقرة. لا بد لنا من دراسة الظاهرة واي توزيع احتمالي تخضع له هذه الظاهرة. وذلك لإعطاء نظرة بدائية حول طبيعة هذه السلسلة. ان من صفات التوزيع الطبيعي يجب ان يكون معامل الالتواء (Skewness) معدوماً. وكذلك يجب ان يكون معامل التفلطح (Kurtosis) يساوي 3. أي ان في هذا الاختبار يتم معرفة هل ان البواقي تتوزع توزيعاً طبيعياً ام لا. عن طريق اختبار (Jarque-Bera) الذي يجمع بين اختبارين (Skewness) و (Kurtosis). ان الصيغة العامة لمعامل الالتواء (Skewness) تكتب بالشكل الآتي.

$$S = \frac{\left[\frac{1}{n} \sum_{i=1}^n (y_{t-m})^3 \right]^2}{\left[\frac{1}{n} \sum_{i=1}^n (y_{t-m})^2 \right]^3} = \frac{\mu_3^2}{\mu_2^3} = \beta_1 \quad (53-2)$$

وان الصيغة العامة لمعامل التفلطح (Kurtosis) تكتب بالشكل الآتي.

$$K = \frac{\frac{1}{n} \sum_{i=1}^n (y_{t-m})^4}{\left[\frac{1}{n} \sum_{i=1}^n (y_{t-m})^2 \right]^2} = \frac{\mu_4}{\mu_2^2} = \beta_2 \quad (54-2)$$

إذ ان μ_i العزم (i) حول الوسط الحسابي للسلسلة المستقرة وان
(i=1,2,3,4)

وان الصيغة العامة لإحصائه (Jarque-Bera) تكتب بالشكل الآتي.

$$JB = \frac{n}{6} \beta_1 + \frac{T}{24} (\beta_2 - 3)^2 \sim \chi_{\alpha(2)}^2 \quad (55-2)$$

وان فرضيات الاختبار تكتب بالشكل الآتي.

H_0 : البواقي تتوزع طبيعي

H_1 : البواقي لا تتوزع طبيعي

فاذا كانت احصائه $JB > \chi_{\alpha(2)}^2$ فأنا نرفض فرضية العدم H_0 أي ان البواقي لا تتبع التوزيع الطبيعي. اما اذا كانت احصائه $JB < \chi_{\alpha(2)}^2$ فلا نرفض فرضية

العدم H_0 ونقبل الفرضية البديلة H_1 أي ان البواقي تتوزع طبيعياً. او عن طريق قيمة P-Value لإحصائه Jarque-Bera فإذا كانت قيمة P-Value أصغر من مستوى المعنوية 0.05 فأننا نرفض فرضية العدم وان البواقي تتبع التوزيع الطبيعي. اما اذاك كانت قيمة P-Value أكبر من مستوى المعنوية 0.05 فلا نرفض فرضية العدم وان البواقي لا تتبع التوزيع الطبيعي [13].

Forecasting

(5-5-2) التنبؤ

ان من اهم مراحل تحليل السلاسل الزمنية هي مرحلة التنبؤ، وتعد الهدف الأساس لعملية تقدير الأنموذج أي استعمال الأنموذج المقدر للتنبؤ بقيمة متغير الاستجابة Y_t وذلك بالاستناد إلى قيم المتغيرات المستقلة للتعرف على سلوك الظاهرة في المستقبل، ويعرف التنبؤ على انه تحليل البيانات في الماضي وتطبيق النتائج في المستقبل وذلك عن طريق استعمال الأنموذج الرياضي الملائم، ويعد التنبؤ مهم في التخطيط ووضع الافتراضات عن الاحداث في المستقبل باستعمال الأساليب الإحصائية ذات العلاقة، ان التنبؤ يؤدي دورا مهما وبارزا في عملية اتخاذ القرارات المستقبلية أي انه رؤية للمستقبل فيما ستكون عليه الظواهر المختلفة او المتغيرات في المستقبل.

(1-5-5-2) مقاييس الدقة التنبؤية Predictive Accuracy Metrics

تعد الدقة التنبؤية المقياس الرئيسي في اختيار طريقة التنبؤ الملائمة. ان اغلب التنبؤات تميل الى نسبة خطأ معينة او عدم الدقة في التنبؤ. لذا لا بد من اخضاع هذه القيم لاختبارات الدقة التنبؤية عن طريق اجراء اختبار بعض المقاييس. وهي كالآتي [1]، [7].

(1-1-5-5-2) متوسط مربعات الأخطاء Mean Square Error

هو أحد مقاييس القوة التنبؤية. ويتم حساب مقياس MSE عن طريق تربيع جميع الأخطاء ومن ثم نأخذ المتوسط لمربعات الخطأ. ويتم قسمتها على عدد المشاهدات n . كما في المعادلة الآتية [1]، [7].

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (56-2)$$

إذ يتم اختيار الأنموذج الذي يعطي مجموع مربعات اقل.

Root Mean Square Error جذر متوسط مربع الأخطاء (2-1-5-5-2)

وهو الجذر التربيعي لمقياس متوسط مربع الأخطاء (Mean Square Error) [1]. ويعبر عنه بالصيغة الآتية.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_t^2} \quad (57-2)$$

إذ يتم اختيار الأنموذج الذي يعطي مجموع مربعات اقل.

(1-3) تمهيد

ان تطوير الخدمات الصحية وتوسيعها يعد من الأهداف الأساسية للتنمية الاجتماعية والاقتصادية والبشرية، والتي تهدف الى الحد من خطورة او تحجيم الامراض المعدية والفيروسات ولاسيما فيروس Covid-19 الذي اجتاح جميع دول العالم وأصبح يشكل خطرا على حياة الانسان. ان الارتقاء بالواقع الخدمي والصحي الذي يوفر حياة أفضل للإنسان كان الدافع الأساس لبداية البحوث والدراسات التي تتعلق بالتنبؤ بالكثير من الامراض المعدية والفيروسات الخطرة، ولاسيما دراستنا في التنبؤ بفيروس Covid-19. سنتناول في هذا الفصل مبحثين يتكون الأول من تمهيد للفصل وكذلك تعريف فيروس Covid-19 واعراض هذا الفيروس وطرائق الوقاية والعلاج اما الثاني سيتم فيه تطبيق كل ما تم ذكره في الفصل الثاني (الجانب النظري) من اختبارات إحصائية ومقاييس أخرى، إذ يعد الجانب التطبيقي من أساسات البحث العلمي والذي يتضح فيه للقارئ او الباحث التعرف على صفات الظاهرة المدروسة، كذلك الحصول على النتائج ومناقشتها. في هذه الرسالة تم تطبيق سلسلة زمنية بواقع 122 مشاهدة تمثل اعداد الإصابات اليومية بفيروس Covid-19 في محافظة كربلاء المقدسة وذلك من (1/6/2020) لغاية (30/9/2020). تم الحصول على البيانات من الموقع الرسمي لدائرة صحة كربلاء المقدسة، تم تطبيق الاختبارات والمقاييس والمعايير الإحصائية التي تم ذكرها في الفصل الثاني (الجانب النظري) على بيانات السلسلة الزمنية للظاهرة المذكورة آنفاً. إذ تم تحليل هذه البيانات وتطبيق الاختبارات كافة وكذلك تقدير معلمات الانموذجات المستعملة في هذه الرسالة باستعمال برنامج لغة R version (3.6.1).

(2-3) فيروس كورونا واعراضه وطرائق الوقاية والعلاج

هو أحد فصائل الفيروسات المعروفة والواسعة الانتشار والتي تسبب اعراضا مثل الحمى ونزلات البرد وغيرها، تم اكتشاف فيروس كورونا / سارس2 من قبل منظمة الصحة العالمية ولأول مرة في الصين في 31 كانون الأول / ديسمبر 2019 في مدينة ووهان الصينية، انتشر هذا الفيروس من مدن الصين الشعبية الى الدول الأخرى وأصبح يمثل جائحة عالمية يهدد حياة الانسان وسبباً أزمة اقتصادية واجتماعية وصحية، تشمل اعراض فيروس Covid-19 ومنها الأكثر شيوعاً والمتثلة بالحمى والسعال الجاف وكذلك الاجهاد والتعب، اما الاعراض الأقل شيوعاً هي فقدان حاستي الشم والذوق وكذلك واحتقان الانف والم في الحلق والصداع والاسهال والطفح الجلدي، ان بعض الأشخاص المصابين بالعدوى لا تظهر عليهم اعراض الإصابة بهذا الفيروس ويتمثلون للشفاء بمدة قصيرة، تكمن خطورة هذا الفيروس في عملية العدوى إذ أن مدة حضانة هذا الفيروس بما يقارب 14 يوماً إذ لا تظهر اعراض الإصابة بهذا الفيروس الا بعد انتهاء فترة الحضانة وفي هذه المدة ممكن للشخص المصاب ان ينقل العدوى لشخص آخر، ان نسبة الأشخاص المصابين الذين يتمثلون للشفاء من دون الحاجة الى علاج تقدر حوالي 80%. وما يقارب 15% من المصابين تظهر عليهم الاعراض ويصابون بمضاعفات ويحتاجون الى الاوكسجين، اما الأقلية منهم وبنسبة 5% يصبحون بحالة حرجة ويحتاجون الى عناية مركزة. تزداد خطورة الإصابة للأشخاص كبار السن بعمر 60 سنة فأكثر كذلك الأشخاص الذين يعانون من امراض مزمنة مثل مشاكل القلب وارتفاع ضغط الدم والسكر السرطان و... الخ. لذا توصي منظمة الصحة العالمية والمختصون في المجال الطبي بحماية أنفسنا من الإصابة بالعدوى وذلك باتخاذ بعض الاحتياطات مثل التباعد الاجتماعي ولبس الكمامات والتعقيم والتعقيم وغيرها من الإجراءات الوقائية. وقد تكون اللقاحات التي ظهرت مؤخراً للوقاية من فيروس Covid-19 هي الامل لأنهاء هذه الجائحة. إذ يقي هذا اللقاح من الإصابة بالعدوى او الوفاة بسبب هذا الفيروس. كذلك يمنع نقل الفيروس للآخرين. وكلما زاد عدد الافراد الملقحين في المجتمع قل انتشار العدوى وزيادة المناعة الجماعية والتي تسمى بمناعة القطيع. وتوجد ثلاث أنواع من اللقاحات

وهي لقاح فايزر ببيونتك ولقاح موديرنا ولقاح يانسن / جونسن. وتعد هذه اللقاحات جميعها فعالة وآمنة [41].

جدول رقم (١) يمثل اعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة							
Date	Injuries	Date	Injuries	Date	Injuries	Date	Injuries
1/6/2020	11	1/7/2020	115	1/8/2020	264	1/9/2020	220
2/6/2020	31	2/7/2020	145	2/8/2020	294	2/9/2020	250
3/6/2020	11	3/7/2020	122	3/8/2020	300	3/9/2020	150
4/6/2020	50	4/7/2020	114	4/8/2020	296	4/9/2020	243
5/6/2020	31	5/7/2020	146	5/8/2020	285	5/9/2020	335
6/6/2020	34	6/7/2020	138	6/8/2020	244	6/9/2020	201
7/6/2020	28	7/7/2020	75	7/8/2020	252	7/9/2020	194
8/6/2020	47	8/7/2020	125	8/8/2020	239	8/9/2020	227
9/6/2020	25	9/7/2020	141	9/8/2020	239	9/9/2020	248
10/6/2020	24	10/7/2020	80	10/8/2020	299	10/9/2020	199
11/6/2020	38	11/7/2020	121	11/8/2020	278	11/9/2020	166
12/6/2020	38	12/7/2020	119	12/8/2020	287	12/9/2020	164
13/6/2020	43	13/7/2020	103	13/8/2020	285	13/9/2020	48
14/6/2020	38	14/7/2020	129	14/8/2020	257	14/9/2020	169
15/6/2020	27	15/7/2020	122	15/8/2020	325	15/9/2020	219
16/6/2020	40	16/7/2020	118	16/8/2020	295	16/9/2020	165
17/6/2020	68	17/7/2020	141	17/8/2020	243	17/9/2020	156
18/6/2020	44	18/7/2020	137	18/8/2020	235	18/9/2020	131
19/6/2020	67	19/7/2020	161	19/8/2020	218	19/9/2020	123
20/6/2020	100	20/7/2020	171	20/8/2020	278	20/9/2020	49
21/6/2020	44	21/7/2020	187	21/8/2020	336	21/9/2020	108
22/6/2020	82	22/7/2020	185	22/8/2020	304	22/9/2020	120
23/6/2020	54	23/7/2020	176	23/8/2020	221	23/9/2020	154
24/6/2020	55	24/7/2020	187	24/8/2020	80	24/9/2020	131
25/6/2020	113	25/7/2020	208	25/8/2020	239	25/9/2020	86
26/6/2020	67	26/7/2020	240	26/8/2020	335	26/9/2020	99
27/6/2020	92	27/7/2020	263	27/8/2020	260	27/9/2020	19
28/6/2020	65	28/7/2020	195	28/8/2020	293	28/9/2020	68
29/6/2020	84	29/7/2020	231	29/8/2020	250	29/9/2020	117
30/6/2020	119	30/7/2020	250	30/8/2020	269	30/9/2020	80
		31/7/2020	275	31/8/2020	251		

تم الحصول على البيانات من الموقع الرسمي لدائرة صحة كربلاء المقدسة. إذ ان الإصابات (Injuries) تمثل متغير الاستجابة Y_t . ويمثل الزمن t المتغير المستقل.

Properties of time series (3-3) خصائص السلسلة الزمنية

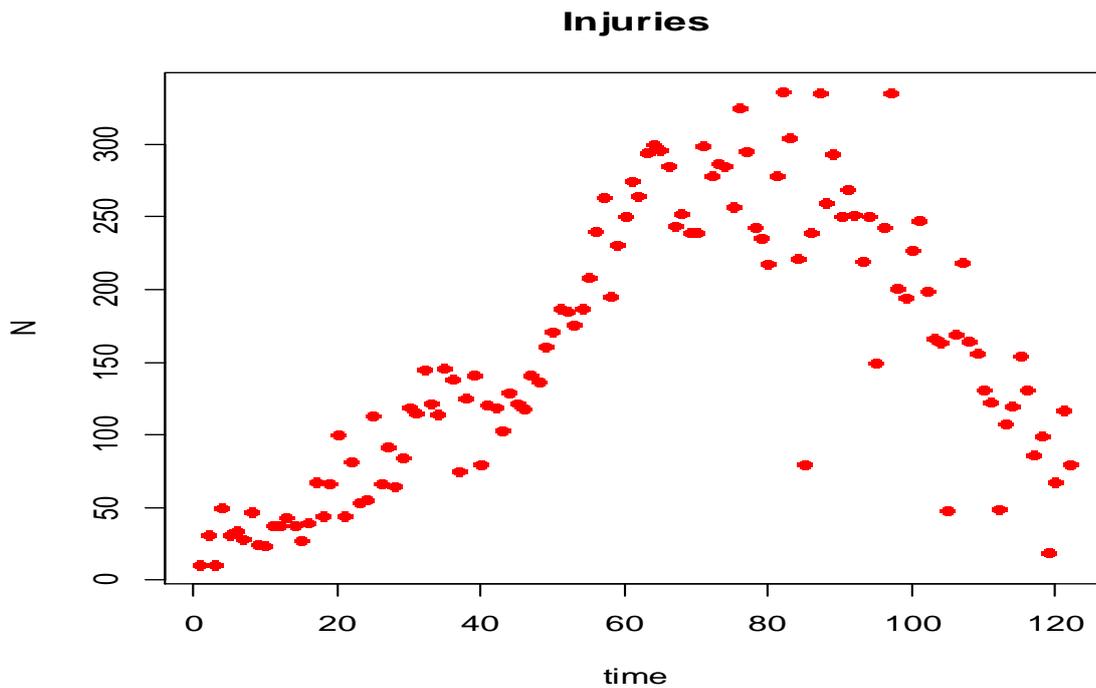
في هذا الفرع سيتم التعرف الى اهم خصائص السلسلة الزمنية لأعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة. وذلك باستعمال برنامج لغة (R version 3.6.1). إذ سيتم التعرف على استقرارية السلسلة الزمنية وكذلك التعرف على خطيتها وهل هي متجانسة ام لا. عن طريق تطبيق الاختبارات الإحصائية الوارد ذكرها في الجانب النظري. وكما يلي

Stationary in time series (1-3-3) استقرارية السلسلة الزمنية

تعد خاصية الاستقرارية في السلاسل الزمنية من الخطوات الأولى في مجال تحليل السلسلة الزمنية. وهناك عدة طرائق واختبارات للكشف عن استقرارية السلسلة الزمنية من عدمها. وسيتم توضيح استقرارية السلسلة الزمنية باستعمال الرسم البياني واختبار Augmented Dicky-Fuller وعلى النحو الآتي.

Graph Method (1-1-3-3) طريقة الرسم البياني

الشكل (8-3) يبين التمثيل البياني لسلسلة اعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة.



الشكل (1-3) يمثل التمثيل البياني للسلسلة الزمنية لأعداد المصابين بفيروس كوفيد-19

يتضح من الشكل البياني (3-8) ان السلسلة الزمنية لأعداد المصابين بفيروس Covid-19 غير مستقرة في المتوسط، ويحتاج تحديد الاستقرار عن طريق الرسم الى خبرة وأحيانا الأشكال البيانية يصعب قراءتها وتحديد هل أن السلسلة مستقرة أو غير مستقرة وخصوصا إذا كان حجم البيانات كبيراً، لذلك يفضل استعمال الاختبارات الرقمية مثل اختبار Augmented Dickey-Fuller.

Augmented Dickey and Fuller test (2-1-3-3) اختبار ديكي-فولر الموسع

يعد اختبار ديكي فولر الموسع Augmented Dickey-Fuller من الاختبارات المهمة للكشف عن استقرار السلسلة الزمنية من عدمها، لذا فان اختبار Augmented Dickey-Fuller ثلاث حالات أو أنماط مختلفة، أختيرت هذه الأنماط مع الاخذ بعين الاعتبار كل من الاتجاه الزمني والحد الثابت. النوع الأول هو في حالة عدم وجود الحد الثابت ولا الاتجاه الزمني. والنوع الثاني في حالة وجود الحد الثابت فقط. اما النوع الثالث في حالة وجود الحد الثابت والاتجاه الزمني. والجدول الآتي يبين نتائج الاختبار في حالة عدم وجود الحد الثابت والاتجاه الزمني. وان فرضيات الاختبار تكتب كالتالي.

$H_0: \delta = 0$ السلسلة تمتلك جذر وحدة وهي غير مستقرة
 $H_1: \delta \neq 0$ السلسلة لا تمتلك جذر وحدة وهي مستقرة

Type 1: No drift no trend

جدول (2-3) اختبار Augmented Dickey-Fuller الموسع في حالة انموذج لا يتضمن الحد الثابت ولا الاتجاه الزمني.

n	Lag	ADF	P-value
1	0	-1.342	0.197
2	1	-1.030	0.309
3	2	-0.611	0.442
4	3	-0.572	0.474
5	4	-0.335	0.497

من الجدول (2-3) وفي حالة عدم وجود الحد الثابت ولا الاتجاه الزمني بينت النتائج أن جميع قيم P-value هي أكبر من مستوى المعنوية (0.05). وهذا يؤدي الى عدم رفض فرضية العدم H_0 التي تنص على ان السلسلة الزمنية تمتلك جذر وحدة وأنها غير مستقرة.

Type 2: With drift no trend

جدول (3-3) اختبار Dickey-fuller المطور في حالة انموذج يتضمن الحد الثابت فقط

n	Lag	ADF	P-value
1	0	-3.02	0.079
2	1	-2.45	0.154
3	2	-1.90	0.368
4	3	-1.60	0.483
5	4	-1.37	0.568

من الجدول (3-3) وفي حالة وجود الحد الثابت فقد بينت النتائج أن جميع قيم P-value هي أكبر من مستوى المعنوية (0.05). وهذا يؤدي الى عدم رفض فرضية العدم H_0 التي تنص على ان السلسلة الزمنية تمتلك جذر وحدة وأنها غير مستقرة.

Type 3: With drift and trend

جدول (4-3) اختبار Dickey-fuller المطور في حالة انموذج يتضمن الحد الثابت والاتجاه الزمني.

n	Lag	ADF	P-value
1	0	-3.0301	0.148
2	1	-2.3083	0.444
3	2	-1.3159	0.860
4	3	-0.8634	0.954
5	4	0.0685	0.990

من الجدول (3-4) أعلاه وفي حالة وجود الحد الثابت والاتجاه الزمني. فقد بينت النتائج أن جميع قيم P-value هي أكبر من مستوى المعنوية (0.05) وبالتالي لا نرفض فرضية العدم H_0 أي ان السلسلة تحتوي على جذر الوحدة وبالتالي فهي غير مستقرة.

أما الصفة او الخاصية الثانية التي يجب دراستها وتحديدها في السلسلة الزمنية هي الخطية كما في المبحث الفرعي الآتي.

Linearly (2-3-3) الخطية

تعد خاصية الخطية في السلاسل الزمنية من الخطوات المهمة في مجال تحليل السلسلة الزمنية. وهناك عدة طرائق واختبارات للكشف عن خطية السلسلة الزمنية من عدمها. سيتم الاعتماد على اختبار Li and Mcleod واختبار Mann-Kendall. وكما يلي

Mcleod and Li اختبار (1-2-3-3)

ان فرضيات الاختبار تكتب كالاتي.

H_0 : السلسلة الزمنية خطية

H_1 : السلسلة الزمنية غير خطية

عن طريق النتائج التي توصلنا اليها لاختبار Li and Mcleod فان قيمة P-value تساوي (0.000) أي أصغر من مستوى المعنوية (0.05) لذا نرفض فرضية العدم وان السلسلة الزمنية غير خطية.

Mann-Kendall (2-2-3-3) اختبار مان كيندال

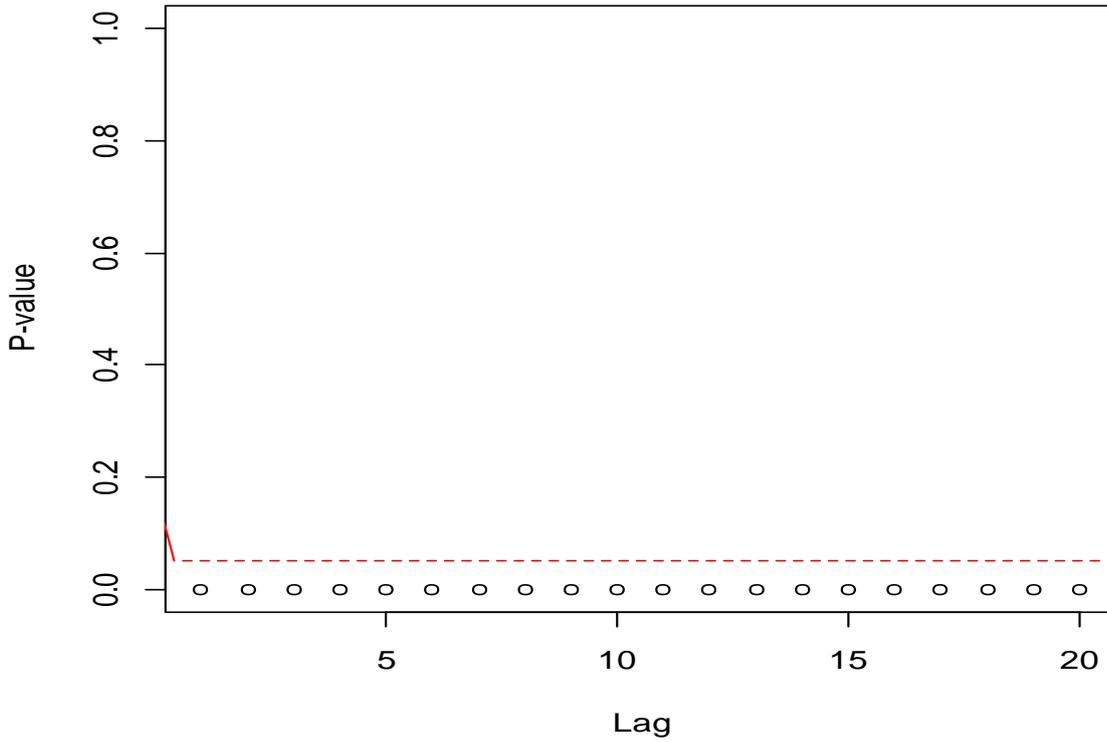
ان فرضيات الاختبار تكتب كالاتي.

H_0 : لا يوجد اتجاه في السلسلة الزمنية وان السلسلة خطية

H_1 : هنالك اتجاه في السلسلة الزمنية وان السلسلة غير خطية

عن طريق النتائج التي توصلنا اليها فان قيمة P-value لاختبار Mann-Kendall تساوي (4.605E-10) أي أصغر من مستوى المعنوية (0.05) لذا نرفض فرضية العدم أي هنالك اتجاه في السلسلة الزمنية وان السلسلة الزمنية غير

خطية. كما ان الشكل البياني الآتي يوضح ان قيمة P-value تقع في منطقة الصفر أي ان السلسلة غير خطية.



الشكل (2-3) يمثل التمثيل البياني لاختبار Mann-Kendal

Homogeneity

(3-3-3) التجانس

هي من الخصائص المهمة في السلاسل الزمنية، وهناك عدة اختبارات للكشف عن تجانس البيانات من عدمها، وهنا سنعتمد على اختبارين فقط وهما اختبار التجانس الطبيعي القياسي (standard normal homogeneity) واختبار فون نيومان (Von Neumann). وكما يأتي

standard normal اختبار التجانس القياسي الطبيعي (1-3-3-3)
homogeneity

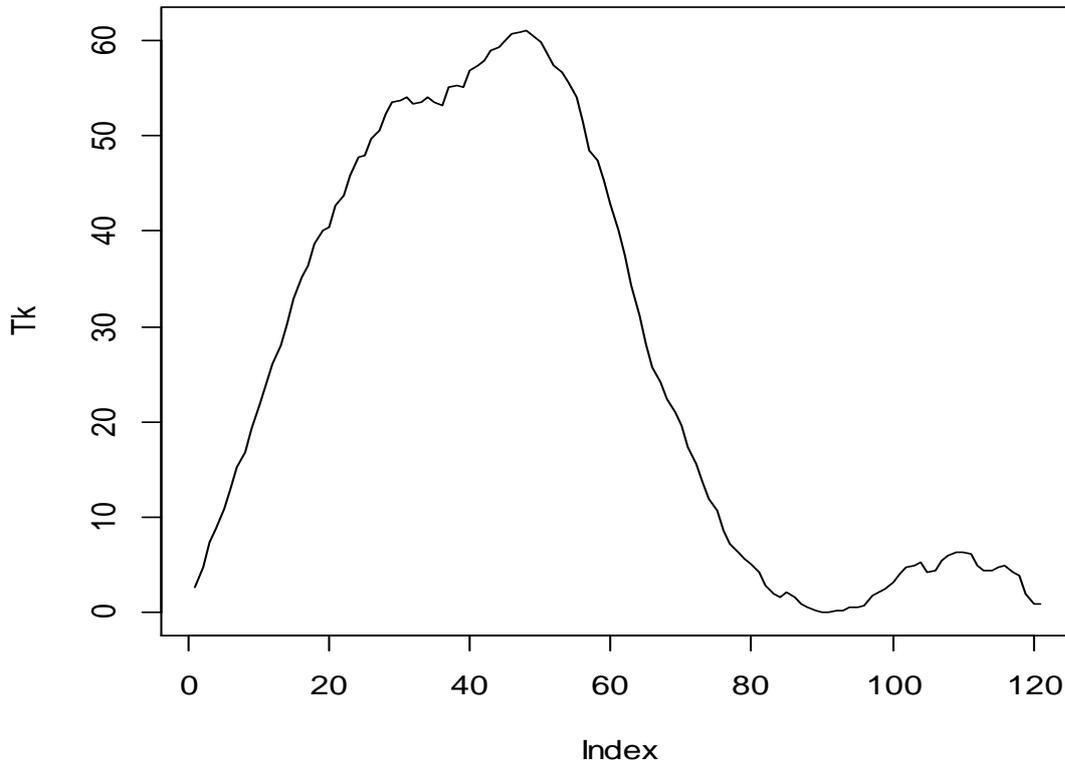
ان فرضيات الاختبار تكتب كالآتي.

H_0 : البيانات تتوزع توزيعا طبيعيا وان البيانات متجانسة

H_1 : البيانات لا تتوزع توزيعا طبيعيا وبالآتي هي غير متجانسة

أن النتائج التي توصلنا اليها تشير الى أن قيمة P-value لاختبار التجانس الطبيعي القياسي تساوي (2.2E-16) وهي أقل من مستوى المعنوية (0.05) لذا ترفض فرضية العدم وان البيانات غير متجانسة. والشكل (3-10) يوضح التمثيل البياني لاختبار التجانس الطبيعي القياسي.

Standard Normal Homogeneity Test (SNHT)



الشكل (3-3) يوضح التمثيل البياني لاختبار التجانس الطبيعي القياسي

Von Neumann test

(2-3-3-3) اختبار فون نيومان

ان فرضيات الاختبار تكتب كالآتي.

H_0 : البيانات متجانسة

H_1 : البيانات غير متجانسة

ان النتائج التي توصلنا اليها تشير الى ان قيمة P-value لاختبار فون نيومان تساوي (2.2E-16) وهي اقل من مستوى المعنوية (0.05) لذا نرفض فرضية العدم وان البيانات غير متجانسة.

بعد دراسة خصائص السلسلة الزمنية والتحقق من استقراريتها وكذلك معرفة خطيتها ومعرفة هل هي متجانسة ام لا. ومعرفة ما الانموذجيات التي تلائم هذه البيانات وهي الانموذجيات اللاخطية. تأتي بعدها مرحلة التقدير.

Estimating

(4-3) تقدير معلمات الانموذجيات اللاخطية

Parameters of Non Linear Models

تم تقدير معلمات الانموذجيات اللاخطية بطريقة Maximum Likelihood على بيانات السلسلة الزمنية لأعداد المصابين بفيروس Covid-19. وبما ان هذه الانموذجيات هي معادلات لا خطية يصعب تقديرها بالطرائق الاعتيادية لذا تم استعمال طريقة Newton-Raphson التكرارية لتقدير هذه المعلمات باستعمال برنامج لغة R version 3.6.1 وتم التوصل الى النتائج الآتية.

Exponential model

(1-4-3) الأنموذج الأسي

كانت القيم المقدرة لمعلمات الأنموذج الأسي Exponential model كما في الجدول الآتي.

جدول (5-3) يمثل المعلمات المقدرة للأنموذج الأسي Exponential model

Parameter	Estimator	Std. Error	t-value	P-Value
α	109.394822	12.44869	8.778	1.30E-14
β	-0.005854	0.001377	-4.251	4.23E-05

بعد تقدير معاملات النموذج الأسّي والتوصل الى النتائج الموضحة كما في الجدول (5-3). فإن النموذج يصبح بالشكل الآتي.

$$\hat{Y}_t = 109.39e^{0.0058t}$$

ولاختبار معنوية معاملات النموذج الأسّي نستعمل اختبار t . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{المعاملات غير معنوية}$$

$$H_1: B_j \neq 0 \quad \text{المعاملات معنوية}$$

ونلاحظ عن طريق الجدول (5-3) ان قيمة P-Value للمعلمتين المقدرتين (α, β) هي $(4.23E-5, 1.30E-14)$ وبالتالي هي اقل من مستوى المعنوية (0.05) . لذا نرفض فرضية العدم H_0 وان المعلمتان المقدرتان (α, β) معنوية احصائيا.

ولاختبار معنوية النموذج الأسّي بشكل عام نستعمل اختبار F . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{انعدام العلاقة المعنوية بين المتغير المستقل ومتغير الاستجابة}$$

$$H_1: B_j \neq 0 \quad \text{وجود علاقة معنوية بين المتغير المستقل ومتغيرات الاستجابة}$$

وعن طريق النتائج التي تم الحصول عليها فإن القيمة الحسابية لاختبار F للنموذج الأسّي هي (235.6653) وهي أكبر من القيمة الجدولية لجدول اختبار F وبدرجة حرية 1 للبسط و120 للمقام وهي (3.92) لذا نرفض فرضية العدم H_0 أي ان النموذج معنوي احصائيا.

تم استعمال اختبار Jarque-Bera لمعرفة هل ان البواقي تتوزع طبيعي ام لا. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0 = 0 \quad \text{البواقي تتوزع طبيعي}$$

$$H_1 \neq 0 \quad \text{البواقي لا تتوزع طبيعي}$$

وعن طريق النتائج التي توصلنا اليها التي أظهرت ان القيمة الاحتمالية P-value لأحصاء Jarque-Bera للنموذج الأسّي تساوي (0.1333) أي أكبر

من مستوى المعنوية 0.05 وبالاتي لا نرفض فرضية العدم H_0 أي البواقي Residual تخضع للتوزيع الطبيعي. وباستعمال هذا الأنموذج تم التوصل الى القيم المقدرة والموضحة في الجدول (6-3) ادناه.

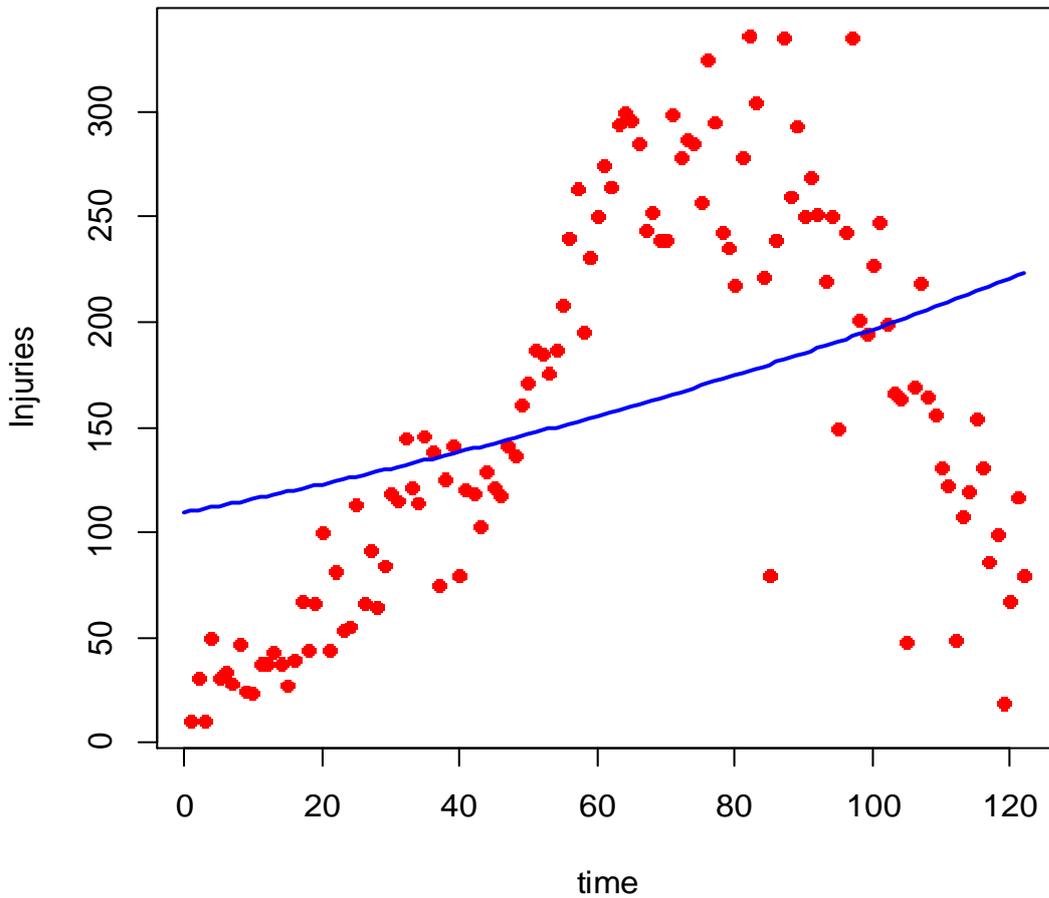
جدول (6-3) يمثل القيم المقدرة للأنموذج الأسّي Exponential model

t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t
1	110.04	32	131.93	63	158.18	93	188.55
2	110.68	33	132.71	64	159.11	94	189.65
3	111.33	34	133.48	65	160.04	95	190.77
4	111.99	35	134.27	66	160.98	96	191.89
5	112.64	36	135.06	67	161.93	97	193.01
6	113.31	37	135.85	68	162.88	98	194.15
7	113.97	38	136.65	69	163.84	99	195.29
8	114.64	39	137.45	70	164.80	100	196.43
9	115.31	40	138.26	71	165.77	101	197.59
10	115.99	41	139.07	72	166.74	102	198.75
11	116.67	42	139.88	73	167.72	103	199.91
12	117.36	43	140.71	74	168.70	104	201.09
13	118.04	44	141.53	75	169.69	105	202.27
14	118.74	45	142.36	76	170.69	106	203.46
15	119.43	46	143.20	77	171.69	107	204.65
16	120.14	47	144.04	78	172.70	108	205.85
17	120.84	48	144.88	79	173.71	109	207.06
18	121.55	49	145.74	80	174.73	110	208.28
19	122.26	50	146.59	81	175.76	111	209.50
20	122.98	51	147.45	82	176.79	112	210.73
21	123.70	52	148.32	83	177.83	113	211.97
22	124.43	53	149.19	84	178.87	114	213.21
23	125.16	54	150.06	85	179.92	115	214.46
24	125.90	55	150.94	86	180.98	116	215.72
25	126.63	56	151.83	87	182.04	117	216.99
26	127.38	57	152.72	88	183.11	118	218.26
27	128.13	58	153.62	89	184.18	119	219.54
28	128.88	59	154.52	90	185.27	120	220.83
29	129.63	60	155.43	91	186.35	121	222.13
30	130.40	61	156.34	92	187.45	122	223.43
31	131.16	62	157.26				

ونلاحظ من الجدول (3-6) ان القيم المقدرة للأنموذج الأسّي Exponential model بعيدة عن القيم الحقيقية لأعداد المصابين بفيروس Covid-19 خصوصا في بداية السلسلة الزمنية وبعدها تزداد تدريجيا لتقترب من القيم الحقيقية، كذلك نلاحظ عن طريق الشكل البياني (3-11) ان الأنموذج الأسّي Exponential model لا يمثل البيانات تمثيلا جيدا وهذا يدل على ان الأنموذج غير ملائم للتنبؤ.

والرسم البياني الآتي يوضح القيم المقدرة للأنموذج الأسّي Exponential model

exponential growth curve



الشكل (3-4) يبين رسم القيم المقدرة للأنموذج الأسّي Exponential model

(2-4-3) أنموذج فون بيرتلانفي Von Bertalanffy model

كانت القيم المقدرة لمعاملات لأنموذج فون بيرتلانفي Von Bertalanffy model كما في الجدول الآتي.

جدول (7-3) يمثل المعلمات المقدرة لأنموذج فون بيرتلانفي Von Bertalanffy

Parameter	Estimator	Std.Error	t-value	P-Value
α	214.99519	14.32978	15.003	2.E-16
β	0.03183	0.00668	4.765	5.36E-06

بعد تقدير معاملات أنموذج Von Bertalanffy والتوصل الى النتائج الموضحة كما في الجدول (7-3). فإن الأنموذج يصبح بالشكل الآتي.

$$\hat{Y}_t = 214.995(1 - e^{-0.031t})$$

ولاختبار معنوية معاملات أنموذج Von Bertalanffy نستعمل اختبار t . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{المعاملات غير معنوية}$$

$$H_1: B_j = 0 \quad \text{المعاملات معنوية}$$

ونلاحظ عن طريق الجدول (7-3) المذكور أنفاً ان قيمة P-Value للمعلمتين المقدرتين (α, β) هي $(2.E-16, 5.36E-06)$ على الترتيب. وبالتالي هي اقل من مستوى المعنوية (0.05) . لذا نرفض فرضية العدم H_0 وان المعلمتين المقدرتين (α, β) معنوية احصائياً.

ولاختبار معنوية أنموذج Von Bertalanffy بشكل عام نستعمل اختبار F . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{عدم وجود علاقة معنوية بين المتغير المستقل ومتغير الاستجابة}$$

$$H_1: B_j \neq 0 \quad \text{وجود علاقة معنوية بين المتغير المستقل ومتغيرات الاستجابة}$$

وعن طريق النتائج التي تم الحصول عليها. فأن القيمة الحسابية لاختبار F لأنموذج فون بيرتالانفي Von Bertalanffy هي (385.099) وهي أكبر من القيمة الجدولية لجدول اختبار F وبدرجة حرية 1 للبسط و120 للمقام وهي (3.92) لذا نرفض فرضية العدم أي ان الأنموذج معنوي احصائيا.

تم استعمال اختبار Jarque-Bera لمعرفة هل ان البواقي تتوزع طبيعي ام لا. وان فرضيات الاختبار تكتب بالشكل الآتي.

H_0 : البواقي تتوزع طبيعي

H_1 : البواقي لا تتوزع طبيعي

وعن طريق النتائج التي توصلنا اليها التي أظهرت ان القيمة الاحتمالية P-value لأحصاء Jarque-Bera لأنموذج Von Bertalanffy تساوي (0.881) أي أكبر من مستوى المعنوية 0.05 وبالتالي لا نرفض فرضية العدم H_0 أي ان البواقي Residual تخضع للتوزيع الطبيعي. وباستعمال هذا الأنموذج تم التوصل الى القيم المقدرة والموضحة في الجدول ادناه.

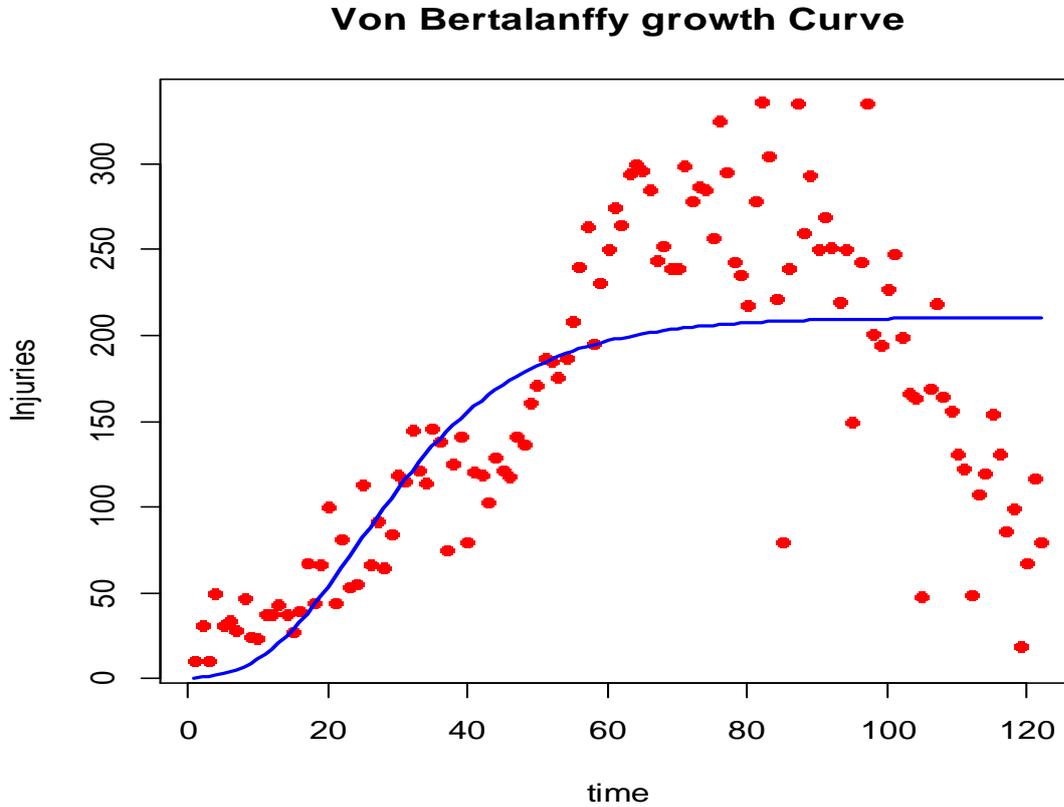
جدول (8-3) يمثل القيم المقدرة لأنموذج فون بيرتلانفي Von Bertalanffy model

t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t
1	6.73488	32	137.35	63	186.05	93	203.85
2	13.2588	33	139.78	64	186.95	94	204.20
3	19.5783	34	142.14	65	187.83	95	204.54
4	25.6999	35	144.42	66	188.68	96	204.87
5	31.6297	36	146.63	67	189.51	97	205.19
6	37.3738	37	148.77	68	190.31	98	205.49
7	42.9379	38	150.85	69	191.08	99	205.79
8	48.3277	39	152.86	70	191.83	100	206.08
9	53.5487	40	154.80	71	192.55	101	206.36
10	58.6061	41	156.69	72	193.26	102	206.63
11	63.5051	42	158.52	73	193.94	103	206.89
12	68.2506	43	160.28	74	194.60	104	207.14
13	72.8475	44	162.00	75	195.24	105	207.39
14	77.3004	45	163.66	76	195.86	106	207.63
15	81.6138	46	165.27	77	196.45	107	207.86
16	85.792	47	166.82	78	197.04	108	208.08
17	89.8394	48	168.33	79	197.60	109	208.30
18	93.76	49	169.80	80	198.14	110	208.51
19	97.5578	50	171.21	81	198.67	111	208.71
20	101.24	51	172.58	82	199.18	112	208.91
21	104.80	52	173.91	83	199.68	113	209.10
22	108.25	53	175.20	84	200.16	114	209.28
23	111.60	54	176.44	85	200.62	115	209.46
24	114.83	55	177.65	86	201.07	116	209.64
25	117.97	56	178.82	87	201.51	117	209.80
26	121.01	57	179.96	88	201.93	118	209.97
27	123.96	58	181.05	89	202.34	119	210.12
28	126.81	59	182.12	90	202.74	120	210.28
29	129.57	60	183.15	91	203.12	121	210.42
30	132.25	61	184.14	92	203.49	122	210.57
31	134.84	62	185.11				

ونلاحظ من الجدول (8-3) ان القيم المقدرة لأنموذج فون بيرتلانفي Von Bertalanffy

model بعيدة عن القيم الحقيقية لأعداد المصابين بفيروس Covid-19 خصوصا في بداية السلسلة الزمنية وبعدها تزداد تدريجيا لتقترب من القيم الحقيقية، كذلك نلاحظ عن طريق الشكل البياني (11-3) ان أنموذج فون بيرتلانفي Von Bertalanffy model لا يمثل البيانات تمثيلا جيدا وهذا يدل على ان الأنموذج غير ملائم للتنبؤ.

والشكل الآتي يوضح التمثيل البياني للقيم المقدرة لأنموذج فون بيرتالانفي Von Bertalanffy.



الشكل (3-5) بين رسم القيم المقدرة لأنموذج فون بيرتالانفي Von Bertalanffy model

(3-4-3) أنموذج جمبرتز Gompertz model

كانت القيم المقدرة لمعاملات أنموذج جمبرتز Gompertz model كما في الجدول الآتي.

جدول (9-3) يمثل المعلمات المقدرة لأنموذج جمبرتز Gompertz model

Parameter	Estimator	Std.Error	t-value	P-Value
α	210.49008	9.21818	22.834	2.E-16
β	6.1107	3.05509	2	4.78E-02
k	0.07505	0.01852	4.052	9.1E-05

بعد تقدير معاملات أنموذج Gompertz والتوصل الى النتائج الموضحة

كما في الجدول (9-3). فإن الأنموذج يصبح بالشكل الآتي.

$$\hat{Y}_t = 210.49e^{-6.11e^{-0.075t}}$$

ولاختبار معنوية معاملات أنموذج Gompertz نستعمل اختبار t . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{المعاملات غير معنوية}$$

$$H_1: B_j \neq 0 \quad \text{المعاملات معنوية}$$

ونلاحظ عن طريق الجدول (3-9) ان قيمة P-Value للمعاملات المقدرة (α, β, k) هي $(2.E-16, 4.78E-02, 9.1E-05)$. وبالتالي هي اقل من مستوى المعنوية (0.05) . لذا نرفض فرضية العدم H_0 وان المعاملات المقدرة (α, β, k) معنوية احصائياً.

ولاختبار معنوية أنموذج Gompertz بشكل عام نستعمل اختبار F . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{عدم وجود علاقة معنوية بين المتغير المستقل ومتغير الاستجابة}$$

$$H_1: B_j \neq 0 \quad \text{وجود علاقة معنوية بين المتغير المستقل ومتغيرات الاستجابة}$$

وعن طريق النتائج التي تم الحصول عليها. فإن القيمة الحسابية لاختبار F هي (295.456) وهي أكبر من القيمة الجدولية لجدول اختبار F وبدرجة حرية 2 للبسط و119 للمقام وهي (3.07) لذا نرفض فرضية العدم أي ان الأنموذج معنوي احصائياً.

تم استعمال اختبار Jarque-Bera لمعرفة هل ان البواقي تتوزع طبيعياً ام لا. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0 \quad \text{البواقي تتوزع طبيعي}$$

$$H_1 \quad \text{البواقي لا تتوزع طبيعي}$$

وعن طريق النتائج التي توصلنا اليها التي أظهرت ان القيمة الاحتمالية P-value لأحصاء Jarque-Bera لأنموذج Gompertz تساوي (0.0524) أي أكبر من مستوى المعنوية (0.05) لذا لا نرفض فرضية العدم H_0 أي ان البواقي تخضع للتوزيع الطبيعي. وباستعمال هذا الأنموذج تم التوصل الى القيم المقدرة والموضحة في الجدول ادناه.

جدول (10-3) يمثل القيم المقدرة لأنموذج جمبرتز Gompertz model

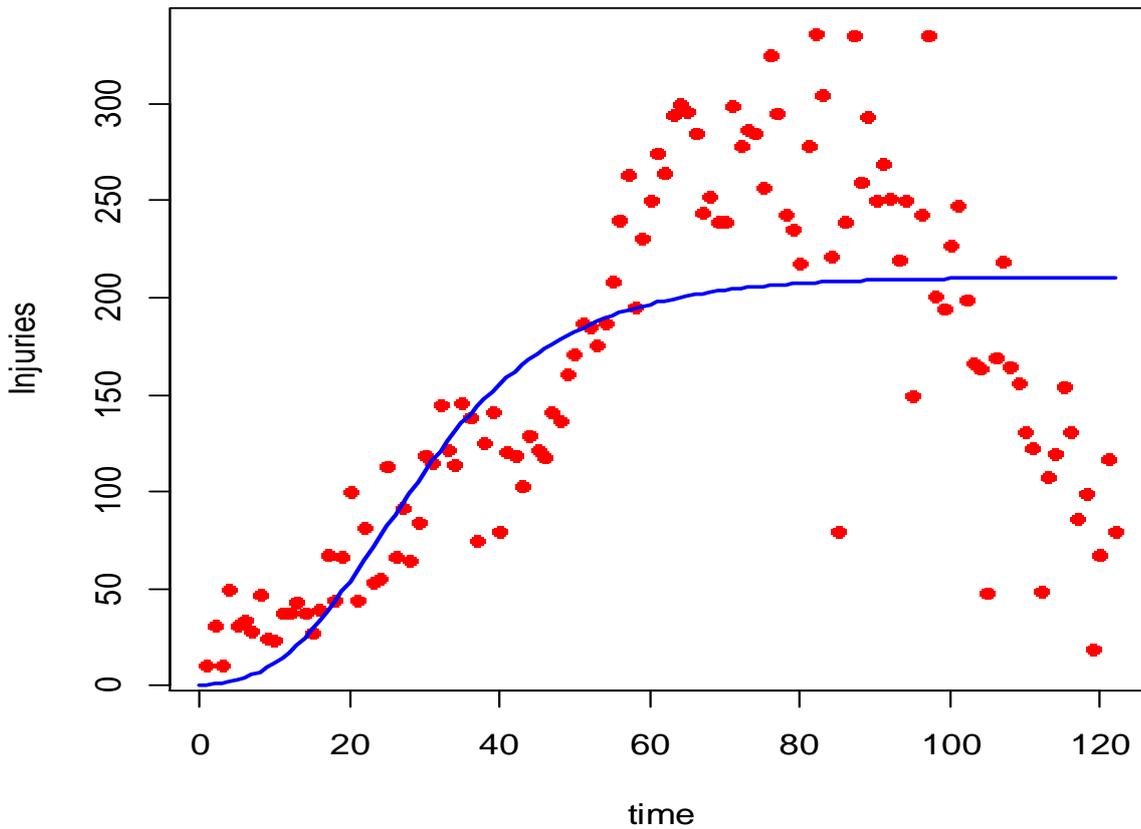
t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t
1	0.72656	32	121.02	63	199.42	93	209.30
2	1.09465	33	125.96	64	200.20	94	209.38
3	1.60107	34	130.73	65	200.93	95	209.46
4	2.27828	35	135.31	66	201.60	96	209.54
5	3.16029	36	139.70	67	202.23	97	209.61
6	4.28125	37	143.90	68	202.82	98	209.67
7	5.6739	38	147.91	69	203.36	99	209.73
8	7.368	39	151.74	70	203.87	100	209.78
9	9.38888	40	155.37	71	204.34	101	209.83
10	11.7562	41	158.82	72	204.78	102	209.88
11	14.483	42	162.09	73	205.19	103	209.93
12	17.5753	43	165.18	74	205.57	104	209.97
13	21.0314	44	168.10	75	205.92	105	210.00
14	24.8426	45	170.85	76	206.25	106	210.04
15	28.9932	46	173.45	77	206.55	107	210.07
16	33.4613	47	175.89	78	206.83	108	210.10
17	38.22	48	178.19	79	207.09	109	210.13
18	43.2377	49	180.35	80	207.34	110	210.16
19	48.4798	50	182.38	81	207.56	111	210.18
20	53.9096	51	184.28	82	207.77	112	210.20
21	59.4892	52	186.06	83	207.97	113	210.22
22	65.1804	53	187.73	84	208.15	114	210.24
23	70.9459	54	189.29	85	208.32	115	210.26
24	76.7496	55	190.74	86	208.48	116	210.28
25	82.5573	56	192.11	87	208.62	117	210.29
26	88.3375	57	193.38	88	208.76	118	210.31
27	94.0609	58	194.57	89	208.88	119	210.32
28	99.7016	59	195.68	90	209.00	120	210.33
29	105.24	60	196.72	91	209.10	121	210.34
30	110.65	61	197.68	92	209.20	122	210.35
31	115.91	62	198.58				

ونلاحظ من الجدول (10-3) ان القيم المقدرة لأنموذج جمبرتز Gompertz model تكون نوعا ما قريبة من القيم الحقيقية لأعداد المصابين بفيروس Covid-19، كذلك نلاحظ

عن طريق الشكل البياني (6-3) ان أنموذج جمبرتز Gomperz model قريب نوعا ما من تمثيل البيانات وهذا يدل على ان الأنموذج ملائم للتنبؤ.

والشكل الآتي يبين رسم القيم المقدرة لأنموذج جمبرتز Gomperz model

Gompertz growth Curve



الشكل (6-3) بين رسم القيم المقدرة لأنموذج جمبرتز Gomperz model

Logistic model

(4-4-3) الأنموذج اللوجستي

كانت القيم المقدرة لمعاملات الأنموذج اللوجستي Logistic model كما في الجدول الآتي.

جدول (11-3) يمثل المعلمات المقدرة للأنموذج اللوجستي Logistic model

Parameter	Estimator	Std. Error	t-value	P-Value
α	211.3218	8.52532	24.788	2.E-16
β	22.07004	15.69987	1.406	1.62E-01
k	0.09921	0.02281	4.35	2.89E-05

بعد تقدير معاملات الأنموذج اللوجستي Logistic model التوصل الى النتائج الموضحة كما في الجدول (3-11). فإن الأنموذج يصبح بالشكل الآتي.

$$\hat{Y}_t = \frac{211.32}{1+22.07e^{-0.0992t}}$$

ولاختبار معنوية معاملات الأنموذج اللوجستي Logistic model نستعمل اختبار t. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{المعاملات غير معنوية}$$

$$H_1: B_j \neq 0 \quad \text{المعاملات معنوية}$$

ونلاحظ عن طريق الجدول (3-11) ان قيمة P-Value للمعاملات المقدره (α, β, k) هي $(2.E-16, 1.62E-01, 2.89E-05)$. وبالتالي هي اقل من مستوى المعنوية (0.05). لذا نرفض فرضية العدم H_0 وان المعاملات المقدره (α, β, k) معنوية احصائياً.

ولاختبار معنوية الأنموذج اللوجستي Logistic model بشكل عام نستعمل اختبار F. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{انعدام العلاقة المعنوية بين المتغير المستقل ومتغير الاستجابة}$$

$$H_1: B_j \neq 0 \quad \text{وجود علاقة معنوية بين المتغير المستقل ومتغيرات الاستجابة}$$

وعن طريق النتائج التي تم الحصول عليها. فإن القيمة الحسابية لاختبار F هي (308.009) وهي أكبر من القيمة الجدولية لجدول اختبار F وبدرجة حرية 2 للبسط و119 للمقام وهي (3.9) لذا نرفض فرضية العدم أي أن الأنموذج معنوي احصائياً.

تم استعمال اختبار Jarque-Bera لمعرفة هل ان البواقي تتوزع توزيعاً طبيعياً ام لا. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: \quad \text{البواقي تخضع للتوزيع الطبيعي}$$

$$H_1: \quad \text{البواقي لا تخضع للتوزيع الطبيعي}$$

وعن طريق النتائج التي توصلنا اليها التي أظهرت ان القيمة الاحتمالية P-value لأحصاء Jarque-Bera للأنموذج اللوجستي تساوي (0.053) أي أكبر

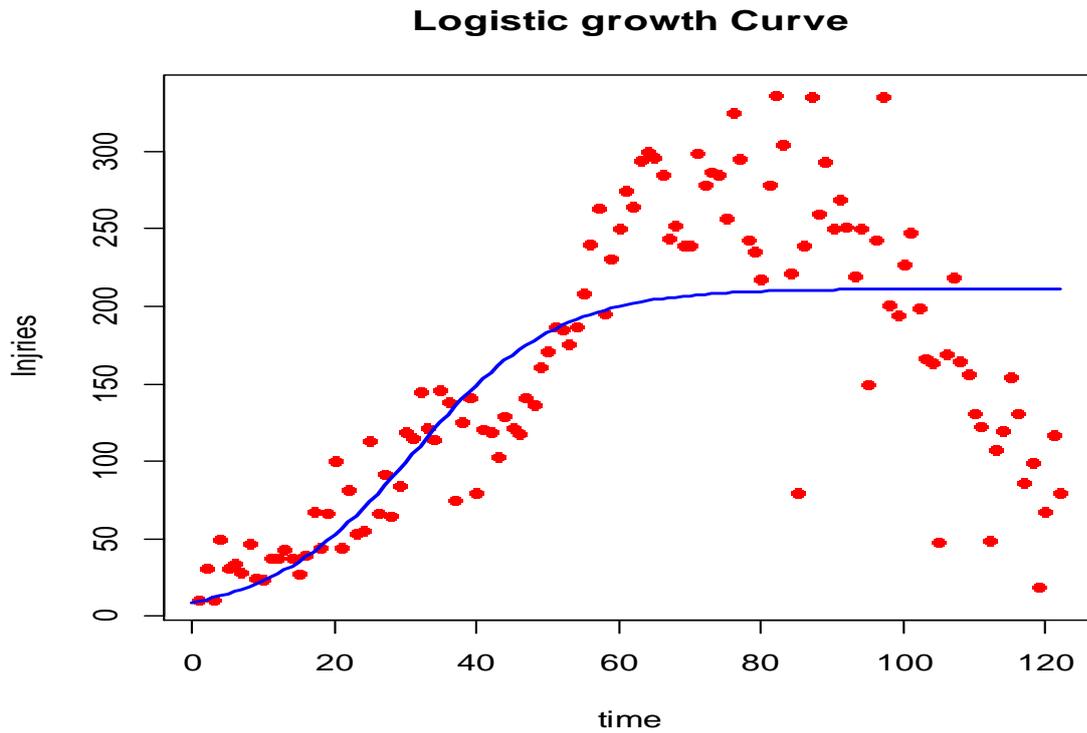
من مستوى المعنوية (0.05) لذا لا نرفض فرضية العدم H_0 أي ان البواقي Residual تخضع للتوزيع الطبيعي. وباستعمال هذا الأنموذج تم التوصل الى القيم المقدرة والموضحة في الجدول الآتي.

جدول (12-3) يمثل القيم المقدرة للأنموذج اللوجستي Logistic model

t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t
1	10.06985	32	109.91	63	202.69	93	210.86
2	11.06512	33	115.13	64	203.47	94	210.91
3	12.15282	34	120.30	65	204.19	95	210.95
4	13.34031	35	125.40	66	204.84	96	210.98
5	14.63531	36	130.41	67	205.44	97	211.01
6	16.04584	37	135.30	68	205.98	98	211.04
7	17.58015	38	140.06	69	206.47	99	211.07
8	19.24672	39	144.67	70	206.92	100	211.09
9	21.05411	40	149.11	71	207.33	101	211.11
10	23.01089	41	153.38	72	207.70	102	211.13
11	25.12552	42	157.45	73	208.04	103	211.15
12	27.40619	43	161.34	74	208.34	104	211.17
13	29.86069	44	165.02	75	208.62	105	211.18
14	32.49617	45	168.51	76	208.87	106	211.20
15	35.31898	46	171.80	77	209.10	107	211.21
16	38.33443	47	174.89	78	209.31	108	211.22
17	41.54656	48	177.78	79	209.50	109	211.23
18	44.95789	49	180.49	80	209.67	110	211.24
19	48.56919	50	183.01	81	209.82	111	211.24
20	52.37924	51	185.36	82	209.96	112	211.25
21	56.38463	52	187.53	83	210.09	113	211.26
22	60.57956	53	189.55	84	210.21	114	211.26
23	64.95576	54	191.41	85	210.31	115	211.27
24	69.50232	55	193.13	86	210.41	116	211.27
25	74.20578	56	194.71	87	210.49	117	211.28
26	79.05011	57	196.17	88	210.57	118	211.28
27	84.01692	58	197.51	89	210.64	119	211.29
28	89.08562	59	198.73	90	210.71	120	211.29
29	94.23375	60	199.86	91	210.76	121	211.29
30	99.43738	61	200.89	92	210.82	122	211.30
31	104.67	62	201.83				

ونلاحظ من الجدول (3-12) ان القيم المقدرة للأنموذج اللوجستي Logistic model تكون أقرب للقيم الحقيقية من بقية الانموذجات الخمسة وكذلك نلاحظ عن طريق الشكل البياني (3-7) ان الأنموذج اللوجستي Logistic model يمثل البيانات تمثيلا جيدا وأفضل من بقية الانموذجات اللاخطية الخمسة المستعملة في هذه الرسالة.

والشكل البياني يبين رسم القيم المقدرة للأنموذج اللوجستي Logistic model



الشكل (3-7) يبين رسم القيم المقدرة للأنموذج اللوجستي Logistic Model

Chapman Richard model (3-4-5) أنموذج جابمان ريتشارد

كانت القيم المقدرة لمعاملات الأنموذج Chapman-Richard model كما في الجدول الآتي.

جدول (3-13) يمثل المعلمات المقدرة للأنموذج جابمان ريتشارد Chapman-Richard

Parameter	Estimator	Std. Error	t-value	P-Value
α	210.5	10.22	20.593	2.E-16
β	0.05	6.644	0.008	9.94E-01
K	0.07503	0.04128	1.817	7.17E-02
θ	121.8	16370	0.007	9.94E-01

بعد تقدير معاملات أنموذج Chapman-Richard والتوصل الى النتائج الموضحة كما في الجدول (3-13) فإن الأنموذج يصبح بالشكل الآتي.

$$Y_t = 210.5(1 - 0.05e^{-0.075t})^{121.8}$$

ولاختبار معنوية معاملات أنموذج جابمان ريتشارد Chapman Richard نستعمل اختبار t . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{المعاملات غير معنوية}$$

$$H_0: B_j = 0 \quad \text{المعاملات معنوية}$$

ونلاحظ عن طريق الجدول (3-13) ان قيمة P-Value للمعاملات المقدره $(\alpha, \beta, k, \theta)$ هي $(2.E-16, 9.94E-01, 7.17E-02, 9.94E-01)$ وبالتالي هي اقل من مستوى المعنوية (0.05) . لذا نرفض فرضية العدم H_0 وان المعاملات المقدره $(\alpha, \beta, k, \theta)$ معنوية احصائياً.

ولاختبار معنوية انموذج جابمان ريتشارد Chapman Richard بشكل عام نستعمل اختبار F . وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: B_j = 0 \quad \text{عدم وجود علاقة معنوية بين المتغير المستقل ومتغير الاستجابة}$$

$$H_1: B_j \neq 0 \quad \text{وجود علاقة معنوية بين المتغير المستقل ومتغيرات الاستجابة}$$

وعن طريق النتائج التي تم الحصول عليها. فإن القيمة الحسابية لاختبار F لأنموذج Chapman Richard هي (308.009) وهي أكبر من القيمة الجدولية لجدول اختبار F وبدرجة حرية 3 للبسط و118 للمقام وهي (2.68) لذا نرفض فرضية العدم أي أن الأنموذج معنوي إحصائياً.

تم استعمال اختبار Jarque-Bera لمعرفة هل ان البواقي تخضع للتوزيع الطبيعي ام لا. وان فرضيات الاختبار تكتب بالشكل الآتي.

$$H_0: \quad \text{البواقي تتوزع طبيعي}$$

$$H_1: \quad \text{البواقي لا تتوزع طبيعي}$$

وعن طريق النتائج التي توصلنا اليها التي أظهرت ان القيمة الاحتمالية P-value لأحصاءة اختبار Jarque-Bera لأنموذج Chapman-Richard تساوي (0.055) أي

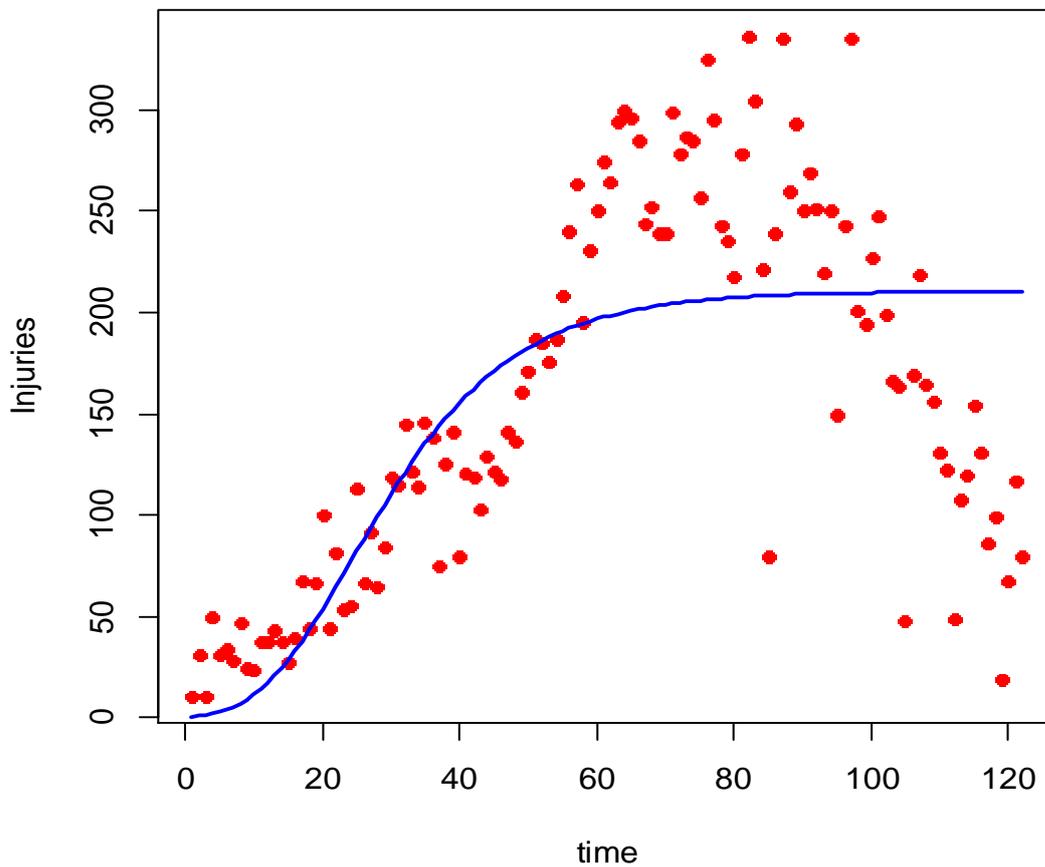
أكبر من مستوى المعنوية 0.05 وبالتالي لا نرفض فرضية العدم H_0 أي ان البواقي تخضع للتوزيع الطبيعي. وباستعمال هذا الأنموذج تم التوصل الى القيم المقدرة والموضحة في الجدول الآتي.

جدول (3-14) يمثل القيم المقدرة لأنموذج جابمان ريتشارد Chapman-Richard model

t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t	t	\hat{Y}_t
1	0.64798	32	121.06	63	199.42	93	209.28
2	0.99349	33	126.01	64	200.20	94	209.37
3	1.47491	34	130.78	65	200.93	95	209.45
4	2.12553	35	135.36	66	201.60	96	209.52
5	2.98039	36	139.76	67	202.23	97	209.59
6	4.07479	37	143.96	68	202.82	98	209.65
7	5.44262	38	147.97	69	203.36	99	209.71
8	7.11472	39	151.80	70	203.87	100	209.77
9	9.11736	40	155.43	71	204.34	101	209.82
10	11.471	41	158.88	72	204.77	102	209.86
11	14.189	42	162.14	73	205.18	103	209.91
12	17.2778	43	165.23	74	205.56	104	209.95
13	20.7356	44	168.15	75	205.91	105	209.99
14	24.5534	45	170.90	76	206.24	106	210.02
15	28.7151	46	173.50	77	206.54	107	210.05
16	33.1984	47	175.94	78	206.82	108	210.08
17	37.9753	48	178.24	79	207.08	109	210.11
18	43.0139	49	180.39	80	207.33	110	210.14
19	48.2787	50	182.41	81	207.55	111	210.16
20	53.7323	51	184.31	82	207.76	112	210.18
21	59.3361	52	186.09	83	207.96	113	210.21
22	65.0514	53	187.75	84	208.14	114	210.22
23	70.8404	54	189.31	85	208.31	115	210.24
24	76.6666	55	190.77	86	208.46	116	210.26
25	82.4955	56	192.13	87	208.61	117	210.27
26	88.2951	57	193.40	88	208.74	118	210.29
27	94.0364	58	194.59	89	208.86	119	210.30
28	99.693	59	195.69	90	208.98	120	210.31
29	105.24	60	196.73	91	209.09	121	210.33
30	110.66	61	197.69	92	209.19	122	210.34
31	115.94	62	198.59				

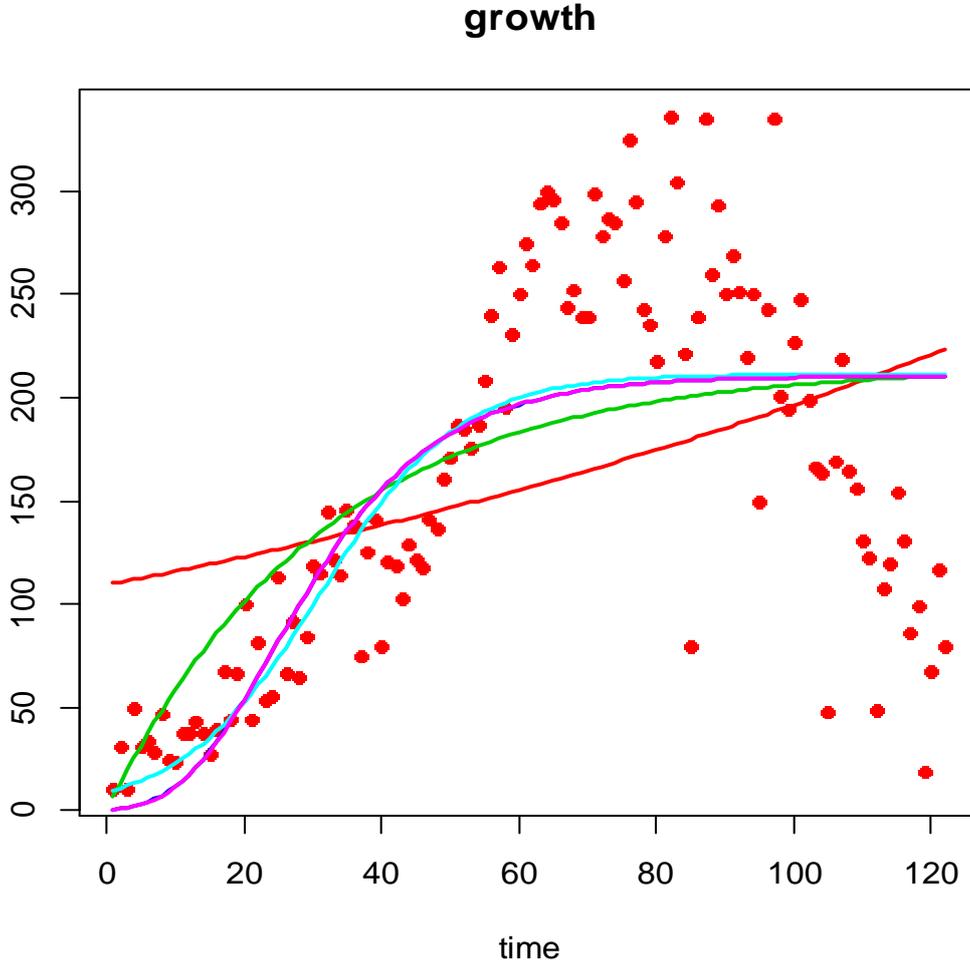
ونلاحظ من الجدول (3-14) ان القيم المقدرة لأنموذج جابمان ريتشارد -Chapman-Richard model بعيدة عن القيم الحقيقية لأعداد المصابين بفيروس Covid-19 خصوصا في بداية السلسلة الزمنية وبعدها تزداد تدريجيا لتقترب من القيم الحقيقية، كذلك نلاحظ عن طريق الشكل البياني (3-8) لأنموذج جابمان ريتشارد -Chapman-Richard model لا يمثل البيانات تمثيلا جيدا وهذا يدل على ان الأنموذج غير ملائم للتنبؤ.

Chapman Richard growth Curve



الشكل (3-8) يبين رسم القيم المقدرة لأنموذج جابمان-ريتشارد -Chapman-Richard

والشكل البياني (9-3) يمثل رسم القيم المقدرة للأنموذجات اللاخطية الخمسة.



الشكل (9-3) يمثل رسم القيم المقدرة للأنموذجات اللاخطية

اختيار الأنموذج الأفضل (5-3) Choosing the Best model

بعد تقدير معاملات الانموذجات اللاخطية وفحص معنوية المعلمات وكذلك معرفة معنوية هذه الانموذجات. تأتي بعدها مرحلة اختيار الأنموذج الأفضل الذي يمثل بيانات سلسلة اعداد المصابين بفيروس Covid-19. عن طريق المعايير الإحصائية وهي (AIC, BIC, H-Q). كما مبين في الجدول الآتي.

جدول (3-15) يمثل قيم معايير المفاضلة بين الأنموذجات اللاخطية

	Exponential	Bertalanffy	Gompertz	Logistic	Chapman-R
AIC	1081.152	1030.89	1017.026	1012.542	1019.075
BIC	1068.76	1036.498	1025.438	1020.945	1030.292
H-Q	1083.43	1033.168	1020.443	1015.959	1023.631

وعن طريق النتائج التي توصلنا اليها والمبينة في الجدول (3-15) وباستعمال المعايير الثلاثة (AIC,BIC,H-Q) إذ اتفقت هذه المعايير على ان الأنموذج اللوجستي Logistic model هو أفضل أنموذج يلائم بيانات اعداد المصابين بفيروس Covid-19 والذي يعطي اقل قيمة لهذه المعايير. ويأتي بعده أنموذج Gompertz وأنموذج Von Bertalanffy وأنموذج Chapman-Richard والأنموذج الأسّي Exponential model على الترتيب. وللمقارنة بين نتائج التنبؤ للنماذج الخمسة على الرغم من ان المعايير اعطت بان الأنموذج اللوجستي هو الافضل لكن تم استعمال النماذج الخمسة للتنبؤ بأعداد الاصابات اليومية لفيروس Covid-19 في محافظة كربلاء المقدسة وكما في المبحث الآتي.

Forecasting

(6-3) التنبؤ

تم استعمال الانموذجات اللاخطية للتنبؤ بأعداد المصابين بفيروس Covid-10 في محافظة كربلاء المقدسة. والجدول الآتية تمثل القيم التنبؤية باستعمال الانموذجات اللاخطية وبعده 12 مشاهدة خارج المدة.

جدول (3-16) يمثل القيم التنبؤية للأنموذج الأسّي Exponential model

1/10/2020	224.7439	7/10/2020	232.7774
2/10/2020	226.0633	8/10/2020	234.1441
3/10/2020	227.3905	9/10/2020	235.5188
4/10/2020	228.7254	10/10/2020	236.9041
5/10/2020	230.0682	11/10/2020	238.2922
6/10/2020	231.4189	12/10/2020	239.6912

جدول (17-3) يمثل القيم التنبؤية لأنموذج فون بيرتلانفي Von Bertanffy model

1/10/2020	210.7068	7/10/2020	211.4523
2/10/2020	210.8411	8/10/2020	211.5633
3/10/2020	210.9713	9/10/2020	211.6708
4/10/2020	211.0973	10/10/2020	211.7749
5/10/2020	211.2194	11/10/2020	211.8758
6/10/2020	211.3377	12/10/2020	211.9735

جدول (18-3) يمثل القيم التنبؤية لأنموذج جميرتز Gompertz model

1/10/2020	210.3641	7/10/2020	210.4098
2/10/2020	210.3732	8/10/2020	210.4156
3/10/2020	210.3817	9/10/2020	210.421
4/10/2020	210.3895	10/10/2020	210.426
5/10/2020	210.3968	11/10/2020	210.4306
6/10/2020	210.4035	12/10/2020	210.4349

جدول (19-3) يمثل القيم التنبؤية لأنموذج اللوجستي Logistic model

1/10/2020	211.2984	7/10/2020	211.3089
2/10/2020	211.3006	8/10/2020	211.3101
3/10/2020	211.3026	9/10/2020	211.3112
4/10/2020	211.3044	10/10/2020	211.3122
5/10/2020	211.3061	11/10/2020	211.3131
6/10/2020	211.3076	12/10/2020	211.3139

جدول (20-3) يمثل القيم التنبؤية لأنموذج جابمان ريتشارد Chapman Richard model

1/10/2020	210.3462	7/10/2020	210.3918
2/10/2020	210.3553	8/10/2020	210.3967
3/10/2020	210.3637	9/10/2020	210.4029
4/10/2020	210.3715	10/10/2020	210.79
5/10/2020	210.3788	11/10/2020	210.4126
6/10/2020	210.3855	12/10/2020	210.4169

عن طريق الجداول (16-3)، (17-3)، (18-3)، (19-3)، (20-3) المذكورة آنفاً نلاحظ ان هنالك تقارب بين القيم الحقيقية والقيم التنبؤية للأنموذجات اللاخطية المستعملة في هذه الدراسة، كذلك نلاحظ التقارب الكبير بين القيم التنبؤية للأنموذجات الخمسة. اما المعلمات المقدرة وحسب اختبار t وقيمة الاحتمال prob المرافقة لكل معلمة نجد ان جميع معلمات الانموذجات اللاخطية هي معلمات معنوية ومؤثرة في الأنموذج وذات دلالة إحصائية. كذلك اختبار F إذ أظهرت النتائج ان جميع الانموذجات معنوية احصائيا وهي صفة جيدة ومرغوبة في الأنموذج وإمكانية الاعتماد على نتائجه للتنبؤ. ولقياس الكفاءة التنبؤية للأنموذجات الخمسة المستعملة في هذه الدراسة سوف نستعمل كل من مقياس (MSE) ومقياس (RMSE).

(7-3) مقاييس الدقة التنبؤية Predictive Accuracy Metrics

تعد مقاييس الدقة التنبؤية مهمة لاختيار الأنموذج الذي يعطي أفضل قيم تنبؤية للظاهرة المدروسة. إذ ان اغلب التنبؤات تمتلك نسبة خطأ معينة او عدم الدقة في التنبؤ. لذا لا بد من اخضاع هذه الانموذجات لهذه المقاييس وهنا سوف نستعمل مقياسين لقياس جودة التنبؤ لهذه الانموذجات وهي مقياس متوسط مربعات الأخطاء (Mean Square Error) ومقياس جذر متوسط مربعات الأخطاء (Root Mean Square Error). وكما مبين في الجدول الآتي.

جدول (21-3) يبين مقاييس القوة التنبؤية للأنموذجات اللاخطية

	Exponential	Von Bertalanffy	Gompertz	Logistic	Chapman-Richard
MSE	6944.074	4599.319	4072.445	3925.481	4108.611
RMSE	83.331	67.818	63.815	62.653	64.098

ومن نتائج الجدول (21-3) وبالاعتماد على مقاييس القوة التنبؤية إذ أظهرت النتائج ان القيم التنبؤية للأنموذج اللوجستي Logistic model هي أكثر دقة من نتائج القيم التنبؤية للأنموذجات الأخرى والذي يعطي اقل قيمة لهذه المقاييس. وهذا يدل على كفاءة الأنموذج المقترح والذي يمثل الأنموذج الملائم للبيانات اليومية لأعداد المصابين بفيروس Covid-19.

Conclusions

(1-4) الاستنتاجات

بعد النتائج التي تم الحصول عليها بواسطة برنامج لغة (R version 3.6.1). وبعد تحليلها ومناقشتها في الجانب التطبيقي توصل الباحث الى بعض الاستنتاجات الموضحة كالآتي

1- ان سلسلة اعداد المصابين بفيروس Covid-19 في محافظة كربلاء المقدسة للمدة من (2020-6-1) الى (2020-9-30). تمثل سلسلة زمنية غير مستقرة وغير خطية وغير متجانسة وذلك حسب الاختبارات الإحصائية المستعملة في هذه الدراسة.

2- عند المقارنة بين الانموذجات اللاخطية للسلاسل الزمنية المستعملة في هذه الدراسة عن طريق المعايير الإحصائية (AIC, BIC, H-Q). وجد ان الأنموذج الأكثر ملائمة هو الأنموذج اللوجستي Logistic model ويأتي بعده أنموذج جمبرتز Gompertz model ثم أنموذج فون بيرتلانفي Von Bertalanffy model ومن ثم انموذج جابمان ريتشارد Chapman-Richard model واخيرا الأنموذج الأسّي Exponential model وحسب اقل قيمة لهذه المعايير.

3- باستعمال اختبار F لاختبار معنوية الانموذجات اللاخطية بينت النتائج ان جميع هذه الانموذجات معنوية احصائيا وهذا يدل على كفاءة الانموذجات للتنبؤ. كذلك باستعمال اختبار t لاختبار معنوية معاملات الأنموذج بشكل خاص، إذ بينت النتائج ان جميع معاملات هذه الأنموذجات معنوية احصائيا.

4- هنالك تقارب بين القيم الحقيقية لأعداد الإصابات بفيروس Covid-19 والقيم المقدرة كما هو موضح في الجانب التطبيقي ولجميع الانموذجات اللاخطية.

5- هنالك تقارب بين القيم التنبؤية للمدد اللاحقة ولجميع الانموذجات اللاخطية. ولكن من حيث الأفضل فان الأنموذج اللوجستي Logistic model هو أفضل أنموذج للتنبؤ بأعداد الإصابات بفيروس Covid-19 في محافظة كربلاء المقدسة. وذلك عن طريق مقاييس القوة التنبؤية (MSE, RMSE). إذ يعطي اقل قيمة لهذه المعايير. وهذا يدل على كفاءة الأنموذج للتنبؤ.

Recommendations

(2-4) التوصيات

في ضوء الاستنتاجات التي تم التوصل اليها قام الباحث بوضع عدة توصيات وهي كالآتي.

- 1- نوصي بدراسة أنموذجات السلاسل الزمنية اللاخطية متعددة المتغيرات للحالات التي تتطلب بحوث للتنبؤ في المجالات الطبية والفيروسات الأخرى التي تتأثر بأكثر من متغير مستقل واحد مثل متغير الجنس والعمر وغيرها.
- 2- نوصي دائرة صحة كربلاء المقدسة والجهات المعنية بالاعتماد على نتائج القيم التنبؤية للأنموذج اللوجستي Logistic model للتنبؤ بأعداد الاصابات ووضع الخطط المستقبلية على ضوء هذه الاعداد لمواجهة هذه الجائحة.
- 3- تعميم هذه الدراسة الى دراسات مشابهة في بقية محافظات العراق والأخذ بالنتائج والأفادة منها في وضع الخطط المستقبلية للتغلب على هذه الجائحة.

المصادر العربية

القرآن الكريم

1. التميمي، رعد فاضل، العنبيكي، عدي طه، "مبادئ السلاسل الزمنية: نماذج التخطيط الاستراتيجي"، بغداد 2013.
2. حسن، يحيى حمود، زكي، حسام الدين، (2012)، "تحليل العلاقة بين أسواق النفط والسياسة النفطية العراقية بالاعتماد على السلاسل الزمنية"، مجلة الغري للعلوم الاقتصادية والإدارية، المجلد 8، العدد 25.
3. حموي، نادر إسكندر، (2018)، "استعمال معياري المعلومات أكايك (AIC) وشوارتز (SC) في المفاضلة بين نماذج النمو غير الخطية لأنواع السمكية المختلفة"، مجلة جامعة البعث، المجلد 40، العدد 3.
4. سكاف، ميشيل، أبو زخم، عبد الله، مثبتوت، شفاء، (2011)، "تغيرات الحرارة في المناطق شبه الجافة والجافة وشديدة الجفاف في سورية وآثارها الكامنة في الغطاء النباتي"، المجلة العربية للبيئات الجافة، العدد 6، ص: (42-54).
5. السيفو، وليد إسماعيل، شلوف، فيصل مفتاح، "أساسات في الاقتصاد القياسي التحليلي: نظرية الاقتصاد القياسي والاختبارات القياسية من الدرجة الأولى"، دار الأهلية للنشر والتوزيع، عمان، الأردن. 2006
6. الصراف، نزار مصطفى، شومان، عبد اللطيف حسن، "السلاسل الزمنية والأرقام القياسية"، دار الدكتور للعلوم الإدارية والاقتصادية، بغداد 2013.
7. عبيد، حميد، "الاقتصاد القياسي"، دار الكتب، موزعون وناشرون، العراق، كربلاء، 2017.
8. عزيز، ميسون مال الله، عبد الله، أسماء عبد المنعم، (2005)، "بناء نظام ديناميكي للسلاسل الزمنية بمعلمات قليلة"، المجلة العراقية للعلوم الإحصائية، العدد 8، ص: (140-165).
9. قاسم، سماح محمد، يونس، عبد الوهاب محمد، عبد الغني، عمر مقداد، (2020)، "فحص التجانس لسلاسل الأمطار في مناطق مختارة من العراق"، مجلة الرافدين الهندسية، المجلد 25، العدد 1، ص: (126-137).

10. الكلابي، عدنان كريم نجم الدين (2018) "استعمال بعض طرائق التنبؤ المختلفة لتحليل اعداد المصابين بالأورام الخبيثة" رسالة ماجستير في علوم الإحصاء، كلية الإدارة والاقتصاد / جامعة كربلاء.
11. المحمد، حسين، (2013)، "تجانس القيم السنوية للهطول في سورية"، المجلة العربية للبيئات الجافة، العدد 7، ص: (2-1).
12. محمد، رملي، "تحليل السلاسل الزمنية"، الجزائر، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، 2020.
13. محمد، شيخي، "طرائق الاقتصاد القياسي"، دار الحامد للنشر والتوزيع، عمان، الطبعة الأولى، 2012.
14. ناصر، جنان عباس، (2012)، "اختبار الفرضية الخطية لمبيعات الشركة العامة للصناعات الكهربائية"، مجلة العلوم الاقتصادية والإدارية، المجلد 18، العدد 67، ص: (304-289).
15. الناصر، عبد المجيد حمزة، جمعة، أحلام احمد (2007) "المقارنة بين طرائق تحديد رتبة انموذج الانحدار الذاتي الطبيعي باستعمال بيانات مولدة وبيانات لبعض العناصر المناخية في العراق"، مجلة العلوم الاقتصادية والإدارية، كلية الإدارة والاقتصاد، جامعة بغداد، العراق، العدد (48)، ص: 272-251.
16. الناصر، عبد المجيد حمزة، رشيد، ظافر حسين، "الأرقام القياسية وتحليل السلاسل الزمنية"، بغداد، جامعة بغداد، 2013.

References

17. Agha, O. M. A. M., Bağçacı, S. Ç., & Şarlak, N. (2017). Homogeneity analysis of precipitation series in North Iraq. IOSR Journal of Applied Geology and Geophysics, 5(03), 57-63.
18. Ahmadi, A., Fadaei, Y., Shirani, M., & Rahmani, F. (2020). Modeling and forecasting trend of COVID-19 epidemic in Iran until May 13, 2020. Medical Journal of the Islamic Republic of Iran, 34, 27.

19. AMARAL, M. T. R., CONCEIÇÃO, K. S., de ANDRADE, M. G., & PADOVANI, C. R. (2020). Generalized growth curve model for covid-19 in brazilian states. *Revista Brasileira de Biometria*, 38(2), 125-146.
20. Celik, S., Ankarali, H., & Pasin, O. (2020). Modeling of COVID-19 Outbreak Indicators in China Between January and June. *Disaster Medicine and Public Health Preparedness*, 1-9.
21. Chakraborty, B., & Basu, A. (2008). A Natural Goodness-of-Fit Testing Procedure for the Logistic Growth Curve Model. *Calcutta Statistical Association Bulletin*, 60(1-2), 53-70.
22. Clark, A. J., Lake, L. W., & Patzek, T. W. (2011, October). Production forecasting with logistic growth models. In *SPE annual technical conference and exhibition*. OnePetro.
23. Franses, P. H. (1994). Gompertz curves with seasonality. *Technological forecasting and social change*, 45(3), 287-297.
24. Gbangboche, A. B., Glele-Kakai, R., Salifou, S., Albuquerque, L. G. D., & Leroy, P. L. (2008). Comparison of non-linear growth models to describe the growth curve in West African Dwarf sheep. *Animal*, 2(7), 1003-1012.
25. Hossein-Zadeh, N. G. (2015). Modeling the growth curve of Iranian Shall sheep using non-linear growth models. *Small Ruminant Research*, 130, 60-66.
26. Jaheen, Z. F. (2003). Prediction of progressive censored data from the Gompertz model. *Communications in Statistics-Simulation and Computation*, 32(3), 663-676.

27. Jia, L., Li, K., Jiang, Y., & Guo, X. (2020). Prediction and analysis of coronavirus disease 2019. arXiv preprint arXiv:2003.05447.
28. Ma, J. (2020). Estimating epidemic exponential growth rate and basic reproduction number. *Infectious Disease Modelling*, 5, 129-141.
29. Mansano, C. F. M., Stéfani, M. V. D., Pereira, M. M., & Macente, B. I. (2012). Non-linear growth models for bullfrog tadpoles. *Ciência e Agrotecnologia*, 36, 454-462.
30. Medina-Mendieta, J. F., Cortés-Cortés, M., & Cortés-Iglesias, M. (2020). COVID-19 Forecasts for Cuba Using Logistic Regression and Gompertz Curves. *MEDICC review*, 22(3).
31. Mohammadi, Y., Mokhtari, M. S., Saghi, D. A., & Shahdadi, A. R. (2019). Modeling the growth curve in Kordi sheep: The comparison of non-linear models and estimation of genetic parameters for the growth curve traits. *Small Ruminant Research*, 177, 117-123.
32. Mueller, L. D., Nusbaum, T. J., & Rose, M. R. (1995). The Gompertz equation as a predictive tool in demography. *Experimental gerontology*, 30(6), 553-569.
33. Ratkowsky, D. A. (1986). Statistical properties of alternative parameterizations of the von Bertalanffy growth curve. *Canadian Journal of Fisheries and Aquatic Sciences*, 43(4), 742-747.
34. Salman Dawood, A., Ashour Akesh, A., & Sagban Khudier, A. (2018). Study of Surface Water Quality and Trends Assessment at Shatt Al-Arab River in Basrah Province. *journal of kerbala university*, 14(1), 215-231.
35. Slavov, N., Budnik, B. A., Schwab, D., Airoidi, E. M., & van Oudenaarden, A. (2014). Constant growth rate can be supported by

decreasing energy flux and increasing aerobic glycolysis. Cell reports, 7(3), 705-714.

36. Tjørve, K. M., & Tjørve, E. (2017). The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family. PloS one, 12(6), e0178691.

37. Topal, M., & Bolukbasi, Ş. C. (2008). Comparison of nonlinear growth curve models in broiler chickens. Journal of Applied Animal Research, 34(2), 149-152.

38. Trappey, C. V., & Wu, H. Y. (2008). An evaluation of the time-varying extended logistic, simple logistic, and Gompertz models for forecasting short product lifecycles. Advanced Engineering Informatics, 22(4), 421-430.

39. Uba, G., Yakasai, H. M., Abubakar, A., & Abd Shukor, M. Y. (2020). Predictive Mathematical Modelling of the Total Number of COVID-19 Cases for Brazil. Journal of Environmental Microbiology and Toxicology, 8(1), 16-20.

40. Wang, X. S., Wu, J., & Yang, Y. (2012). Richards model revisited: Validation by and application to infection dynamics. Journal of theoretical biology, 313, 12-19.

41. <https://www.who.int/ar/news-room/q-a-detail/coronavirus-disease-covid-19>.

ملحق (1): اختبارات السلسلة الزمنية (الاستقرارية، الخطية، التجانس)

```
#install.packages("aTSA")
#install.packages("TSA")
#install.packages("trend")

library(aTSA)
library(TSA)
library(trend)

growth1 <-
c(11,31,11,50,31,34,28,47,25,24,38,38,43,38,27,40,68,44,67,100,44,8
2,54,55,113,67,92,65,84,119,115,145,122,114,146,138,75,125,141,80
,121,119,103,129,122,118,141,137,161,171,187,185,176,187,208,240
,263,195,231,250,275,264,294,300,296,285,244,252,239,239,299,278
,287,285,257,325,295,243,235,218,278,336,304,221,80,239,335,260,
293,250,269,251,220,250,150,243,335,201,194,227,248,199,166,164,
48,169,219,165,156,131,123,49,108,120,154,131,86,99,19,68,117,80)

growth <- growth1

n<- length(growth)

years<- c(1:n)

# ADF test for growth data

adf <- adf.test(growth)

#adf
```

#McLeod and Li test

dev.new()

ML<- McLeod.Li.test(y=growth)

ML

#Mann Kendall test

MK<-mk.test(growth, continuity = TRUE)

MK

#Standard Normal Homogeneity Test (SNHT)

(out <- snh.test(growth))

dev.new()

plot(out)

#Performs a rank version of von Neumann's ratio test

#as proposed by Bartels. The null hypothesis

#of randomness is tested against the alternative hypothesis

bartels.test(growth)

الملاحق

ملحق (2): تتضمن هذه البكجات (تقدير معلمات الأنموذجات اللاخطية)، اختبار t واختبار F واختبار Jarque-Bera، (معايير المعلومات للمفاضلة بين الأنموذجات)، (التنبؤ بهذه الأنموذجات)، (مقاييس الدقة التنبؤية).

```
#install.packages("tseries")

library(FME)

library(tseries)

growth1 <-
c(11,31,11,50,31,34,28,47,25,24,38,38,43,38,27,40,68,44,67,100,44,8
2,54,55,113,67,92,65,84,119,115,145,122,114,146,138,75,125,141,80
,121,119,103,129,122,118,141,137,161,171,187,185,176,187,208,240
,263,195,231,250,275,264,294,300,296,285,244,252,239,239,299,278
,287,285,257,325,295,243,235,218,278,336,304,221,80,239,335,260,
293,250,269,251,220,250,150,243,335,201,194,227,248,199,166,164,
48,169,219,165,156,131,123,49,108,120,154,131,86,99,19,68,117,80)

growth <- growth1

n<- length(growth)

years<- c(1:n)

## ===== ##
## tests criteria ##
## ===== ##

aic.test<- function(RSS, n, k){
```

```
aic <- n*log(RSS/n) + 2 * k
```

```
return(aic)
```

```
}
```

```
bic.test<- function(RSS, n, k){
```

```
bic <- n*log(RSS/n) + k * log(n)
```

```
return(bic)
```

```
}
```

```
HQ.test<- function(RSS, n, k){
```

```
HQ <- n*log(RSS/n) + 2 * k * log(log(n))
```

```
return(HQ)
```

```
}
```

```
## ===== ##
```

```
## Exponential growth model ##
```

```
## ===== ##
```

```
Data.EX <- data.frame(
```

```
  time = years,
```

```
  N = growth
```

```
)
```

```
## initial "guess"
```

```
parms.EX <- c(alpha = 0.06, beta = 0.06)
```

```
## analytical solution
```

```
model.EX <- function(parms,time)
```

```
  with (as.list(parms), return(alpha * exp(-beta * time)))
```

```
## FITTING algorithm 1
```

```
ModelCost.EX <- function(P) {
```

```
  out <- model.EX(P, years)
```

```
  return(Data.EX$N-out) # residuals
```

```
}
```

```
Fita.EX <- modFit(f = ModelCost.EX, p = parms.EX, method =  
"Newton",lower = -0.05,
```

```
control=list(stepol = 1e-20, iterlim = 10000, check.analyticals =  
TRUE))
```

```
times <- 0:n
```

```
dev.new()
```

```
plot(Data.EX[, "time"], Data.EX[, "N"], pch = 16, col = "red",
```

```
main = "exp growth", xlab = "time", ylab = "N")
```

```
lines(times, model.EX(Fita.EX$par, times), lwd = 2, col = "blue")
```

```
summary(Fita.EX)
```

```
mse.EX<- Fita.EX$ssr/Fita.EX$df.residual
```

```
par.EX<- c(Fita.EX$par,0,0)
```

```
RSS.EX <- Fita.EX$ssr
```

```
AIC.EX<- aic.test( RSS.EX , n, k=2)
```

```
BIC.EX<- bic.test( RSS.EX , n, k=2)
```

```
HQ.EX<- HQ.test( RSS.EX , n, k=2)
```

```
JB.EX<-jarque.bera.test(residuals(Fita.EX))
```

```
TSS<- sum(growth^2)
```

```
RSS.EX<- sum(model.EX(Fita.EX$par, times)^2)
```

```
ESS.EX<- sum(Fita.EX$residuals^2)
```

```
F.EX<- (RSS.EX/2)/(ESS.EX/(n-2))
```

```
## ===== ##
## Von Bertalanffy growth model  ##
## ===== ##

Data.Vb <- Data.EX

## initial "guess"

parms.VB <- c(alpha = 0.06, beta = 0.06)

## analytical solution

model.VB <- function(parms,time)

  with (as.list(parms), return(alpha * (1-exp(-beta * time))))

## FITTING algorithm 1

ModelCost.VB <- function(P) {

  out <- model.VB(P, years)

  return(Data.VB$N-out) # residuals

}

Fita.VB <- modFit(f = ModelCost.VB, p = parms.VB, method =
"Newton",lower = -0.5,
```

```
control=list(steptol = 1e-20, iterlim = 10000, check.analyticals =  
TRUE))
```

```
times <- 0:n
```

```
dev.new()
```

```
plot(Data.VB[,"time"], Data.VB[,"N"], pch = 16, col = "red",
```

```
main = "Von Bertalanffy growth", xlab = "time", ylab = "N")
```

```
lines(times, model.VB(Fita.VB$par, times), lwd = 2, col = "blue")
```

```
summary(Fita.VB)
```

```
mse.VB<- Fita.VB$ssr/Fita.VB$df.residual
```

```
par.VB<- c(Fita.VB$par,0,0)
```

```
RSS.VB <- Fita.VB$ssr
```

```
AIC.VB<- aic.test( RSS.VB , n, k=2)
```

```
BIC.VB<- bic.test( RSS.VB , n, k=2)
```

```
HQ.VB<- HQ.test( RSS.VB , n, k=2)
```

```
JB.VB<-jarque.bera.test(residuals(Fita.VB))
```

```
TSS<- sum(growth^2)
```

```
RSS.VB<- sum(model.VB(Fita.VB$par, times)^2)
```

```
ESS.VB<- sum(Fita.VB$residuals^2)
```

```

F.VB<- (RSS.VB/2)/(ESS.VB/(n-2))

## ===== ##
## Gompertz growth model  ##
## ===== ##

Data.GO <- Data.EX

## initial "guess"

parms.GO <- c(alpha = 0.06, beta = 0.06, k = 0.06)

## analytical solution

model.GO <- function(parms,ti)

  with (as.list(parms), return(alpha * exp(-beta * exp(-k*ti))))

## FITTING algorithm 1

ModelCost.GO <- function(P) {

  out <- model.GO(P, years)

  return(Data.GO$N-out) # residuals

}

Fita.GO <- modFit(f = ModelCost.GO, p = parms.GO, method =
"Newton",lower = 0.05,

```

```
control=list(steptol = 1e-20, iterlim = 10000, check.analyticals =  
TRUE))
```

```
times <- 0:n
```

```
dev.new()
```

```
plot(Data.GO["time"], Data.GO["N"], pch = 16, col = "red",
```

```
main = "Gompertz growth", xlab = "time", ylab = "N")
```

```
lines(times, model.GO(Fita.GO$par, times), lwd = 2, col =  
"blue")
```

```
summary(Fita.GO)
```

```
mse.GO<- Fita.GO$ssr/Fita.GO$df.residual
```

```
par.GO<- c(Fita.GO$par,0)
```

```
RSS.GO <- Fita.GO$ssr
```

```
AIC.GO<- aic.test( RSS.GO , n, k=3)
```

```
BIC.GO<- bic.test( RSS.GO , n, k=3)
```

```
HQ.GO<- HQ.test( RSS.GO , n, k=3)
```

```
JB.GO<-jarque.bera.test(residuals(Fita.GO))
```

```
TSS<- sum(growth^2)
```

```
RSS.GO<- sum(model.GO(Fita.GO$par, times)^2)

ESS.GO<- sum(Fita.GO$residuals^2)

F.GO<- (RSS.GO/3)/(ESS.GO/(n-3))

## ===== ##
## Logistic growth model  ##
## ===== ##

Data.LO <- Data.EX

## initial "guess"

parms.LO <- c(alpha = 0.06, beta = 0.06, k = 0.06)

## analytical solution

model.LO <- function(parms,ti)

  with (as.list(parms), return(alpha/(1 + beta * exp(-k * ti))))

## FITTING algorithm 1

ModelCost.LO <- function(P) {

  out <- model.LO(P, years)

  return(Data.LO$N-out) # residuals

}
```

```
Fita.LO <- modFit(f = ModelCost.LO, p = parms.LO, method =  
"Newton", lower = 0.05,  
  
control=list(stepTol = 1e-5, iterLim = 10000, check.analyticals =  
TRUE))  
  
times <- 0:n  
  
dev.new()  
  
plot(Data.LO[, "time"], Data.LO[, "N"], pch = 16, col = "red",  
      main = "Logistic growth", xlab = "time", ylab = "N")  
  
lines(times, model.LO(Fita.LO$par, times), lwd = 2, col = "blue")  
  
summary(Fita.LO)  
  
mse.LO <- Fita.LO$ssr/Fita.LO$df.residual  
  
par.LO <- c(Fita.LO$par, 0)  
  
RSS.LO <- Fita.LO$ssr  
  
AIC.LO <- aic.test( RSS.LO , n, k=3)  
  
BIC.LO <- bic.test( RSS.LO , n, k=3)  
  
HQ.LO <- HQ.test( RSS.LO , n, k=3)  
  
JB.LO <- jarque.bera.test(residuals(Fita.LO))  
  
TSS <- sum(growth^2)
```

```

RSS.LO<- sum(model.LO(Fita.LO$par, times)^2)

ESS.LO<- sum(Fita.LO$residuals^2)

F.LO<- (RSS.LO/3)/(ESS.LO/(n-3))

## ===== ##
## Chapman-Richard growth model ##
## ===== ##

Data.CR <- Data.EX

## initial "guess"

parms.CR <- c(alpha = 3, beta = 0.06, k = 2, theta=0.1)

## analytical solution

model.CR <- function(parms,ti)

  with (as.list(parms), return(alpha * ( 1-(beta * exp(-k *
ti))))^theta))

## FITTING algorithm 1

ModelCost.CR <- function(P) {

  out <- model.CR(P, years)

  return(Data.CR$N-out) # residuals

}

```

```
Fita.CR <- modFit(f = ModelCost.CR, p = parms.CR, method =  
"Newton",lower = 0.05,  
  
control=list(stepcol = 1e-20, iterlim = 10000, check.analyticals =  
TRUE))  
  
times <- 1:n  
  
dev.new()  
  
plot(Data.CR[,"time"], Data.CR[,"N"], pch = 16, col = "red",  
      main = "Chapman-Richard growth", xlab = "time", ylab =  
"N")  
  
lines(times, model.CR(Fita.CR$par, times), lwd = 2, col = "blue")  
  
summary(Fita.CR)  
  
mse.CR<- Fita.CR$ssr/Fita.CR$df.residual  
  
par.CR<- c(Fita.CR$par)  
  
RSS.CR <- Fita.CR$ssr  
  
AIC.CR<- aic.test( RSS.CR , n, k=4)  
  
BIC.CR<- bic.test( RSS.CR , n, k=4)  
  
HQ.CR<- HQ.test( RSS.CR , n, k=4)  
  
JB.CR<-jarque.bera.test(residuals(Fita.CR))
```

```
TSS<- sum(growth^2)

RSS.CR<- sum(model.CR(Fita.CR$par, times)^2)

ESS.CR<- sum(Fita.CR$residuals^2)

F.CR<- (RSS.CR/4)/(ESS.CR/(n-4))

## ===== ##

aic <- cbind(AIC.EX, AIC.VB, AIC.GO, AIC.LO, AIC.CR)

Bic <- cbind(BIC.EX, BIC.VB, BIC.GO, BIC.LO, BIC.CR)

hq <- cbind(HQ.EX, HQ.VB, HQ.GO, HQ.LO, HQ.CR)

par<-cbind(

par.EX,

par.VB,

par.GO,

par.LO,

par.CR)

rownames(par)<- c("alpha", "beta", "k", "theta")

dev.new()

plot(Data.CR[, "time"], Data.VB[, "N"], pch = 16, col = "red",
      main = "growth", xlab = "time", ylab = " ")
```

```
lines(times, model.EX(Fita.EX$par, times), lwd = 2, col = 2)
```

```
lines(times, model.VB(Fita.VB$par, times), lwd = 2, col = 3)
```

```
lines(times, model.GO(Fita.GO$par, times), lwd = 2, col = 4)
```

```
lines(times, model.LO(Fita.LO$par, times), lwd = 2, col = 5)
```

```
lines(times, model.CR(Fita.CR$par, times), lwd = 2, col = 6)
```

```
#== predicted values ==#
```

```
pr=12
```

```
Ftimes<- (n+1):(n+pr)
```

```
pre.EX<- model.EX(Fita.EX$par, Ftimes)
```

```
pre.VB<- model.VB(Fita.VB$par, Ftimes)
```

```
pre.GO<- model.GO(Fita.GO$par, Ftimes)
```

```
pre.LO<- model.LO(Fita.LO$par, Ftimes)
```

```
pre.CR<- model.CR(Fita.CR$par, Ftimes)
```

```
summary(Fita.EX)
```

```
summary(Fita.VB)
```

```
summary(Fita.GO)
```

```
summary(Fita.LO)
```

```
summary(Fita.CR)
```

```
par
```

cbind(mse.EX, mse.VB, mse.GO, mse.LO, mse.CR)

aic

Bic

hq

#== Jarque Bera Test ==#

JB.EX

JB.VB

JB.GO

JB.LO

JB.CR

#== F Test ==#

F.EX

F.VB

F.GO

F.LO

F.CR

#== predicted values ==#

pre.EX

pre.VB

pre.GO

pre.LO

pre.CR

write.csv(model.EX(Fita.EX\$par, times),"C:/Users/DELL/Desktop/2/fit
EX.csv")

write.csv(model.VB(Fita.VB\$par, times),"C:/Users/DELL/Desktop/2/fit
VB.csv")

write.csv(model.GO(Fita.GO\$par, times),"C:/Users/DELL/Desktop/2/fit
GO.csv")

write.csv(model.LO(Fita.LO\$par, times),"C:/Users/DELL/Desktop/2/fit
LO.csv")

write.csv(model.CR(Fita.CR\$par, times),"C:/Users/DELL/Desktop/2/fit
CR.csv")

Abstract

Abstract

The increase in the number of the daily infections with Covid-19 virus around the world, especially in the holy governorate of holy Karbala. Also, the human, material and economic losses where this virus have been left behind, the social and psychological effects for all segments of society was the main motivation for presenting this study. Where this study aims to choose the best model to forecast the number of the daily infections with Covid-19 virus in the city of holy Kerbala among the nonlinear time series models which utilized in this study via using the statistical criteria of (AIC, BIC, H-Q).

This study included two basic parts which are the theoretical and practical part. The most important content of the theoretical part is the characteristics of the time series (stationarity, linearity, homogeneity) and the most important statistical tests that are utilizes in order to test these characteristics. It also includes some nonlinear models of the time series which are the Exponential model, Von Bertalanffy model, Gompertz model, Logistic model and Chapman Richard model. Furthermore, it includes the method of estimating the parameters of these models as well as some tests and statistical criteria to measure the significance of parameters and the quality of the models. As well as the criteria for comparison between the models and measures of the predictive power.

As for the practical part, all of what was mentioned in the theoretical part was applied on a sample of people that infected by Covid-19 virus with the size of 122 patients and for the period from (1-6-2020) to (30-9-2020). The results showed that it is a nonstationary, non-linear and heterogeneous series. Also, the best model is the logistic model via

Abstract

utilizing the statistical information criteria (AIC, BIC, H-Q) which gives a lowest value for these criteria. The results also showed that there is a convergence between the real values and the predictive values of the logistic model. Two criteria are applied to measures the predictive power (MSE, RMSE) where the results showed that the logistic model gives the lowest values for these criteria. Finally, these results indicated the efficiency of the proposed model for predicting the numbers of people that infected by Covid-19 virus.

Keywords: Nonlinear time series, Logistic model, Forecasting, Properties of the time series, Covid-19.

**Republic of Iraq
Ministry of Higher Education
And Scientific Research
University of Kerbala
Faculty of Management
And Economics
Department of Statistics**



**Forecasting the number of infected people with
covid-19 in the Holy Kerbala Governorate
Using nonlinear time series models**

**A thesis submitted to the council of the college of
Administration & Economics\ University of Kerbala as
partial fulfillment of the requirements for the
master's degree of Statistics Sciences**

By

Mohammed Saleh Hashem ALaaraji

Under supervision

Prof. Dr. Jassim Nasir Hussain

2021 AD

1443 AH

Holy Kerbala