



وزارة التعليم العالي والبحث العلمي
جامعة القادسية
كلية الادارة والاقتصاد
قسم الاحصاء
الدراسات العليا

تقدير المعالم لعملية Hawkes مع التطبيق

رسالة قدمها الطالب

أيمن عباس نجم

الى

مجلس كلية الادارة والاقتصاد في جامعة القادسية وهي جزء من
متطلبات نيل درجة الماجستير في علوم الإحصاء

بإشراف

أ. د مهند فائز السعدون

2023م

1444هـ

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

يَرْفَعُ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا

الْعِلْمَ دَرَجَاتٍ وَاللَّهُ بِمَا تَعْمَلُونَ خَبِيرٌ

صدق الله العلي العظيم

(المجادلة الآية 11)

الشكر والتقدير

الحمد لله الذي له الحمدُ أكملهُ، ولهُ الثناءُ أجملهُ، ولهُ العلمُ أحكمهُ، ولهُ القولُ أبلغهُ،
والصلاة والسلام على المبعوثِ هُدىً للعالمين، مُحَمَّدٍ الصادقِ الأمين، وعلى آله
الغُرِّ الميامين.

من دواعي رد الجميل الى اهله، تقديراً وعرفاناً أتقدم بالشكر الى أستاذي الدكتور
(مهند فائز السعدون) لتفضله بالأشراف على رسالتي هذه وللجهود القيمة التي بذلها
وتوصياته القيمة التي كان لها الأثر الكبير في إتمام هذه الرسالة.
وانتقدم بخالص شكري وأمتناني الى أستاذتي الأفاضل، رئيس وأعضاء لجنة المناقشة
لتفضلهم بقبول مناقشة هذه الرسالة.

كما أتقدم بالشكر الجزيل والامتنان الى جميع أستاذتي الأفاضل في قسم الإحصاء
الذين كانوا بحق شمساً من العلم أنارت عقولنا فلهم مني كل الود والاحترام.
كما أرى من الواجب عليّ أن أتقدم بالشكر الى جميع اخوتي وزملائي في السنة
التحضيرية لما جسدوا من روح الأخوة الحقيقية والتعاون اسأل الله ان يحفظهم ويجعل
توفيقيهم سرمدياً الى كل ما فيه خير وصلاح.

الباحث





الى.....

باب مدينة العلم وسراج نور الله في الأرض والسماء ...

سيدي امير المؤمنين علي بن ابي طالب (عليه السلام)

امام العصر والدهر ومنقذ البشرية من الظلم والجور ...

الحجة بن الحسن المنتظر (عجل الله فرجة الشريف)

من أسندُ إليه حين خذلتني قدماي، من انحنى ظهره لأستقيم ...

أبي الحبيب (رحمة الله)

من بها أعلو، وعليها أرتكر، وبدعائها تفرج الهموم، ورضاها السبيل الوحيد الى التوفيق

أمي الغالية (حفظها الله)

الى أسرتي وأخوتي وأصدقائي ...

اليكم جميعا ... أهدي جهدي هذا

أمين

المستخلص:

تناولت هذه الرسالة دراسة عمليات التحفيز الذاتي العشوائية المسماة بعمليات هوكس، إذ غالباً ما تشكل الكثير من الحوادث اثناء حدوثها بيانات عبر الزمن تتمثل بما يسمى بالأحداث العنقودية (Cluster) أي ان عمليات الوصول او حدوث الحدث تتمثل بالعينات العنقودية التي فيها يكون حدوث كل حدث يحفز على حدوث حدث اخر بوتيرة متسارعة يشبه العنقود، أي يمكن القول ان عمليات هوكس هي عمليات عشوائية تعتمد في تحليلها للحوادث على الاثار الناتجة من حدوث حوادث سابقة وهذا ما يسمى بتأثير الحافز الذاتي للحدث.

تم اجراء تجربة محاكاة لدراسة سلوك عملية هوكس مستخدمين خوارزمية لتقدير معالم عملية هوكس الأولى الخوارزمية العددية المباشرة لتعظيم لوغاريتم الإمكان (DNMLL) والخوارزمية الثانية هي تعظيم التوقع التقريبية (AEMA) إذ تم استخدام معيار الجودة التحيز والانحراف المعياري للحكم على اداء خوارزميات التقدير من حيث جودة المقدرات وتبين تفوق خوارزمية (DNMLL) في التقدير فضلاً عن دراسة سلوك بيانات حوادث اعداد المصابين بفايروس COVID-19 في محافظة كربلاء المقدسة إذ قدرنا معالم عملية هوكس بعد ان تم التأكد وباستخدام اختبار كولموكروف-سمرنوف ان هذه البيانات تتبع عملية هوكس وكذلك اختبار استقرارية السلسلة لعملية هوكس.

المحتويات

رقم الصفحة	الموضوع	ت
1-13	الفصل الأول منهجية الرسالة والدراسات السابقة	
1-3	المقدمة	1-1
4	مشكلة الرسالة	1-2
4	هدف الرسالة	1-3
5-13	الاستعراض المرجعي	1-4
14-39	الفصل الثاني	
14	المقدمة	2-1
15-22	تعريف أساسية	2-2
15-16	العملية النقطية	2-2-1
16-17	عملية العد	2-2-2
18	الخاصية الاستقلالية	2-2-3
18	الخاصية الاستقرارية	2-2-4
18-19	التجانس	2-2-5
19-20	عملية بواسون المتجانسة	2-2-6
20-21	عملية بواسون غير المتجانسة	2-2-7
21-22	دالة الكثافة الشرطية	2-2-8
22	عملية ماركوف	2-2-9
22-24	عملية هوكس	2-3
24-25	عملية التحفيز الذاتي	2-3-1
25-29	دالة كيرنل الاسية	2-3-2
29-31	اشتقاق دالة الإمكان لعملية هوكس	2-4
31-36	دالة الإمكان باستخدام الاضمحلال الاسي	2-5
36-37	خوارزميات حل مقدر الإمكان الاعظم	2-6
37	خوارزمية التقدير DNMLL	2-7
38-39	خوارزمية التقدير AEME	2-8
39	اختبار جودة المطابقة لعملية هوكس	2-9

المحتويات

40-50	الفصل الثالث	
40	المقدمة	3-1
40-41	مفهوم المحاكاة	3-2
41-42	بناء نموذج المحاكاة	3-3
42-45	نتائج المحاكاة	3-4
45-50	تحليل البيانات الحقيقية	3-5
51-52	الفصل الرابع	
51-52	الاستنتاجات	4-1
52	التوصيات	4-2
53-56	المصادر	
57	الخلاصة باللغة الإنكليزية	

جدول المختصرات

الاسم باللغة الإنكليزية	الرمز
Expectation-Maximization	EM
Approximation expectation-Maximization algorithm	AEMA
Direct numerical Maximization of log-likelihood	DNMLL
Kolmogorov-Simirnov	K-S
Standard deviation	S.D
Markov Chain Monte Carlo	MCMC

جدول الأشكال

الصفحة	العنوان	رقم الشكل
15	العلاقة بين عملية العد وعملية النقطة المرتبطة بها	2-1
26	دالة Kernel الأسية عبر الزمن	2-2
28	سلوك دالة الكثافة الشرطية عند قيم مختلفة لمعلمها	2-3
43	عدد الحوادث المتولدة من عملية هوكس للمجموعة الأولى	3-1
43	عدد الحوادث المتولدة من عملية هوكس للمجموعة الثانية	3-2
48	عدد الإصابات خلال فترة الدراسة	3-3

الجدول

الصفحة	العنوان	رقم الجدول
42	القيم الافتراضية الابتدائية لمعالم عملية هوكس	3-1
44	القيم المقدرة للتحيز والانحراف المعياري للنماذج المختلفة للتجربة الاولى	3-2
45	القيم المقدرة للتحيز والانحراف المعياري للنماذج المختلفة للتجربة الثانية	3-3
46-48	بيانات وباء COVID-19 تمثل اعداد الاصابات	3-4
49	نتائج تقدير الخوارزميتين AEMA, DNMLL	3-5
50	إحصاءات اختبار K-S	3-6

الفصل الاول

منهجية الرسالة والدراسات السابقة

1-1 المقدمة Introduction

تعد عمليات هوكس (Hawkes processes) بمثابة تعميم لعمليات بواسون المتجانسة (Homogeneous Poisson processes) التي يكون فيها معدل حدوث حدث معين يعتمد على وقت وتاريخ الحدث أي انه يؤخذ بنظر الاعتبار بالتسلسل الزمني لحدوث الحدث. ان عمليات Hawkes تم اقتراحها من قبل الباحث (Hawkes,1971) إذ كانت فكرته هي نمذجة الزلازل الارضية، أي ان الفكرة الأساسية للبحث هي اعتماد بيانات حدوث الزلازل والاستفادة منها في دراسة احتمالية حدوث زلزال اخر. وبعد اقتراح هذا النموذج والذي يعد عملية عشوائية فأن الكثيرين من الباحثين حاولوا استخدام عمليات Hawkes في تطبيقات أخرى، مثل التطبيق في علم الاعصاب، العلوم المالية، العلوم الاجتماعية وعلم الجرائم. قبل استخدام وظهور عمليات Hawkes كنماذج يمكن تطبيقها في البيانات المالية كانت تعد بمثابة عمليات Poisson، ولكن بعد ظهور عمليات Hawkes تم استبدال قفزات بواسون (Poisson jump) بما يسمى عمليات Hawkes.

ان الحوادث التي تمت مشاهدتها او ملاحظتها في كثير من الأحيان غالباً ما تتجمع او تحدث بشكل عنقودي (cluster)، فمثلا حدوث زلزال في منطقة جغرافية معينة يؤدي الى شد او ضغط جيولوجي في المنطقة التي حدث فيها مما يعني انه من المحتمل ان تتبعه هزات ارتدادية (تتابع)، كذلك فأن اندلاع حرب شوارع بين عصابات قد يؤدي الى اندلاع موجة من الاعمال الانتقامية الاجرامية، ان بيع كمية كبيرة من الأسهم في بورصة معينة قد يؤدي الى حدوث اضطراب تجاري، (او ان انهيار الاستثمارات في سوق المال لبنك استثماري معين قد يؤدي الى موجة صدمات تجارية لأسواق المال العالمية).

وبالرجوع الى بحث (Hawkes,1971) تجده بالأساس استخدم عملية نقطية (point process) والمسماة self-exciting اي التحفيز الذاتي ، والتي مؤخرا أطلق عليها اسم (Hawkes process)، أي ان عملية هوكس هي نموذج لعملية نقطية (point process) تستخدم لنمذجة نقاط تمثل حوادث معينة موجودة في خط الاعداد الحقيقي (R)، إذ يعرض هذا النموذج التشتت العالي لهذه النقاط ضمن عمليات بواسون (Poisson process)، ومن هنا يمكن القول بأن عملية Hawkes هي عملية عشوائية (stochastic process) وتعد عمليات نقطية تتمثل خصائصها بانها (self-exciting) وهذا يعني أن كل وصول (arrival) او حدث يزيد من معدل الواصلين او الحوادث بالمستقبل لفترة زمنية معينة. ومن هنا يمكن القول ان عملية Hawkes هي عملية عد (Counting process) تشكل نمودجا لسلسلة من حوادث الوصول (arrivals) خلال الزمن مثل الهزات الأرضية، عنف العصابات الاجرامية، الاوامر التجارية (trade orders) والتخلف عن السداد المصرفي.

يتضح مما ورد سابقا ان عملية Hawkes هي نموذج رياضي للعمليات العشوائية ذاتية التحفيز (self-exciting) أي يكون فيها كل (arrival) يحفز (excites) لحدوث حوادث اخرى للعملية العشوائية، بمعنى ان فرصة الوصول اللاحق تزداد في فترة زمنية معينة بعد حدوث الحدث الأساس (Initial arrival)، وبهذا فهي امتداد لعمليات Poisson التي لا تعتمد على اسس سلاسل ماركوف. إذ توحى بعض البيانات الخاصة بعدد الشركات المتخلفة عن سداد قروضها للمصارف بانها عملية محفزة لذاتها (self-exciting)، بمعنى ان عملية العد التي تمثل العملية العشوائية للشركات المتخلفة عن سداد قروضها في فترة زمنية معينة تزيد من فرص عدم تسديد الشركات الأخرى لقروضها (أي ان الشركات التي لا تسدد القروض تحفز الشركات الأخرى على عدم

التسديد وهذا واضح من خلال البيانات عبر الزمن). فضلاً عن ذلك فإن استخدام عمليات Poisson لنمذجة عدد التبادلات التجارية لسهم معين (arrivals) تكون غير مناسبة تماماً، لأن المشاركين بهذه العملية التجارية يتبعون سياسية سلوك القطيع (herding behaviors) أي كل واحد يتأثر بما يفعله الآخر، وتضمنت هيكل الرسالة أربعة فصول

تضمن الفصل الأول مقدمة عامة عن عمليات هوكس وبعض العمليات الأساسية بوصفها مدخلاً لعمليات هوكس مثل عمليات بواسون والعمليات النقطية، فضلاً عن ذلك تناول هذا البحث الدراسات السابقة لعمليات هوكس ذات العلاقة بموضوع الرسالة التي تتضمن آخر الدراسات حول موضوع تقدير معالم عمليات هوكس.

الفصل الثاني تضمن الجانب النظري لعمليات هوكس واشتمل على تعاريف أساسية في العمليات العشوائية وتعاريف العمليات النقطية وعمليات بواسون، فضلاً عن طرائق تقدير معالم عمليات هوكس.

الفصل الثالث تناول توظيف تجارب محاكاة لدراسة عمليات هوكس من ناحية تقدير معالم هذه العملية العشوائية، كما تناول هذا الفصل دراسة تطبيقية لعمليات هوكس على بيانات فايروس كورونا التي تم تجميعها من موقع وزارة الصحة العراقية لمحافظة كربلاء المقدسة. أما الفصل الرابع فقد تناول أهم الاستنتاجات والتوصيات التي توصل إليها الباحث.

1-2 مشكلة الدراسة Thesis Problem

بينت المعلومات المنشورة بالتقارير الطبية الخاصة لانتشار جائحة كورونا في بداية عام 2020 ان هناك عواقب وخيمة من انتشار هذا الفيروس في حالة عدم وجود تدخلات صحية عامة ذات دلالة، إذ لاحظنا ان هناك الكثير من الإجراءات الوقائية مثل اغلاق المدارس وتوقف بعض الاعمال التجارية، فضلاً عن اجراءات الحضر المنزلي للعامة، إذ اعتمدت الكثير من الدول على النماذج الرياضية والتوقعات من اجل اتخاذ القرار حول متى وكيف تفرض الإجراءات الصحية العامة، وقد لوحظ في السنوات الأخيرة طريقة انتشار وباء Covid-19 الذي يعد بمثابة ظاهرة تنتشر بياناتها (حوادثها) بشكل عينة عنقودية مع الزمن.

ان اتجاه تطور انتشار وباء كورونا بشكل عالٍ التكرار مع الزمن اثار انتباه الكثير من الباحثين للاهتمام بهذا الوباء لأن انتشاره يعد عنقودياً والانتشار العنقودي يتغير خلال دقائق وحتى خلال ثواني أيضاً، ان الباحثين الراغبين بدراسة تطور سلوك بيانات وباء كورونا من خلال دراسة الخاصية العنقودية لانتشار الوباء يطمحون بنمذجة الخاصية العنقودية لعملية الانتشار، لذلك في هذه الرسالة سيتبنى الباحث نماذج هوكس في تتبع سلوك انتشار وباء كورونا.

1-3 هدف الرسالة Thesis Objectives

ان الهدف من هذه الرسالة هو دراسة نوع من العمليات العشوائية المسماة بعمليات هوكس (Hawkes process) التي تعد بمثابة امتداد لعمليات بواسون (Poisson process) لكنها تتميز بخاصية التحفيز الذاتي (Self-excitation)، فضلاً عن ذلك سيتم تقدير معالم عمليات هوكس بخوارزميتين الأولى الخوارزمية العددية المباشرة لتعظيم لوغاريتم الإمكان (DNMLL) والثانية خوارزمية تعظيم التوقع التقريبية (AEMA) وكذلك سيتم دراسة سلوك عملية هوكس من خلال عدة سيناريوهات لأمثلة محاكاة مع اجراء تطبيق عملي على بيانات انتشار وباء Covid-19 في محافظة كربلاء المقدسة لسنة 2021 وللاشهر (نيسان، أيار، حزيران، تموز، اب، أيلول).

1-4 الاستعراض المرجعي Literature Review

بهدف دراسة عمليات Hawkes بشكل دقيق سوف نتطرق الى اهم الدراسات التي تناولت هذا النوع من العمليات العشوائية وفترات تطورها وعلاقتها بالعمليات العشوائية الأخرى، كالعمليات النقطية وعمليات العد، إذ يمكن عد العمليات النقطية نماذج إحصائية تستخدم لوصف أنماط النقاط في فضاء معين، إذ ان النقاط يمكن ان تمثل أوقات حدوث حوادث معينة، مثل النقاط الزمنية التي تمثل ظهور مرض معين عند تحليلها عبر تاريخ هذه الحادثة او مكانها ضمن فضاء وزمان معين كما هو في مركز حدوث الهزة الأرضية إذ يتم تسجيل وقت الحدوث (نقطة الزمن) لحدوث الزلزال ومكان الزلزال.

ان ابسط أنواع العمليات النقطية هي عمليات Poisson التي تحدث فيها النقاط (points) بشكل مستقل الواحدة عن الأخرى وبالتالي يمكن اعتبار ان عمليات Hawkes هي عمليات نقطية يكون فيها ظهور حدث معين (point) يؤثر في احتمالية ظهور النقاط الأخرى (points).

- في عام 1971 قدم الباحث Alan Hawkes بحثاً عن التحليل الطيفي (spectral) للعمليات النقطية ذاتية التحفيز، إذ قدم في هذا البحث نموذجاً رياضياً يدرس فيه أوقات حدوث الزلازل وما يتبعها من هزات ارتدادية ودراسة وقتها من خلال وضع نموذج احتمالي يمكن من خلال دراسة احتمالات الحوادث المستقلة. في هذا البحث عد الباحث ان النموذج المقترح لسلسلة الهزات الأرضية بوصفه عملية عد خلال الزمن، ويقصد بذاتية التحفيز (self-exciting) بان وقوع كل هزة أرضية يحفز عملية حدوث الهزات الأخرى إذ ان احتمال حدوث هذه الهزات الأخرى يتزايد خلال فترة زمنية معينة بافتراض ان حالة هزة أرضية (حدث) قد وقع فعلا (Initial).

- وأيضاً قدم الباحث Hawkes في نفس السنة 1971 بحثاً عن العمليات النقطية ذاتية التحفيز، إذ بين الباحث ان الحوادث المتمثلة بالأمراض المعدية (مثل مرض الحصبة وغيرها) وبين ايضاً ان ظهور حوادث معينة يزيد من احتمالية ظهور حوادث أخرى في المستقبل القريب، فقد افترض الباحث نموذجاً لمثل هذا النوع من العمليات فضلاً عن ذلك تم اشتقاق مقدار رياضي يعبر عن مصفوفات الحوادث (Events matrices) وعدها نقاطاً طيفية (spectral points).

- في عام 1979 قدم الباحث Ozaki بحثاً اقترح فيه استخدام طريقة مقدرات الإمكان الأعظم لتقدير معالم عملية Hawkes من خلال استخدام لوغاريتم دالة الإمكان لنموذج Hawkes مستخدماً متجه gradient ومصفوفة Hessian للوغاريتم دالة الإمكان. في هذا البحث اقترح الباحث خوارزمية لتوليد بيانات عملية Hawkes وهو اول بحث حاول توليد بيانات عملية Hawkes إذ عد هذا البحث اساساً لتطوير دراسات مستقبلية لتوليد بيانات عملية Hawkes.

- في عام 1988 قدم الباحث Ogata بحثاً ناقش فيه نماذج عشوائية محاولاً نمذجة أوقات حدوث الهزات الأرضية، إذ قارن النماذج اعتماداً على بيانات لهزات أرضية حدثت في اليابان للفترة من 1885-1980 مستخدماً طرائق الإمكان (Likelihood)، في هذا البحث تم تناول عمليات Poisson والعمليات النقطية في نمذجة الهزات الأرضية، إذ بين الباحث في هذا البحث ان حدوث الزلزال يؤدي غالباً الى زيادة في التوتر الجيولوجي في طبقات الأرض التي حدث فيها الزلزال ومن المحتمل ان تتبعه هزات ارتدادية. فقد توصل الى هذه النتيجة من خلال التحقق باستخدام مبدئ التنبؤ اعتماداً على النماذج المدروسة.

ملاحظة: قام الباحث بالإشارة لبدایات عمليات هوكس من سنة 1971 ولغاية 1988 وتم الاستعراض للدراسات من سنة 2005 لغاية سنة 2022.

- في عام 2005 قدم الباحثان Moller و Rasmussen بحثاً عن دراسة محاكاة عمليات هوكس، إذ كان هدف البحث هو بناء خوارزمية لتحليل المحاكاة وإيجاد الحل الأمثل لعمليات هوكس. إن تحليل المحاكاة لعمليات هوكس غالباً ما يعاني من بعض الصعوبات الحسابية وهذا ما تمت معالجته في هذا البحث، كذلك فإن البحث طور العمليات الحسابية لعمليات هوكس بوصفها عمليات Poisson العنقودية (Cluster) من خلال استخدام خاصية الاستقلال الشرطي وعمليات التفرع (branching) وتم الحصول على نتائج مثالية من خلال امثلة محاكاة وامثلة تطبيقية.

- في عام 2013 قدم الباحثان Zhao و Dassios بحثاً تناول خوارزمية عددية كفوة لمحاكاة عمليات هوكس بافتراض ان معلمة الكثافة (intensity parameter) تسلك سلوكاً اسياً. إن الطريقة العددية المقترحة يمكن من خلالها توليد بيانات للعمليات النقطية عن طريق توليد عينات لأوقات الوصول البينية بواسطة دوال توزيع تحليلي، فضلاً عن ذلك عد الباحثان ان الطريقة الحسابية هي طريقة مرنة في توليد النقاط (الحوادث) في حالة الاستقرار وعدم الاستقرار في معدل الحدوث، كذلك تمت دراسة الطريقة الحسابية في حالة وجود عمليات Hawkes ذات البعد الواحد (الزمن) وذات البعدين (الزمن والموقع).

- في عام 2015 قدم الباحث Bacry وآخرون بحثاً عن استخدام عمليات Hawkes في الجانب المالي إذ تناول هذا البحث مراجعة عامة حول تطبيق عمليات Hawkes في البيانات المالية لكونها تشكل صنفاً معيناً من العمليات النقطية المتعددة، إذ يتصف هذا النوع من العمليات العشوائية بمرونتها وسهولة تطبيقها، فضلاً عن ذلك أشار البحث الى البحوث التي تناولت عمليات Hawkes وكيف انهم قد شاركوا بنجاح من خلال تناولهم لتقدير التقلبات في بيانات التحويلات المالية وكذلك اجراء تقدير لاستقرار الأسواق المالية، والاستدلال على التنفيذ الأمثل للاستراتيجيات المالية ودراسة المخاطر المالية.

– في عام 2015 قدم الباحث Laub وآخرون تقريراً حول عمليات Hawkes لكونها من العمليات العشوائية ذات الاهتمام الخاص والتي يمكن تطبيقها في مجالات متنوعة مثل الهزات الأرضية والتحليلات المالية وغيرها من المجالات التطبيقية كذلك أشاروا الى ان عمليات Hawkes هي عمليات نقطة تتميز او تتصف بانها ذاتية التحفيز (self-excite) أي ان كل حادثة وصول (arrival) تزيد من معدل وصول الحوادث المستقبلية خلال فترة زمنية معينة، تناول الباحثون بهذا البحث الدراسات المتعلقة بهذا النوع من العمليات العشوائية وتاريخ تطورها.

– في عام 2016 قدم الباحث Simon بحثاً حول عمليات Hawkes واستخدامها في مجال المالية (finance) من خلال اجراء تحليل المحاكاة لهذه العملية، إذ بين الباحث ان عمليات Hawkes هي نماذج حصينة ومرنة لمحاكاة الكثير من العمليات النقطية ذاتية التحفيز فضلاً عن ذلك ناقش هذا البحث تقدير كرنل (Kernel)، تقدير المعالم بطريقة الإمكان الأعظم والعمليات النقطية العنقودية كذلك تناول البحث خوارزميات تنفيذ دراسة المحاكاة.

– في عام 2017 قدم الباحثان Chen و Stindl بحثاً حول امتداد لعمليات هوكس تسمى عمليات هوكس المجددة (Renewal Hawkes process) والتي بين الباحثان ان لها تطبيقات عديدة في الواقع العملي للظواهر التي تتميز بالتحفيز الذاتي (Self-exciting)، إذ بين الباحثان ان هذا النوع من العمليات العشوائية يحتاج في تقدير المعالم الى عمليات حسابية الى دالة الإمكان بافتراض ان الزمن يتبع التوزيع الاسي، إذ تم توظيف خوارزمية تعظيم التوقع (EM) لحساب مقدر الإمكان الأعظم، إذ استخدم الباحثان هذه الخوارزمية بافتراض وجود (Quadratic time) وكانت نتائج المحاكاة وتحليل البيانات من الواقع العملي وبعد اجراء اختبار حسن المطابقة تبين تفوق الخوارزمية في تقدير معالم العملية العشوائية.

– في عام 2018 قدم الباحث Gao وآخرون بحثاً حول عمليات هوكس بوصفها عمليات نقطية ذات تحفيز ذاتي ولها آثار عنقودية في رسم سلوك الظاهرة حيث تم افتراض أن معلمة الكثافة (Intensity parameter) تتصف بأنها (time-dependent) وأن دالة التحفيز هي دالة عامة وأن حجوم القفزات لعملية الكثافة (intensity) هي بمثابة متغير عشوائي غير سالب متماثل التوزيع ومستقل، وبين الباحثون أن هذا النوع من العمليات العشوائية لا تتصف بأنها عمليات تتبع خصائص سلاسل ماركوف، وقد تم إجراء توظيف لهذا النوع من العمليات على بيانات حقيقة تمثل أعداد الواصلين لمسيح معين ودراسة الخصائص المهمة لهذه الظاهرة.

– في عام 2019 قدم الباحث Magris بحثاً عن دراسة لمحاكاة عمليات Hawkes من خلال استخدام دوال (Lambert-w) إذ استخدم الباحث الطريقة القديمة المستخدمة لمحاكاة عمليات Hawkes المسماة بطريقة تحويل عكس المعاينة (inverse Sampling transform) ومناقشتها بشكل ملائم من خلال توظيف دوال (Lambert-w) بوصفها طريقة بديلة فعالة، كذلك اتضح من خلال تنفيذ خوارزميات المحاكاة أن الطريقة المقترحة تعطي حلاً أمثلاً وكفؤاً وبشكل عام يكون أداؤها أفضل من الطرائق المتوفرة الأخرى.

– في عام 2019 قدم الباحث Landstrom رسالة ماجستير عن عمليات Hawkes بوصفها نموذجاً تحفيزياً ذاتياً لعمليات بواسون تدعى (Poisson shot noise) إذ تم استعمال عمليات (Ornstein-Uhlenbeck) المشتقة من خلال عمليات Poisson، فضلاً عن استخدام عمليات العد ذات التحفيز الذاتي المسماة (Shot process).

– في عام 2020 قدم الباحثان Short وOlinde بحثاً عن عمليات Hawkes ذاتية التحفيز مستهدفين تفسير نموذج Hawkes وعملية تقدير معالم النموذج من خلال تطبيق النموذج على بيانات الجريمة.

- في عام 2020 قدم الباحث Hawkes بحثاً حول عمليات هوكس بافتراض ان القفزات الحاصلة عبر الزمن للحوادث لها تقلبات عشوائية وقد تم توظيف هذه الفكرة على البيانات المالية، إذ تناول البحث دراسة مفهوم التقلبات (Volatility) في البيانات المالية ولا سيما تقلبات القفزات التي تفترض بانها تتبع عمليات بواسون او عمليات Levy، كذلك تناول هذا البحث مناقشة التطبيقات لعمليات التحفيز الذاتي وعمليات التحفيز الذاتي المستقلة.

- في عام 2021 قدم الباحث Wehrli وآخرون بحثاً حول استخدام عمليات Hawkes لنمذجة تغيرات الأسعار، إذ تناول البحث تغير أسعار العملات كاليورو والدولار الأمريكي المتداولة إلكترونياً للبنوك من خلال خدمات الوساطة فضلاً عن تغير العملات في مؤشر (S&P500) في بورصة شيكاغو إذ تناول البحث أيضاً استخدام العمليات النقطية باستخدام تقديرات خوارزمية (EM).

- في عام 2021 قدم الباحث Swishchuk وآخرون بحثاً عن استخدام عمليات Hawkes في بيانات التأمين إذ تم استخدام عمليات Hawkes المركبة (Compound) بوصفه نموذج مخاطرة وتبين ان هذه العمليات تلائم نمذجة بيانات التأمين على المخاطرة. تم الاعتماد على نظرية الغاية المركزية وقانون الاعداد الكبيرة من اجل اشتقاق صيغة تقريبية للتقلبات الحاصلة بالبيانات، إذ تم استخدام هذا التقريب لدراسة تأثير تبديل عمليات الوصول ل (Poisson) الكلاسيكية بعمليات Hawkes المركبة في دراسة استراتيجيات الاستثمار المثلى.

- في عام 2021 قدم الباحث Unwin وآخرون بحثاً عن استخدام عمليات Hawkes لنمذجة انتشار عدوى الامراض من اجل متابعة التحولات التي تحصل بانتشار هذه الأوبئة، إذ تم توظيف عمليات Hawkes لدراسة سلوك انتشار مرض الملاريا والتنبؤ بتحول المرض وسيناريوهات انتشاره في بقعة معينة من دولة الصين، إذ تم تطوير نموذج باستخدام عمليات Hawkes لبيانات مرض الملاريا وانتشاره وتم تقييم عمل هذا النموذج باستخدام المحاكاة، وبعد ذلك تم تطبيق النموذج

المقترح على بيانات حقيقية، إذ أظهرت النتائج ان النموذج المقترح هو مقياس جيد لحسن المطابقة والتنبؤ لمدة (35) يوماً بالمستقبل.

- في عام 2021 قدم الباحث Bonnet وآخرون بحثاً عن تقدير معالم عمليات Hawkes باستخدام طريقة الإمكان الأعظم، في هذا البحث عرض نظرية الإمكان الأعظم بوصفها طريقة لتقدير معالم عمليات Hawkes ذات المتغير الواحد (Univariate)، إذ تم تطبيق طريقة الإمكان الأعظم لأن عملية Hawkes هي عملية تحفيز ذاتي، فقد تم تطبيق مقدر الإمكان الأعظم المقترح لأن سلوك الحوادث هو بمعدل يتبع دالة (exponential kernel). أظهرت النتائج ان المقدر المقترح هو مقدر دقيق ويتفوق على الطرائق الموجودة البديلة.

- في عام 2021 قدم الباحث Li رسالة ماجستير حول تحليل عمليات هوكس للسلوك الانتحاري بالاعتماد على بيانات متوفرة في مقاطعة Wales في إنكلترا، إذ عد الباحث انه اول من قام بتطبيق عمليات هوكس في مثل هذا النوع من التطبيقات، وتناول الباحث دراسة سلوك ظاهرة الانتحار بهدف التعرف اكثر على الطبيعة التي تتكاثر بها هذه الظاهرة العنقودية إذ عد ان ظاهرة الانتحار ظاهرة يقلد بها الافراد بعضهم بعضاً، إذ ركز الباحث على التنبؤ من خلال استخدام دوال Kernel لتمثيل لقانون القوى power-law بوصفها دالة تحفيز، كذلك بين الباحث ان الدالة Power-law افضل من دالة exponential kernel.

- في عام 2022 تناول الباحث Chiang وآخرون نمذجة بيانات وباء كورونا-19 من خلال عمليات Hawkes، إذ أشار الباحثون الى ان عمليات Hawkes يمكن استخدامها لنمذجة بيانات الأوبئة والتي غالباً ما تكون حوادثها (بياناتها) تتجمع بشكل عنقودي (Cluster)، كذلك حاول الباحثون تطوير نماذج لتمثيل انتقال بيانات كورونا بالاعتماد على زمان ومكان ونوع الإصابات، إذ تمت نمذجة حالات الوفاة وحالات الإصابة بهذا الوباء في الولايات المتحدة اعتماداً على بيانات

متوفرة من موقع كوكل، فقد تم تقدير اعداد الانتشار الداينمي للإصابات باستخدام خوارزمية توقع والتعظيم (EM) لنموذج الانحدار، وظهرت النتائج ان عمليات Hawkes تفوقت على عدة طرائق تستخدم حالياً لتتبع انتشار الوباء.

- في عام 2022 قدم الباحث Lesage وآخرون بحثاً عن عمليات Hawkes إذ تم فيها استخدام دالة كثافة كما بوصفها دالة تحفيز ذاتي وتم تطبيق النموذج المقترح على بيانات التأمين على الكوارث الطبيعية إذ بين الباحثون ان معظم الدراسات المتعلقة بعمليات Hawkes تعاني من نوعين من العيوب عند تطبيقها على بيانات التأمين، العيب الأول هو ان عمليات Hawkes تستخدم صيغة الاضمحلال الاسي (exponential decaying) لوصف سلوك التحفيز، مما يعني انها لا تسمح لحدوث تأخير (delay) بين ظهور حادثة معينة وأثارها في الحوادث المتعاقبة مما يعني أنها لا تمثل بيانات التأمين افضل تمثيل، والعيب الثاني هو ان الجانب النظري لعمليات Hawkes وتطویر نموذج صالح للاستعمال فقط عندما يكون زمن المشاهدات (الحوادث) تؤول الى اللانهاية (Infinity)، بينما الأفق الزمني (التسلسل الزمني) لبيانات التأمين غالباً ما تكون لعدة اشهر او عدة سنوات، ففي هذا البحث تم استخدام دالة تحفيز تتبع دالة كثافة كما بمعنى ان عملية تقدير المعالم واجراء المحاكاة وحسن مطابقة النموذج جميعها تم بالاعتماد على الدالة المقترحة بدلا من دالة الاضمحلال الاسي، إذ تم استخدام بيانات التأمين لمدينة لوكسمبورغ.

- في عام 2022 استخدم الباحث Wang H نظريات الغاية المركزية (central limit theorem) لدراسة سلوك عمليات Hawkes ذات الزمن المتقطع، إذ أشار الباحث الى ان عمليات Hawkes هي عمليات ذات زمن مستمر ولا سيما في التطبيقات البيئية الاجرامية والمالية، كما أشار الباحث الى ان الكثير من التطبيقات هي ذات ازمان متقطعة وبالتالي ومن اجل نمذجة العمليات الزمنية فإنه غالباً ما يكون استخدام عمليات Hawkes ذات الزمن المتقطع مرغوباً بها.

وما يخص هذه الدراسة فقد اعتمد الباحث على بيانات المصابين بمرض Covid-19 في محافظة كربلاء المقدسة والمنشورة في الموقع الرسمي لوزارة الصحة العراقية، إذ تمت دراسة سلوك ونمذجة انتشار هذا الوباء باستخدام عمليات Hawkes، ويعتقد الباحث ان هذه الدراسة هي الأولى على مستوى الدراسات العليا في الجامعات العراقية، وتعد إضافة علمية لمكتبات جامعاتنا العراقية والعربية بوصفها مصدراً باللغة العربية تناول هذا النوع المهم من العمليات العشوائية، وبالتالي تعد البذرة الأولى في هذا المجال للدارسين في مجال العمليات العشوائية للاستفادة منها في المستقبل لتطوير هذا النوع من العمليات العشوائية سواءً من ناحية الجانب النظري او الجانب التطبيقي.

الفصل الثاني

الجانب النظري

2-1 المقدمة :

يتطلب فهم موضوع عمليات Hawkes بشكله الدقيق استعراض مفهوم عمليات العد Counting Process لأن عمليات Hawkes تعد بمثابة حالة خاصة من عمليات العد، أيضا ومن اجل فهم عمليات العد بشكل وافٍ فإننا يجب ان نستعرض مفهوم وتعريف العمليات النقطية Point Process، هذا يعني ان عمليات العد والعمليات النقطية هي من اساسيات العمل بعمليات Hawkes، إذ ان عمليات النقطة هي مجموعة من النقاط العشوائية تقع في فضاء معين مثل الوقت والمكان.

عمليات النقطة توفر نماذج إحصائية لوصف توقيت وخصائص الحوادث، ففي أسواق المال يمكن ان تكون النقطة تمثل صفقة شراء او صفة بيع والتي غالباً ما تؤثر بأسعار الصفقات المستقبلية واحجام التعاملات، كذلك في علم البيئية يتم تسجيل الحوادث على شكل مجموعة من النقاط تمثل مكان ملاحظة ظاهرة معينة، أيضا في وسائل التواصل الاجتماعي عبر الانترنت يمكن ان تكون الاحداث إجراءات المستخدم (user actions) بمرور الزمن إذ لكل اجراء خصائص معينة مثل تأثير المستخدم، موضوع الاهتمام، والاتصال بالشبكة المحلية، أيضا في علم الجيوفيزياء قد تكون النقطة هي حدث زلزال من خلاله يمكن الاستدلال على احتمال وقوع زلزال اخر في المنطقة المجاورة في المستقبل القريب.

عملية النقطة تتوضح على خط الاعداد الحقيقية غير السالبة، إذ يمثل خط الاعداد غير السالبة الزمن، وهي عملية عشوائية تتكون من أوقات حدث معين T_1, T_2, \dots التي تقع على خط الاعداد.

2-2 تعاريف أساسية :

في هذه البحث سنتطرق الى تعاريف وخصائص كل من عمليات العد والعمليات النقطية والتي ستقودنا الى تعريف عمليات Hawkes.

تعريف 2-2-1: العملية النقطية Point Process (Obra1,2016)

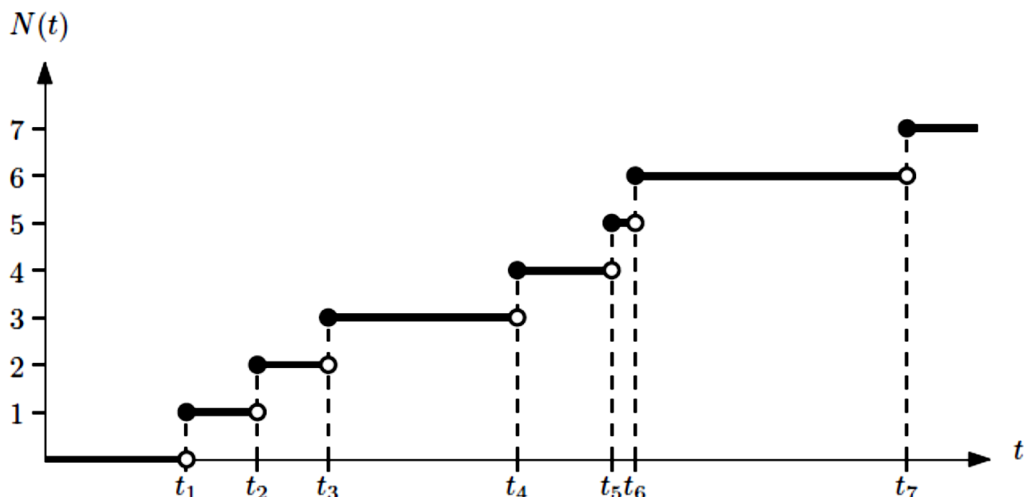
تسمى سلسلة المتغيرات العشوائية غير السالبة $T = \{T_1, T_2, \dots\}$ التي تمثل أوقات حدوث حدث معين (أوقات الوصول العشوائي) بحيث ان $T_i < T_{i+1}$ بالعملية النقطية البسيطة (simple) إذا تحققت البديهيات الاتية:

$$(1) \text{ ان قيم السلسلة ضمن } \mathbb{R}^+ \cup \{\infty\}$$

$$(2) \text{ ان } \Pr(0 < T_0 \leq T_1 \leq T_2 \leq \dots) = 1 \text{ وان}$$

$$\Pr(T_i < T_{i+1}, T_i < \infty) = \Pr(T_i < \infty) \text{ لكل } i \geq 1$$

(3) من المؤكد تقريبا (almost surely) ان عدد النقاط (points) في منطقة محدودة (bounded region) هو عدد محدود (finite).



شكل (2-1) يوضح مثالا بين عملية النقطة (t_1, t_2, \dots) وعملية العد $N(t)$ المرتبطة بها.

(Laub et al., 2015).

نلاحظ من الشكل أعلاه ان عملية العد $N(t)$ مستمرة لليمين وانه عند الزمن $t = 0$ نجد ان

$$N(0) = 0 \text{ وأيضا ان القفزات في عملية العد أعلاه حدثت في } T = \{1,2,3,4,5,6,7\}.$$

واعتمادا على تعريف العملية النقطية في التعريف (1) الذي يستند إلى أوقات الوصول البينية

يمكن تعريف عملية العد كما يأتي:

تعريف 2-2-2: عملية العد Counting process

(Laub et al.,2015) (Obral,2016)

تعد عملية العد بمثابة عملية عشوائية $\{N(t); t \geq 0\}$ وتعرف قيمها بالمجموعة N_0 بحيث ان

$N(0) = 0$ وهي مؤكده تقريبا بانها (almost surely) كذلك فان عملية العد هي عملية محدودة

(finite) ومستمرة من جهة اليمين بخطوات ذات قفزات بحجم +1، او يمكن تعريفها على انها

تمثل العدد الكلي للحوادث التي تحدث بمرور الزمن t .

ان عملية العد $N(t)$ هي عملية احتمالية عشوائية تعتمد على الزمن وتعرف على $t \geq 0$ وتأخذ

قيماً صحيحة $\{1,2,\dots\}$ ، القيمة فيها تمثل عدد الحوادث في عملية النقطة بالزمن t ، لذلك فأنها

تحدد من خلال سلسلة من المتغيرات العشوائية غير السالبة والتي تحقق الشرط $T_i < T_{i+1}$ اذا

كان $T_i \leq \infty$. بمعنى اخر ان $N(t)$ هي عملية عد للحوادث لحين الزمن t ، أي ان: $N(t) =$

$\sum_{i \geq 1} I \{t \leq T_i\}$ إذ يمكن ان نلاحظ ان $N(t)$ هي ثابت ثنائي (piecewise) وبممتلك قفزة

بمقدار +1 عند زمن الحدث T_i ، وبهذا يمكن القول ان مجموعة الأوقات T_1, T_2, \dots وما تقابلها

من عملية العد هي مكافاة لعملية النقطة.

لنفرض ان لدينا العملية العشوائية $\{T_i ; i \in \mathbb{N}\}$ والتي تمثل عملية النقطة، فان العملية العشوائية التالية ذات القفزات المستمرة من اليمين تدعى بعملية العد المرافقة للعملية $\{T_i ; i \in \mathbb{N}\}$:

$$N(t) = \sum_{i \in \mathbb{N}} I_{\{T_i \leq t\}}$$

إذ ان $I_{\{T_i \leq t\}}$ هو دالة المؤشر T_i :

$$I_{\{T_i \leq t\}} = \begin{cases} 1 & \text{if } T_i \leq t \\ 0 & \text{if } T_i > t \end{cases}$$

وتجدر الإشارة الى ان مشتقة عملية العد يمكن كتابتها على انها مجموع لدوال دلتا ورياضيا تكتب بالشكل الاتي:

$$dN(t) = \sum_{t_i=t} \delta(t - t_i)$$

حيث ان دالة دلتا $\delta(\cdot)$ هي مشتقة عملية العد $N(t)$ اذ ان عملية العد تعتبر بمثابة عملية قفز وعادة ما تعطي قيمة ثابتة (Constant) لكل قيم T فان دالة دلتا تتلاشى ولا يكون لها وجود في أي نقطة في الزمن.

لكن دالة القفز هي دالة غير مستمرة عند $t=0$ وهذا يعني ان مشتقتها غير محددة عند هذه النقطة لذلك فان:

$$\delta(\cdot) = \frac{d}{dt} N(t) = \begin{cases} 0 & \text{if } t \neq 0 \\ \infty & \text{if } t = 0 \end{cases}$$

وغالبا ما يتم استخدام دالة δ ضمن العمليات الحسابية للتكاملات أي ان

$$\int_{-\infty}^{\infty} f(t) \delta(t) dt = 0$$

وان

$$\int_{-\infty}^{\infty} \delta(t) dt = 1$$

التعاريف الآتية تبين اهم الخصائص لعملية العد.

تعريف 2-2-3: الخاصية الاستقلالية Independence (Beichelt,2018)

يقال ان عملية العد تمتلك زيادات مستقلة (independent increments) إذا كان لدينا $t_0, t_1, t_2, \dots, t_n$ بحيث $t_0 < t_1 < t_2 < \dots < t_n$ وان المتغيرات العشوائية $N_{t_0}, N_{t_1}, -N_{t_0}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ مستقلة او مكافئة الى $N_{t+T} - N_T$ تكون مستقلة عن N_s لكل $s < T$ وبالتالي فان العملية $[N_t; t \in T]$.

تعريف 2-2-4: الخاصية الاستقرارية stationarity (Beichelt,2018)

يقال ان عملية العد $\{N(t); t \in T\}$ تمتلك زيادات مستقرة (Stationary increments) إذا كان $t_1, t_2 \in T$ فان الزيادات $N(t_1 + T), N(t_2 + T)$ لها نفس التوزيع الاحتمالي ولكل قيم T , حيث ان $t_1 + T \in T$ و $t_2 + T \in T$ وان t_2, t_1 هي قيم ثابتة fixed واعتباطية arbitrary.

تعريف 2-2-5: التجانس Homogeneity (Obra,2016)

يقال ان العملية العشوائية بانها متجانسة في الفضاء اذا كان الاحتمال الانتقالي بين أي قيمتين للحالة في وقتين معينين يعتمد فقط على الفرق بين قيم تلك الحالة، كذلك فان العملية العشوائية يقال بانها متجانسة في الوقت (time) اذا كان احتمال الانتقال بين أي قيمتين لحالة معينة عند أي وقتين (زمنين) يعتمد فقط على الفرق بين تلك الأوقات.

نتيجة : Corollary (Obrol,2016)

من التعريف رقم (2-2-2) نجد ان عملية العد $N(t)$ يجب ان تحقق البديهيات الاتية:

$$N(t) \geq 0 \quad (1)$$

$$N(t) \text{ هي قيمة ذات عدد صحيح} \quad (2)$$

$$N(s) \leq N(t) \text{ فان } s < t \quad (3)$$

$$(4) \text{ لكل } s < t \text{ فان } N(t) - N(s) \text{ تساوي عدد الحوادث التي تحدث في الفترة } (s, t).$$

تعريف 2-2-6: عملية بواسون المتجانسة Homogeneous Poisson

Process (Rizoiu et al.,2017)(Obrol,2016)

ان عملية العد $\{N(t), t \geq 0\}$ يقال بانها عملية بواسون المتجانسة بمعدل $\lambda > 0$ اذا تحققت

البديهيات الاتية:

$$N(0) = 0 \quad (1)$$

$$(2) \text{ ان العملية } \{N(t), t \geq 0\} \text{ يمتلك زيادات مستقلة}$$

$$P[N(t+h) - N(t) \geq 2] = 0(h) \quad (3)$$

$$P[N(t+h) - N(t) = 1] = \lambda h + 0(h) \quad (4)$$

إذ ان h يمثل طول فترة زمنية معينة تكاد تكون قريبة للصفر وان الدالة $f(h)$ يقال بانها $0(h)$

$$\lim_{h \rightarrow 0} \frac{f(h)}{h} = 0 \quad \text{إذا}$$

افرض ان $\{t_i, i \geq 1\}$ هي سلسلة من المتغيرات العشوائية المستقلة والتي لها نفس التوزيع الاسي

بالمعلمة λ وان أوقات الحدث هي:

$$T_n = \sum_{i=1}^{N(t)} T_i$$

ان العملية $\{N(t), t \geq 0\}$ والمعروفة على

$$N(t) = \sum_{i \geq 1} I_{\{t \geq T_i\}}$$

تدعى بعملية Poisson بمعدل كثافة λ وتعد هذه العملية ابسط فئة او صنف من عمليات النقطة. إذ ان T_i تسمى اوقات الوصول البينية. وتجدر الاشارة الى ان اوقات الوصول البينية للحوادث تتبع التوزيع الاسي بالدالة الاتية:

$$f(t) = \lambda e^{-\lambda t}; t \geq 0$$

$$E(t) = \frac{1}{\lambda} \quad \text{إذ ان}$$

هذا يعني ان الحوادث تصل بالمتوسط بمعدل مساوٍ لـ λ خلال وحدة الزمن، إذ ان الوقت المتوقع بين اوقات الحدث هو $\frac{1}{\lambda}$ ، وبهذا فان عملية Poisson تمتلك كثافة λ . بشكل عام فان الكثافة λ (متوسط اوقات الوصول البينية) ليس بالضرورة لان تكون ثابتة فقد تكون دالة بالزمن وتكتب $\lambda(t)$ وهذا ما يسمى بعملية Poisson غير المتجانسة.

تعريف 2-2-7: عملية بواسون غير متجانسة Nonhomogeneous Poisson

(Laub et al.,2015) Process

يقال ان عملية العد $\{N(t); t \geq 0\}$ بانها عملية بواسون غير المتجانسة بدالة كثافة (Intensity function) $\lambda(t)$ إذ ان $t \geq 0$ إذا تحققت البديهيات الاتية:

$$N(0) = 0 \quad (1)$$

(2) ان العملية $\{N(t), t \geq 0\}$ يمتلك زيادات مستقلة

$$\Pr [N(t+h)-N(t) \geq 2] = 0(h) \quad (3)$$

$$\Pr [N(t+h)-N(t) = 1] = \lambda(t)h+0(h) \quad (4)$$

إذ ان دالة الكثافة $\lambda(t)$ لعملية بواسون غير المتجانسة تعد بمثابة دالة ثابتة (deterministic)،

$$\lambda = R^+ \rightarrow R^+$$

كذلك إذا كانت $\lambda(t) = \lambda$ لكل $t \geq 0$ فان العملية $N(t)$ تصبح عملية بواسون المتجانسة.

تعد عملية بواسون غير المتجانسة ذات أهمية خاصة لدراسة عمليات Hawkes.

تعريف 2-2-8: دالة الكثافة الشرطية Conditional intensity function

(Rodríguez,2019) (Laub et al.,2015)

لنفرض ان لدينا عملية العد $N(.)$ وافرض ان $H(.)$ هي عملية عشوائية للبيانات

التاريخية (History) للعملية $N(.)$ فاذا كان $\lambda^*(t)$ موجودة ومعرفة بالآتي:

$$\lambda^*(t) = \lim_{h \rightarrow 0} \frac{E[N(t+h)-N(t)|H(t)]}{h}$$

ويمكن أيضا كتابة $\lambda^*(t)$ بالشكل الآتي:

$$\lambda^*[t|H(t)] = \lim_{h \rightarrow 0} \frac{1}{h} P[N(t+h) - N(t) > 0|H(t)]$$

فان $\lambda^*[t|H(t)]$ تدعى دالة شرطية لـ $N(.)$ وأن $\lambda^*(t)$ تمثل المعدل المتوقع للواصلين

(arrivals expected rate) اعتمادا على قيمة $H(t)$.

ان دالة الكثافة الشرطية تعد أحد اهم الخصائص المميزة للعمليات النقطية والتي يمكن عددها بالأصل على انها دالة المخاطرة (Hazard function)، فاذا كان حدوث حدث معين (وصول) يسبب زيادة في دالة الكثافة الشرطية عندئذ فان العملية العشوائية تدعى التحفيز الذاتي (self-exciting)، وبالتالي يتسبب هذا السلوك للحوادث المتتالية (أوقات الحدوث) بما يسمى بالتجمع العنقودي (clustering) المؤقت للأوقات T .

تعريف 2-2-9: عملية ماركوف Markov Process

(Engelke et al,2018)

العمليات العشوائية $\{X(t); t \in R\}$ التي تمثل مجموعة من المتغيرات العشوائية تدعى عملية ماركوف اذا تحقق الشرطان الآتيان:

1- خاصية ماركوف

$$p[X(t+s) = j | X(u) ; 0 \leq u \leq t] = p[X(t+s) = j | X(t)]$$

لكل قيم $s, t \geq 0$

2- التجانس

$$P[X(t+s) = j | X(t) = i] = P[X(s) = j | X(0) = i]$$

لكل قيم $s, t \geq 0$ وان $i, j \in E$

حيث ان E هي مجموعة قابلة للعد (Countable set).

2-3 عملية هوكس Hawkes proces (Laub et al,2015,Wang,2015)

تعد عمليات النقطة من المواضيع الإحصائية المهمة التي جلبت انتباه الكثير من الباحثين، إذ عمل العالم (Cox,1955) على ما يسمى بعمليات بواسون العشوائية المضاعفة (double) او ما

تسمى بعمليات Cox، وكذلك طور العالم (Bartlett,1963,1964) طرقاً إحصائية في العمليات النقطية.

قدم الباحث (Lewis,1964) بعض مفاهيم عمليات Hawkes من خلال العمليات النقطية، ومن هنا يمكن القول اننا مهتمون بعمليات العد $\{N(t) ; t \geq 0\}$ التي تكون فيها قيمة دالة الكثافة عند الزمن t غير ثابتة (not fixed)، بل تعتمد على بعض المدخلات العشوائية (random inputs)، بضمنها تاريخ العملية العشوائية (أي الحوادث التي تحدث قبل حدوث الحالة الحالية أي جميع الحوادث السابقة).

وبافتراض ان $H(t)$ تمثل تاريخ العملية (History of process) لحين الزمن t ، ومن المفترض ان $H(t)$ تلخص كل ما يدخل في تكوين دالة الكثافة، فقد أشار الباحث (Obrol,2016) الى ان عملية Hawkes هي عملية تحفيز ذاتي تكون فيها دالة التحفيز تتبع احدى دوال الاضمحلال الاسي (exponential decay) او تسمى (exponential kernel)، وبهذا يمكن القول ان عملية هوكس أحادية المتغير (univariate) هي تعميم لعملية Poisson المتجانسة، إذ ان معدل حدوث الحوادث في عملية بواسون يعتمد على الوقت وتاريخه (history)، كذلك فان دالة الكثافة تعد بمثابة عملية عشوائية عند دراسة عمليات التحفيز الذاتي، وتكون كل قفزة (Jump) قد حدثت مسبقاً تزيد من دالة الكثافة (او تسمى دالة القفزة)، هذا يظهر الاعتمادية على الماضي إذ يختلف هذا عن عمليات بواسون المتجانسة وغير المتجانسة كونها تعتمد على توزيعات تسمى توزيعات عديمة الذاكرة (Memoryless). كذلك فان القفزات اللحظية (instantaneous jumps) غير محتملة المشاهدة وهذا واضح من خلال تعريف عمليات بواسون المتجانسة، طالما ان القفزة الجديدة

تعتمد على الفقرة الأخيرة وان توزيع القفزات تم تحديده بشكل تام من خلال دالة الكثافة الشرطية فان الزوج المرتب $(N(t), \lambda^*(t))$ يشكل عملية ماركوف (Markov process).

افتراض ان $\{N(t), t \in R\}$ هي عملية عد وترافقها العملية $\{H(t), t \in R\}$ وتحقق الشروط الاتية:

$$P[N(t+h) - N(t) = m | H(t)] = \begin{cases} \lambda^*[t|H(t)]h + 0(h), & m = 1 \\ 0(h), & m > 1 \\ 1 - \lambda^*[t|H(t)]h + 0(h), & m = 0 \end{cases}$$

ولهذا فان العملية $N(t)$ تسمى عملية Hawkes.

بشكل خاص تطرق العالم (Hawkes,1971) الى نوعين من دوال التحفيز، الأولى تدعى الاضمحلال الاسي (exponential decay) والدالة الثانية تدعى الاضمحلال المحتمل (Potential decay).

في هذه الرسالة سوف يفترض الباحث ان دالة التحفيز هي من نوع دالة (exponential decay).

تعريف 2-3-1: عملية التحفيز الذاتي Self-exciting process

(Obrol,2016)

ان العملية $\{N(t), t \geq 0\}$ تسمى عملية تحفيز ذاتي اذا تحقق الشرط الاتي:

$$\lambda^*[t|H(t)] = \lambda_0(t) + \int_0^t \mu(t-s) dN(s) \quad (2-1)$$

$$= \lambda_0(t) + \sum_{T_i \leq t} \mu(t - T_i)$$

إذ ان معرفة بالآتي: $\lambda_0(t) : R \rightarrow R^+$ وهي دالة ثابتة (deterministic) تسمى دالة الكثافة الأساسية (base) وان الدالة $\mu(.) : R^+ \rightarrow R^+$ هي دالة Kernel وتمثل التأثير الإيجابي للحوادث السابقة T_i على القيمة الحالية لعملية الكثافة وتسمى أيضا بدالة التحفيز (exciting)، وتتبع احدى دوال الاضمحلال الاسي (exponential decay) او (exponential kernel).

تعريف 2-3-2: دالة كيرنل الاسية Exponential Kernel (Haghdan,2017)

تسمى أيضا دالة كيرنل لقانون القوى (Kernel Power law)، إذ تعد الدالة الاسية من أكثر دوال Kernel استخداما، إذ ان اول من استخدمها هو العالم (Hawkes,1971) وهي تملك الصيغة الآتية:

$$g(t) = \alpha e^{-\beta t} ; \quad \alpha > 0, \beta > 0 \quad (2-2)$$

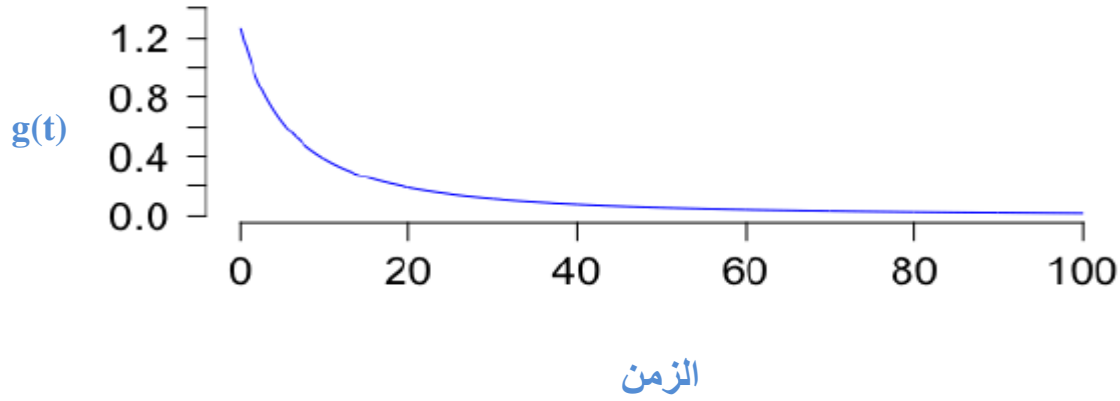
إذ ان:

α تمثل حجم قفزات التحفيز الذاتي (Size of self-excited jumps)

β هي معدل الانحلال الاسي (exponential decay rate).

ان المعادلة (2-2) تدعى دالة Kernel الاسية وكذلك تدعى دالة الانحلال (الاضمحلال) الاسي (exponential decay) طالما ان تأثير ظهور الحالات السابقة (التاريخية) على دالة الكثافة الشرطية هو تأثير يتبع الانحلال الاسي بين أي حدثين متعاقبين. ان دالة كيرنل الاسية تعبر عن التأثير الايجابي للحوادث السابقة على القيمة الحالية لدالة الكثافة.

الرسم الآتي يوضح دالة Kernel الاسية عند قيم محددة لمعاملها.



الشكل (2-2) يوضح دالة Kernel الاسية عبر الزمن (Haghdan,2017)

ولكي نجعل عملية Hawkes مستقرة (stationary) فان متوسط أوقات الوصول البينية (interarrival) يجب ان يكون ثابتا خلال الزمن. ان متوسط وقت الوصول البيني يحسب من خلال $1/\mu$ ، إذ ان μ هو متوسط الكثافة (Intensity mean) ويحسب بالشكل الآتي:

$$\mu = E[\lambda(t)] = E \left[\lambda_0 + \int_{-\infty}^t v(t-s) dN(s) \right]$$

$$\mu = \lambda_0 + E \left[\int_{-\infty}^t v(t-s) \lambda(s) ds \right]$$

$$\mu = \lambda_0 + \int_{-\infty}^t v(t-s) \mu ds$$

$$\mu = \lambda_0 + \mu \int_0^{\infty} v(r) dr$$

$$\begin{aligned} \mu - \mu \int_0^{\infty} v(r) dr &= \lambda_0 \\ \mu(1 - \int_0^{\infty} v(r) dr) &= \lambda_0 \\ \mu &= \frac{\lambda_0}{1 - \int_0^{\infty} v(r) dr} = \frac{\lambda_0}{1 - \frac{\alpha}{\beta}} \\ \therefore \mu &= \frac{\lambda_0}{1 - \frac{\alpha}{\beta}} \quad (2 - 3) \end{aligned}$$

ويمكن استخدام المعادلة الأخيرة لتحديد القيمة المقدرة لعدد الحوادث وحسب الوحدة الزمنية. ولكي تكون عملية هوكس عملية مستقرة (أي ان معدل حدوث الحوادث مستقر تقريبا) يجب ان يتحقق الشرط الاتي:

$$\frac{\alpha}{\beta} < 1$$

إذ تسمى النسبة $\frac{\alpha}{\beta}$ بنسبة التفرع (Branching ratio) إذ ان هذه النسبة غالباً ما تمثل العدد المتوقع (expected number) للحوادث المستقبلية. فقد أشار الباحثان (Filimonov, Sornette, 2012) والباحث (Lorenzen, 2012) الى ان نسبة التفرع تعد المفتاح الرئيسي للتنبؤات بحركة اسعار الاسهم والحوادث المهمة لأسعار الاسهم. وبافتراض ان الدالة (.) v تمثل دالة كرنل الاسية (2-2) يمكننا ان نكتب دالة الكثافة الشرطية كالاتي:

$$\lambda^*(t_n) = \lambda_0 + \sum_{t_i < t} \alpha e^{-\beta(t-t_i)} \quad (2 - 4)$$

ولاستخدام عمليات هوكس في نمذجة ظاهرة معينة فان الامر يتطلب تحديد شرط اولي وبهذا فان دالة الكثافة الشرطية هي بمثابة حل معادلة تفاضلية عشوائية، اي ان الدالة (2-4) يمكن غالباً كتابتها بشكل معادلة تفاضلية عشوائية كما يأتي: (Laub et al, 2015)

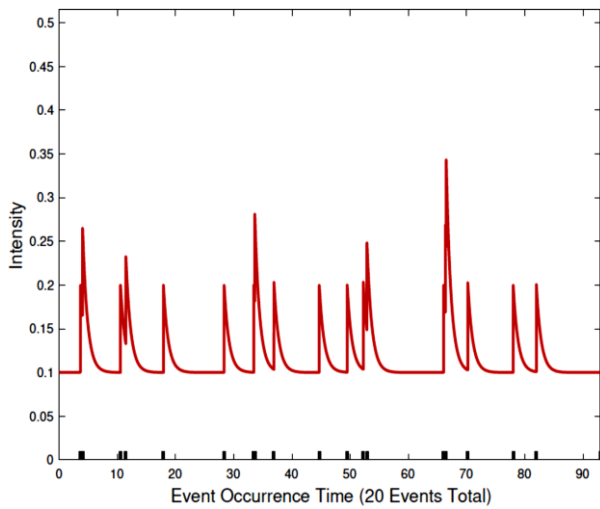
$$\lambda^*(t) = \beta(\lambda_0 - \lambda^*(t))dt + \alpha dN(t) \quad t \geq 0$$

وباستخدام التفاضل والتكامل العشوائي فان الحل العام سيكون كما يأتي:

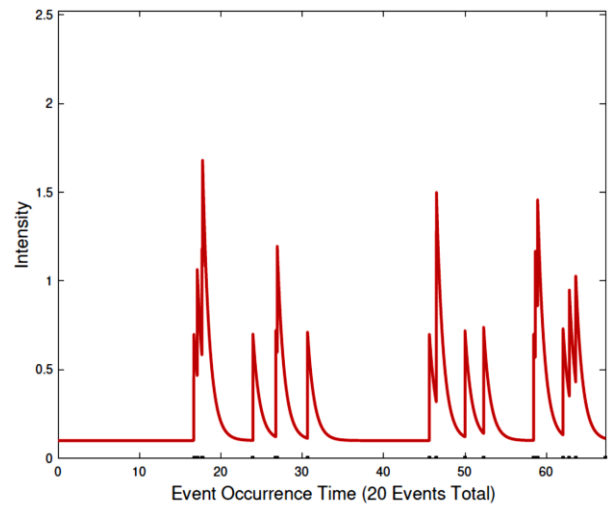
$$\lambda^*(t) = e^{-\beta t}(\lambda_0 - \lambda) + \lambda + \int_{-\infty}^t \alpha e^{-\beta(t-s)} dN(s), t \geq 0$$

وبافتراض قيم مختلفة لكل من λ_0 , α و β فانه يمكن رسم سلوك هذه الدالة وكما موضح

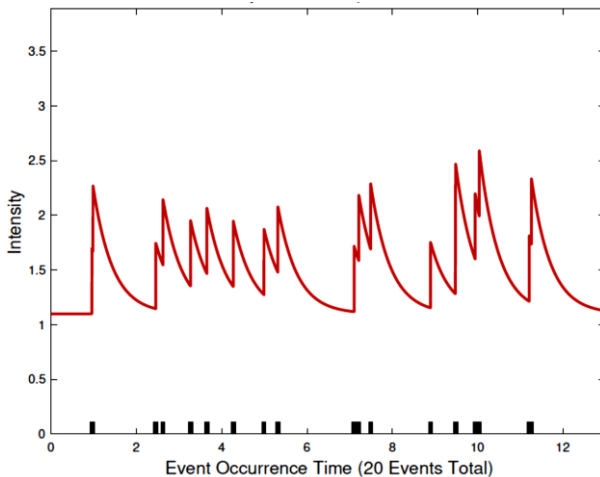
بالأشكال الاتية: (Haghdan,2017)



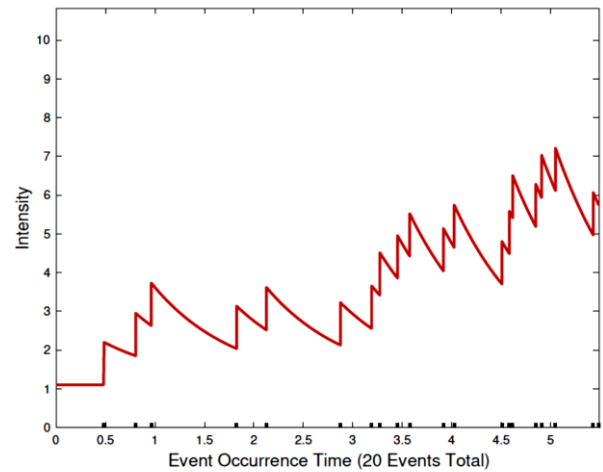
(a) $\alpha = 0.1$, $\beta = 1.2$, $\lambda_0 = 0.1$



(b) $\alpha = 0.6$, $\beta = 1.2$, $\lambda_0 = 0.1$



(c) $\alpha = 0.6$, $\beta = 2.2$, $\lambda_0 = 1.1$



(d) $\alpha = 1.1$, $\beta = 1.2$, $\lambda_0 = 1.1$

الشكل (2-3) اعلاه يوضح سلوك دالة الكثافة الشرطية عند قيم مختلفة لمعاملها

(Haghdan,2017)

إذ يتبين وبوضوح من الرسوم اعلاه ان مسار (pathway) دالة الكثافة الشرطية $\lambda^*(t)$ يتبع دالة الاضمحلال الاسي بين قفزة واخرى عبر الزمن.

2-4 اشتقاق دالة الإمكان لعملية هوكس: (Rizoiu et al. , 2017)

Hawkes process likelihood function derivation

في هذه الرسالة سوف نناقش مشكلة تقدير معالم عملية Hawkes، إذ نفترض مسبقا ان أوقات الوصول او اوقات حدث معين هي $\{T_1, T_2, T_3, \dots, T_n\}$ للعملية $N(t)$ والتي تمثل عملية نقطة معرفة بالفترة $[0, T]$ عندئذ فان دالة الإمكان لعملية Hawkes موضحة بالاتي:

$$\begin{aligned} f(T_1, T_2, \dots, T_n) &= \prod_{i=1}^n f(T_i / T_1, T_2, \dots, T_{i-1}) \\ &= \prod_{i=1}^n f^*(T_i) \end{aligned}$$

ان $f^*(t)$ تمثل دالة شرطية $f(t|H(t))$ ، حيث ان $H(t)$ تمثل الأوقات السابقة (history).

ان دالة الكثافة $\lambda(t)$ يمكن كتابتها بدلالة الدالة الشرطية f^* والدالة التجميعية (F^*) التي تقابلها اي ان:

$$\lambda(t) = \frac{f^*(t)}{1 - F^*(t)} \quad (2 - 5)$$

إذ يمكن إعادة كتابة الدالة اعلاه بصيغة الدالة التجميعية وكالاتي:

$$\lambda(t) = \frac{f^*(t)}{1 - F^*(t)} = \frac{\partial / \partial t F^*(t)}{1 - F^*(t)} = \frac{\partial}{\partial t} \log(1 - F^*(t))$$

ويأخذ التكامل لكل من الطرفين خلال الفترة (T_n, t) فان:

$$\int_{T_n}^t \lambda^*(s) ds = -[\log(1 - F^*(t))] - \log[1 - F^*(T_n)]$$

ان عملية Hawkes هي عملية نقطية بسيطة (Simple) فهذا يعني ان حدوث اكثر من حالة

وصول (Multiple arrivals) غير ممكن الظهور في الوقت نفسة وبالتالي فان $F^*(T_n) = 0$

عندما $T_{n+1} > T_n$ وان:

$$\int_{T_n}^t \lambda^*(s) ds = -\log[1 - F^*(t)]$$

وبإعادة كتابة المعادلة اعلاه نحصل على ما يأتي:

$$F^*(t) = 1 - \exp\left[-\int_{T_n}^t \lambda^*(s) ds\right]$$

وبضرب الطرفين في الوسطين في المعادلة (5-2) نحصل على ما يأتي:

$$\begin{aligned} f^*(t) &= \lambda(t)(1 - F^*(t)) \\ &= \lambda(t)\exp\left[-\int_{T_n}^t \lambda(s) ds\right] \end{aligned} \quad (2 - 6)$$

ومن خلال معادلة رقم (6-2) ودالة الإمكان لمجموعة المعالم θ ستكون لدينا دالة الامكان الاتية:

$$L(\theta) = \prod_{i=1}^n f^*(T_i) = \prod_{i=1}^n \lambda(T_i) \exp\left[-\int_{T_{i-1}}^{T_i} \lambda(u) du\right]$$

$$= \prod_{i=1}^n \lambda(T_i) \exp\left[-\int_0^{T_n} \lambda^*(u) du\right] \quad (2-7)$$

من تعريف دالة الإمكان رقم (2-7) يمكن توضيح النظرية في ادناه.

نظرية : دالة الإمكان لعملية هوكس Hawkes process likelihood function

(Laub et al.,2015) (Daley and Vere-Jones, 2003)

لتكن $N(\cdot)$ عملية نقطية معرفة بالفترة $[0, T]$ ، إذ $T \geq 0$ وان t_1, t_2, \dots, t_n تمثل أوقات الحوادث للعملية $N(\cdot)$ خلال الفترة $[0, T]$ ، عندئذ فان دالة الإمكان للعملية $N(\cdot)$ معرفة بالدالة الآتية:

$$L = \left[\prod_{i=1}^n \lambda^*(t_i) \right] \exp\left[-\int_0^T \lambda^*(u) du\right] \dots (2-8)$$

2-5 دالة الإمكان باستخدام الاضمحلال الاسي Likelihood function with

exponential decay (Ozaki,1979) (Laub et al.,2015)

باستخدام دالة الإمكان في النظرية المشار اليها بالمعادلة رقم (2-8)، فان لوغاريتم دالة الإمكان للفترة $[0, t_n]$ يمكن اشتقاقها بالشكل الآتي:

$$\begin{aligned} \log L(\theta) &= \log \prod_{i=1}^n f^*(T_i) \\ &= \log \prod \lambda^*(T_i) + \log\left[\exp\left(-\int_0^{t_n} \lambda^*(u) du\right)\right] \end{aligned}$$

$$= \sum_{i=1}^n \log \lambda^*(T_i) - \int_0^{t_n} \lambda^*(u) du \quad (2-9)$$

ان التكامل في المعادلة (2-9) يمكن تجزئته الى مجموعة من التكاملات بفترات مختلفة وكما يأتي:

لنفرض ان

$$\begin{aligned} \Lambda(t_n) &= \int_0^{t_n} \lambda^*(u) du \\ &= \int_0^{t_1} \lambda^*(u) du + \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} \lambda^*(u) du \end{aligned} \quad (2-10)$$

وبالاعتماد على فكرة ان الدالة $\lambda^*(\cdot)$ تتبع بسلوكها دالة الاضمحلال الاسي (exponential decay) أي ان:

$$\lambda^*(t_n) = \lambda + \sum_{t_i < t} \mu(t - t_i)$$

إذ ان الدالة $\mu(t)$ هي: $\mu(t) = \alpha e^{-\beta t}$

وبالتعويض عن $\mu(t)$ في المعادلة (2-10) نحصل على ما يأتي:

$$\Lambda(t) = \int_0^{t_1} \lambda du + \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} \lambda + \sum_{t_i < u} \alpha e^{-\beta(u-t_i)} du$$

$$\begin{aligned}
 &= \lambda t_1 + \alpha \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} \sum_{j=1}^i e^{-\beta(u-t_j)} du \\
 &= \lambda t_1 + \alpha \sum_{i=1}^{n-1} \sum_{j=1}^i \int_{t_i}^{t_{i+1}} e^{-\beta(u-t_j)} du \\
 &= \lambda t_1 - \frac{\alpha}{\beta} \sum_{i=1}^{n-1} \sum_{j=1}^i [e^{-\beta(t_{i+1}-t_j)} - e^{-\beta(t_i-t_j)}] \\
 \Lambda(t) &= \lambda t_1 - \frac{\alpha}{\beta} \sum_{i=1}^n [e^{-\beta(t_n-t_i)} - 1]
 \end{aligned}$$

وبتعويض صيغة $\Lambda(t_n)$ الأخير في لوغاريتم دالة الإمكان رقم (2-9) نحصل على ما يأتي:

$$\log L(\theta) = \sum_{i=1}^n \log[\lambda + \alpha \sum_{j=1}^{i-1} e^{-\beta(t_i-t_j)}] - \lambda t_n + \frac{\alpha}{\beta} \sum_{i=1}^n [e^{-\beta(t_n-t_i)} - 1] \quad (2-11)$$

ان اجراء عملية حسابية من المعادلة (2-11) يعد غير ممكن طالما ان الحد الثاني يمتلك رمزين للمجموع، فقد اشار (Ogata,1978) الى إمكانية حساب المعادلة

(2-11) من خلال فرض ما يأتي:

لنفرض ان

$$A(i) = \sum_{j=1}^{i-1} e^{-\beta(t_i-t_j)} \quad (2-12)$$

إذ يمكن تعريف $A(i)$ بدلالة $A(i-1)$ وكما يأتي:

$$\begin{aligned}
 A(i) &= \sum_{j=1}^{i-1} e^{-\beta t_i + \beta t_j} \\
 &= e^{-\beta t_i + \beta t_{i-1}} e^{\beta t_i + \beta t_{i-1}} \sum_{j=1}^{i-1} e^{-\beta t_i + \beta t_j} \\
 &= e^{-\beta t_i + \beta t_{i-1}} \sum_{j=1}^{i-1} e^{-\beta t_{i-1} + \beta t_j} \\
 &= e^{-\beta(t_i - t_{i-1})} \left[1 + \sum_{j=1}^{i-2} e^{-\beta(t_{i-1} - t_j)} \right] \\
 A(i) &= e^{-\beta(t_i - t_{i-1})} [1 + A(i-1)]
 \end{aligned}$$

علما ان $A(1) = 0$ وبالتالي فان لوغاريتم دالة الإمكان تعرف بالشكل الاتي:

$$\ell = \sum_{i=1}^n \log(\lambda + \alpha A(i)) - \lambda t_n + \frac{\alpha}{\beta} \sum_{i=1}^n [e^{-\beta(t_n - t_i)} - 1] \quad (2-13)$$

ان تقدير المعالم α, β, λ يكون من خلال إيجاد مشتقات الدالة (إيجاد الميل) لكل معلمة، بكلام

اكثر دقة اننا سوف نجد مقدر الإمكان الأعظم لمتجه المعالم $(\hat{\theta})$ إذ ان:

$$\hat{\theta} = \operatorname{argmax}_{\theta} [\log L(\theta)]$$

ان حل مشكلة التعظيم اعلاه تتم من خلال ايجاد المشتقات الرياضية الاتية:

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^n \frac{1}{\beta} (e^{-\beta(t_n - t_i)} - 1) + \sum_{i=1}^n \left[\frac{A(i)}{\lambda + \alpha A(i)} \right]$$

$$\frac{\partial \log L}{\partial \beta} = -\alpha \sum_{i=1}^n \left[\frac{1}{\beta} (t_n - t_i) e^{-\beta(t_n - t_i)} + \frac{1}{\beta^2} (e^{-\beta(t_n - t_i)}) \right]$$

$$- \sum_{i=1}^n \left[\frac{\alpha \beta(i)}{\lambda + \alpha A(i)} \right]$$

$$\frac{\partial \log L}{\partial \lambda} = -t_n + \sum_{i=1}^n \left[\frac{1}{\lambda + \alpha A(i)} \right]$$

إذ إن

$$B(i) = \sum_{t_j < t_i} (t_i - t_j) e^{-\beta(t_i - t_j)}$$

لقيم $i \geq 2$ و $B(1) = 0$.

في ادناه مكونات مصفوفة Hessian التي تمثل مصفوفة التباين والتباين المشترك للمعالم، هذه

المصفوفة تحتوي المشتقات الجزئية من المرتبة الثانية لدالة الامكان:

$$\frac{\partial^2 \log L}{\partial \alpha^2} = - \sum_{i=1}^n \left[\frac{A(i)}{\lambda + \alpha A(i)} \right]^2$$

$$\frac{\partial^2 \log L}{\partial \beta \partial \alpha} = - \sum_{i=1}^n \left[\frac{1}{\beta} (t_n - t_i) e^{-\beta(t_n - t_i)} + \frac{1}{\beta^2} (e^{-\beta(t_n - t_i)} - 1) \right]$$

$$+ \sum_{i=1}^n \left[\frac{-B(i)}{\mu + \alpha A(i)} + \frac{\alpha A(i) B(i)}{(\lambda + \alpha A(i))^2} \right]$$

$$\begin{aligned} \frac{\partial^2 \log L}{\partial \beta^2} = & \alpha \sum_{i=1}^n \left[\frac{1}{\beta} (t_n - t_i)^2 e^{-\beta(t_n - t_i)} + \frac{2}{\beta^2} (t_n - t_i) e^{-\beta(t_n - t_i)} \right. \\ & \left. + \frac{2}{\beta^3} (e^{-\beta(t_n - t_i - 1)}) \right] + \sum_{i=1}^n \left[\frac{\alpha c(i)}{\lambda + \alpha A(i)} - \left(\frac{\alpha B(i)}{\lambda + \alpha A(i)} \right)^2 \right] \end{aligned}$$

إذ ان:

$$c(i) = \sum_{t_j < t_i} (t_i - t_j)^2 e^{-\beta(t_i - t_j)}$$

لقيم $i \geq 2$ و $c(1) = 0$

$$\frac{\partial^2 \log L}{\partial \lambda^2} = \sum_{i=1}^n \left[\frac{-1}{(\lambda + \alpha A(i))^2} \right]$$

$$\frac{\partial^2 \log L}{\partial \alpha \partial \lambda} = \sum_{i=1}^n \left[\frac{-A(i)}{(\lambda + \alpha A(i))^2} \right]$$

$$\frac{\partial^2 \log L}{\partial \beta \partial \lambda} = \sum_{i=1}^n \left[\frac{\alpha B(i)}{(\lambda + \alpha A(i))^2} \right]$$

2-6 خوارزميات حل مقدر الإمكان الاعظم:

سوف نستعرض في هذه الرسالة خوارزمتين لتقدير معالم عملية هوكس هما الطريقة العددية المباشرة لتعظيم لوغاريتم الإمكان (direct numerical maximization of log-likelihood) واختصارا هي (DNMLL) والطريقة الثانية هي خوارزمية تعظيم التوقع التقريبية (Approximation expectation-Maximization algorithm) واختصارا (AEMA).

الخوارزمية الأولى (DNMLL) قدمها الباحث (Vere-Jones,1978) وكذلك قدمها الباحث (Ozaki,1979) بعد اجراء بعض التعديلات في الخوارزمية الحسابية، وقدم الباحث

(Ogata,1988) امثلة تطبيقية على استخدام خوارزمية (DNMLL) وكذلك الباحث (Bowsher,2077) وظف هذه الخوارزمية على بيانات مالية، اما عملية حساب المعالم وفق طريقة الإمكان الأعظم باستخدام خوارزمية (Expectation-Maximization) فقد تم استخدامها على سبيل المثال من قبل الباحثين (Veen,Schoenberg,2008) وكذلك الباحثين (Olson,Carley,2013).

في هذه البحث سوف نستخدم طريقة (AEMA) بدلا من (EM) من اجل اختصار العمليات الحسابية في الخوارزمية (EM).

2-7 خوارزمية التقدير DNMLL: (Wang,2015)

بفرض ان لدينا عملية عشوائية نقطية نطويه $\{X_t\}_{t \in (0,T]}$ بافتراض وقوع الاحداث عند الأوقات t_1, t_2, \dots, t_{Nt} وعلى فرض ان دالة الإمكان هي:

$$L(\theta) = \prod_{i=1}^{Nt} \lambda_{ti} \exp\left(-\int_0^T \lambda_s ds\right)$$

وبعد اخذ اللوغاريتم للدالة أعلاه نحصل على

$$\log L(\theta) = \sum_{i=1}^{NT} \log(\lambda_s) + \int_0^T (1 - \lambda_s) ds$$

وباستخدام دالة exponential decay في عمليات هوكس وبفرض وجود قيمة ابتدائية هي λ_0 فانه يمكن إغالباً كتابة الدالة اللوغاريتمية الأخيرة بشكل يسهل العملية الحسابية كما يأتي:

$$\log L(\theta) = T - T\lambda_0 - \frac{\alpha}{\beta} \sum_{i=1}^{NT} (1 - e^{-\beta(T-ti)}) + \sum_{i=1}^{NT} \log(\lambda_0 + \alpha A(i)) \quad (2 - 14)$$

وبهذا يمكن ان نقول ان دالة الكثافة الشرطية Conditional intensity يمكن ان تكتب كما يأتي:

$$\lambda_t = \lambda_0 + \alpha A(i)$$

إذ ان A(i) تم تعريفها بالمعادلة (2-12)

2-8 خوارزمية التقدير AEMA: (Wang,2015)

تعد طريقة (EM) بمثابة خوارزمية تكرارية لحساب قيم مقدرات الإمكان الأعظم عندما يتم النظر الى ان المشاهدات على انها غير كاملة (Incomplete) (Dempster et al.1977) ومن اجل اختصار العمليات الحسابية للخوارزمية (EM) فانه تم توضيح طريقة (AEMA) لعمليات هوكس بوجود دالة exponential decay تحت افتراض انه يمكن الحصول على صيغ تقريبية لمعادلات تقدير معالم عملية هوكس، ومن خلال اجراء بعض التعديلات والتقريبات الرياضية في خطوات خوارزمية (EM) فانه يمكن اجراء العملية الحسابية لتقدير المعالم (λ, α, β) حسب المعادلات الاتية:

$$\lambda_{\infty}^{(K+1)} = \frac{\sum_{i=1}^{N_T} \Pr\{u_i = i | H_{t_1}, \theta^{(K)}, t_i\}}{T}$$

$$\alpha^{(K+1)} = \frac{\beta^{(K+1)} \sum_{i=2}^{N_T} \sum_{j=1}^{i-1} \Pr\{u_i = j | H_{t_1}, \theta^{(K)}, t_i\}}{N_T}$$

$$\beta^{(K+1)} = \frac{\sum_{i=2}^{N_T} \sum_{j=1}^{i-1} \Pr\{u_i = j | H_{t_1}, \theta^{(K)}, t_i\}}{\sum_{i=2}^{N_T} \sum_{j=1}^{i-1} (t_i - t_j) \Pr\{u_i = j | H_{t_1}, \theta^{(K)}, t_i\}}$$

حيث ان:

$$P\{u_i = i | H_{t_i}, \theta^k, t_i\} = \frac{\lambda_{\infty}^{(k)}}{\lambda_{\infty}^{(k)} + \sum_{n: t_n < t_i} \alpha^{(k)} e^{-\beta^{(k)}(t_i - t_n)}}$$

$$P\{u_i = i | H_{t_i}, \theta^k, t_i\} = \frac{\alpha^{(k)} e^{-\beta^{(k)}(t_i - t_j)}}{\lambda_{\infty}^{(k)} + \sum_{n: t_n < t_i} \alpha^{(k)} e^{-\beta^{(k)}(t_i - t_n)}}$$

اذ ان:

T الزمن الكلي

N_T عدد الحالات الكلي

θ قيمة المعالم (α, β, λ)

K عدد التكرارات التي تم اجراها في الخوارزمية (iteration)

λ_{∞} معلمة الكثافة الابتدائية (initial)

2-9 اختبار جودة المطابقة لعملية هوكس:

اشار الباحث (Wang,2015) الى ان ابسط وأسهل اختبار لحسن المطابقة لعمليات هوكس هو من خلال حساب البواقي الناتجة من المعالم المقدره للعملية وبالتالي مقارنة توزيع البواقي بالتوزيع الاسي بمعلمة واحد عدد صحيح، إذ استخدم الباحث اختبار كولموكروف-سيمرنوف (K-S) (Kolmogorov-Simirnov)

إذ ان إحصاء الاختبار لهذا الاختبار تحسب بالمعادلة الآتية:

$$D_n = \sup_x |F_n(x) - F(x)|$$

إذ ان:

$F_n(x)$: هي دالة التوزيع التجريبي

$F(x)$: هي دالة التوزيع الاسي

وبافتراض ان فرضية العدم والفرضية البديلة هما كالآتي:

H_0 : Exp(1) تتبع توزيع

H_1 : Exp(1) لا تتبع توزيع

وتحت مستوى معنوية 1%.

الفصل الثالث

المحاكاة والتطبيق العملي

3-1 المقدمة:

يحتوي هذا الفصل على جزأين رئيسيين إذ يحتوي الجزء الأول على تجربة المحاكاة من خلال افتراض عدة سيناريوهات للتجربة عن حجوم عينات مختلفة وقيم معالم ابتدائية مختلفة مستخدمين البرمجة بلغة R، ومن اجل المقارنة بين أداء الطرائق المستخدمة في تقدير معالم عملية هوكس تم استخدام معيار التحيز (Bias) والانحراف المعياري (Standard deviation)، وتجدر الإشارة هنا الى اننا سوف نستخدم خوارزميتين لتقدير معالم عملية هوكس هما الخوارزمية العددية المباشرة لتعظيم لوغاريتم الإمكان (direct numerical maximization of log-likelihood) (DNMLL) وخوارزمية تعظيم التوقع التقريبية (-Approximation expectation Maximization algorithm) (AEMA) مستخدمين تقنية (Chain Monte Markov) (MCMC) (Carlo).

اما الجزء الثاني فيحتوي على توظيف الخوارزميات المذكورة في أعلاه لتقدير معالم عمليات هوكس من خلال استخدام بيانات حقيقية تمثل اعداد الإصابات لمرض Covid-19 المسجلة في محافظة كربلاء المقدسة لسنة 2021 للأشهر (نيسان، أيار، حزيران، تموز، اب، ايلول) وبالتالي استخدام معيار التحيز والانحراف المعياري للحكم على أفضلية خوارزمية للتقدير.

3-2 مفهوم المحاكاة:

يعد أسلوب المحاكاة أحد اهم الأساليب التي تطورت مع ظهور الحواسب السريعة لكونها تعتمد في عملها على تطبيق لخوارزميات رقمية لتمثل سلوك ظاهرة افتراضية معينة، إذ تستخدم المحاكاة في الحالات التي يكون فيها من الصعب الحصول على بيانات كافية لدراسة سلوك ظاهرة معينة إذ ربما تتمثل هذه الصعوبة بتوفر الوقت او الجهد او التكاليف، فضلاً عن ذلك يلجأ الباحث الى استخدام أسلوب المحاكاة في دراسة سلوك العمليات الرياضية او النظريات او الطرائق الرياضية المقترحة من اجل فهم آلية العمل والظروف المثلى التي تعمل بها هذه الطرائق الرياضية وبالتالي يساعد على فهم ما يجري في العالم الحقيقي، وبهذا فإنه يمكن القول ان المحاكاة هي تقليد لبناء نموذج او نظام يشابه النموذج او النظام الحقيقي (وجود حد خطأ صغير جداً).

وهذا يعني انه بصورة عامة فان تقنية المحاكاة تستند في عملها الى وصف كامل للعلاقات الرياضية المنطقية التي تسهل الية عمل نظام معين او فهم تجربة معينة مستخدمين بذلك برنامجاً حاسوبياً لتسهيل العمليات الحسابية.

وبالنتيجة يمكن القول ان تقنية المحاكاة غالباً ما يستعمل لمراقبة وملاحظة التغيرات التي تسلكها تجربة معينة عند تنفيذها وبالتالي فان تعدد التجارب في المحاكاة يولد لدى الباحث رؤيا واضحة عن المتغيرات التي تحدث في نتائج التجربة وبالتالي يمكن اختبار التجربة التي تحاكي الواقع العملي للبيانات المراد دراستها (البيانات الحقيقية) أي ان المحاكاة هي أساس من خلالها يمكن ان يحدد الباحث أفضل أسلوب لتحليل ناتج للبيانات الحقيقية.

والجدير بالذكر ان تقنية MCMC تعد من أكثر التقنيات او الخوارزميات المستخدمة في تنفيذ تجارب المحاكاة من خلال تنفيذ عدة خوارزميات كل منها له ظروف تشغيل تتطابق مع الأسلوب الرياضي او المعادلات الرياضية التي وصف بها النموذج المدروس، إذ نعتمد بالأساس على توليد عينات(بيانات) بأحجام مختلفة للتجربة المدروسة، أي ان الحاسوب يقوم بتوليد اعداد عشوائية تحاكي من خلالها عملية توليد البيانات إذ غالباً ما يستخدم التوزيع المنتظم في توليد الأرقام العشوائية للعملية المراد دراستها لأنها تعطي قيما او مشاهدات لها نفس فرص الظهور عند السحب العشوائي لأي توزيع من التوزيعات الاحتمالية المعروفة.

3-3 بناء نموذج المحاكاة:

كما بينا سابقا فان الباحث سوف يعتمد على خوارزميتين لتقدير معالم عملية هوكس، إذ تم اجراء تجربتين للمحاكاة كل واحدة بقيم مختلفة ابتدائية لمعالم عملية هوكس (λ, α, β) مع توليد (100) عينة كل واحدة منها تكون فيها البيانات المتولدة من عملية هوكس ضمن المدى $[0, 1000]$ أي ان المرحلة الأولى كانت هي تعيين القيم الابتدائية الافتراضية (initial values) للمعالم المراد تقديرها، هذه المدخلات لتجربة المحاكاة من خلالها يتم بناء نموذج المحاكاة الذي يمثل سلوك التجربة، والجدول الآتي يوضح القيم الافتراضية الابتدائية لمعالم عملية هوكس.

جدول (3-1) القيم الافتراضية الابتدائية لمعالم عملية هوكس

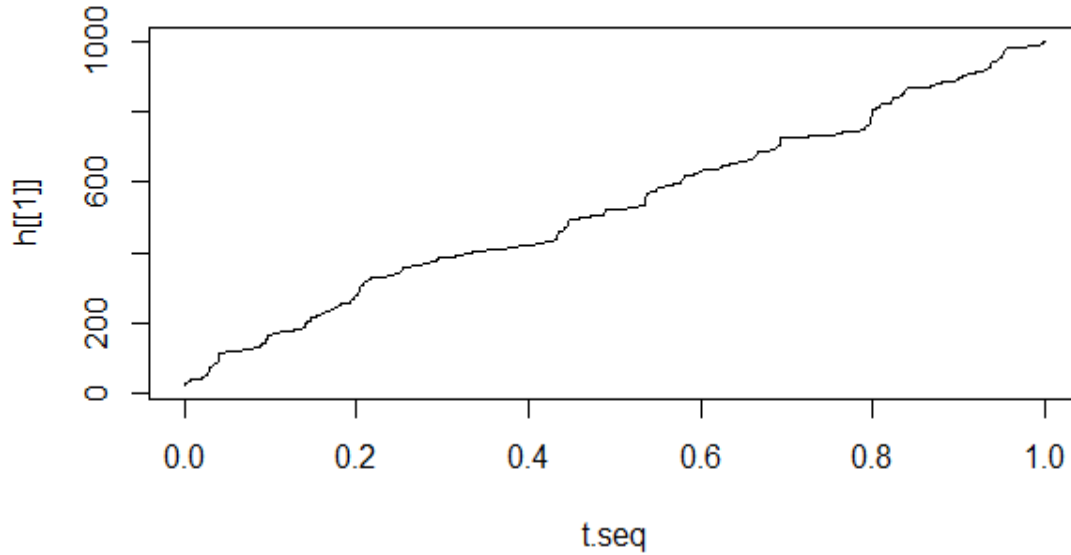
المجموعة	λ	α	β
1	0.15	0.20	0.5
2	0.05	0.05	0.06

إذ كان الأساس في اختيار قيم هذه المعالم هو من خلال الاطلاع من أسلوب دالة (exponential decay) في وصف الانتقال لعملية هوكس من حالة الى أخرى وبافتراض استقراره عملية هوكس وهذا واضح من خلال تحقق شرط نسبة التفرع $\frac{\alpha}{\beta}$ الذي قيمته اقل من واحد عدد صحيح.

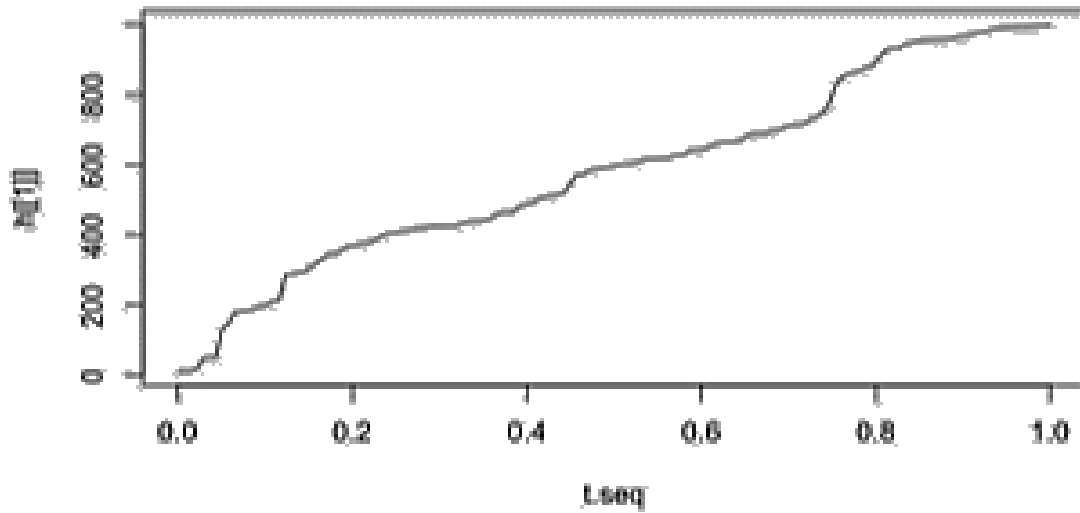
وبعد ان تم بيان كيفية تحديد الية توليد المشاهدات من قبل عملية هوكس والقيم الافتراضية للمعالم ستكون المرحلة الآتية إيجاد المقدرات الحقيقية لمعالم عملية هوكس وتلخيص نتائج التقدير بحساب قيم كل من التحيز والانحراف المعياري وبالتالي نصل الى المرحلة الأخيرة وهي المفاضلة بين خوارزمية التقدير.

3-4 نتائج المحاكاة:

تم ادراج جميع النتائج التي حصلنا عليها في تنفيذ خوارزمية MCMC بالمعطيات المشار اليها سابقا، اولاً سيتم رسم البيانات التي ولدتها عملية هوكس وكما موضحة بالشكل ادناه.



شكل (3-1) يمثل عدد الحوادث المتولدة من عملية هوكس للمجموعة الأولى



شكل (3-2) يمثل عدد الحوادث المتولدة من عملية هوكس للمجموعة الثانية

من الواضح من الأشكال اعلاه ان رسم البيانات المتولدة من عملية هوكس مستمرة نحو اليمين مما يؤكد على انها عملية تحفيز ذاتي وهذا ما ينطبق مع الاساس النظري لعملية هوكس.

بعد ذلك تم عرض نتائج الخوارزميتين (DNMLL) و (AEMA) من خلال ادراج قيم التحيز والانحراف المعياري من اجل الحصول على أفضل تقدير لمعالم عملية هوكس الجدول (2-3) يتضمن قيم معيار التحيز والانحراف المعياري للقيم الافتراضية للمجموعة الاولى.

جدول (3-2) القيم المقدرة للتحيز والانحراف المعياري للنماذج المختلفة للتجربة الأولى

النموذج	$\lambda = 0.15$		$\alpha = 0.2$		$\beta = 0.5$	
	Bias	S.D.	Bias	S.D.	Bias	S.D.
DNMLL	0.0017	0.0184	0.0065	0.0529	0.0245	0.127
AEMA	0.0019	0.0165	0.0089	0.0510	0.0388	0.1309

من الجدول (3-2) اظهرت نتائج المحاكاة بطريقة مونت-كارلو وباستخدام المجموعة الاولى من القيم الافتراضية لمعالم عملية هوكس وللخوارزميتين DNMLL و AEMA تقارب نتائج تقدير التحيز والانحراف المعياري، ومن الواضح ان متوسط مقدرات المعلمة λ قريبة من القيمة الافتراضية ($\lambda=0.15$)، ان طرائق التقدير اعطت تقريبا نفس النتائج لهذه المعلمة وهذا واضح من خلال قيم التحيز والانحراف المعياري، اما بالنسبة للمعلمة α فنجد ان مقدرات هذه المعلمة قريبة من القيمة الافتراضية (0.2) ولكلا الخوارزميتين إذ كانت نتيجة التقدير بخوارزمية DAMLL أقرب للقيمة المفترضة، كذلك تبين ان متوسط مقدرات المعلمة β عانت من تحيز عالٍ في كل من خوارزميتي التقدير. الجدول (3-3) يتضمن قيم معيار التحيز والانحراف المعياري للقيم الافتراضية للمجموعة الثانية.

جدول (3-3) القيم المقدرة للتحيز والانحراف المعياري للنماذج المختلفة للتجربة الثانية

النموذج	$\lambda = 0.05$		$\alpha = 0.05$		$\beta = 0.06$	
	Bias	S.D.	Bias	S.D.	Bias	S.D.
DNMLL	0.0059	0.0206	0.0005	0.0149	0.0079	0.0279
AEMA	0.0119	0.0224	0.0059	0.0166	0.0248	0.0269

نتائج الجدول (3-3) تشير خوارزمية التقدير AEMA الى وجود تحيزا طفيفا لكن الطريقة الأولى DNMLL تبقى الاقرب من قيم المعالم الافتراضية.

3-5 تحليل البيانات الحقيقية:

يتضمن هذا المبحث توظيف خوارزميتي التقدير (DNMLL) و (AEMA) لمعالم عملية هوكس من اجل إيجاد مقدرات هذه العملية من خلال البيانات الحقيقية المتمثلة بوباء Covid-19، إذ تم الاعتماد على بيانات تمثل الإصابة بهذا الوباء في محافظة كربلاء المقدسة المنشورة رسميا على موقع وزارة الصحة العراقية فمن خلال اخذ عدد الإصابات وعلى أساس يومي للأشهر (نيسان، أيار، حزيران، تموز، اب، ايلول) لسنة 2021 والتي تعد ذروة الوباء، فتم الاعتماد على عدد المصابين في كل يوم، وكان مجموع الاصابات خلال الفترة المذكورة (27694) حالة إصابة، وبناءً على هذه البيانات تم تطبيق نموذج الدراسة في المعادلة (4-2) الذي يمثل الصدمات (القفزات) لأعداد المصابين خلال أيام الشهر وتقدير معالم لهذا النموذج.

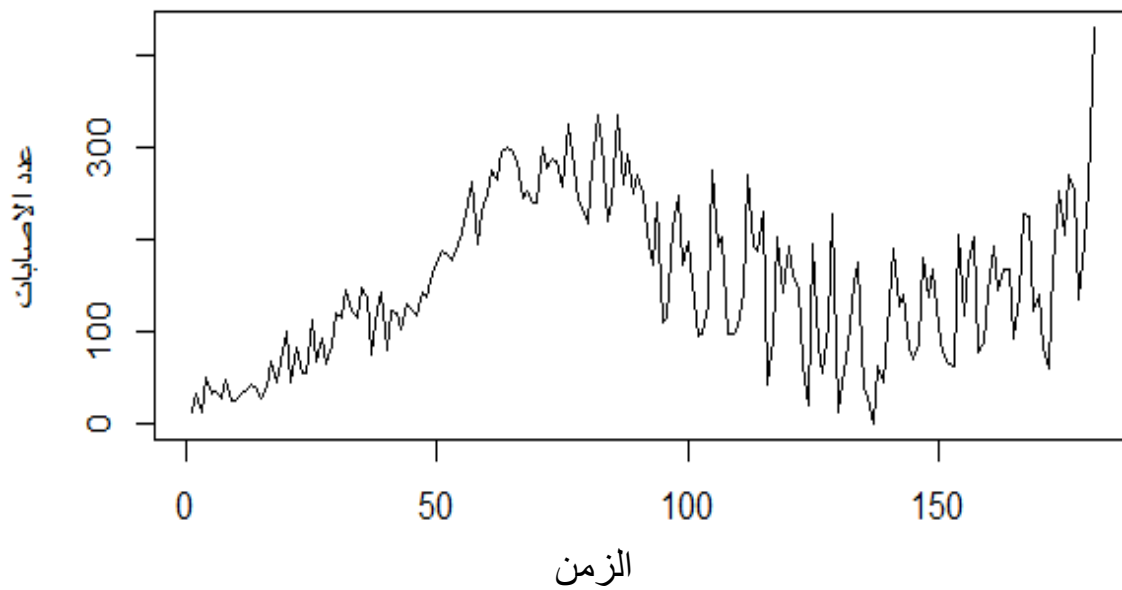
جدول (3-4) بيانات وباء Covid-19 تمثل اعداد الإصابات

I	N	I	N	I	N
1	11	14	38	27	92
2	31	15	27	28	65
3	11	16	40	29	84
4	50	17	68	30	119
5	31	18	44	31	115
6	34	19	67	32	145
7	28	20	100	33	122
8	47	21	44	34	114
9	25	22	82	35	146
10	24	23	54	36	138
11	33	24	55	37	75
12	38	25	113	38	125
13	43	26	67	39	141
I	N	I	N	I	N
40	80	64	300	89	250
41	121	65	296	90	269
42	119	66	285	91	251
43	103	67	244	92	208
44	129	68	252	93	173
45	122	69	239	94	241
46	118	70	239	95	109
47	141	71	299	96	115

48	137	72	278	97	211
49	161	73	287	98	248
50	171	74	257	99	173
51	187	75	325	100	197
52	185	76	295	101	152
53	176	78	243	102	95
54	187	79	235	103	96
55	208	80	218	104	130
56	240	81	278	105	274
57	263	82	336	106	193
58	195	83	304	107	203
59	231	84	221	108	97
60	250	85	239	109	98
61	275	86	335	110	105
62	264	87	260	111	141
63	294	88	293	112	271
I	N	I	N	I	N
113	192	136	29	159	88
114	186	137	0	160	142
115	229	138	62	161	193
116	41	139	44	162	145
117	91	140	103	163	168
118	203	141	189	164	168
119	142	142	126	165	93
120	192	143	139	166	140
121	159	144	82	167	227
122	148	145	70	168	224
123	65	146	88	169	123

124	20	147	179	170	139
125	195	148	136	171	81
126	85	149	167	172	59
127	54	150	106	173	170
128	108	151	77	174	253
129	228	152	64	175	205
130	13	153	62	176	270
131	56	154	205	177	252
132	84	155	117	178	135
133	152	156	173	179	195
134	175	157	202	180	255
135	39	158	76	181	430

الشكل ادناه يوضح اعداد الاصابات بفايروس covid-19 للأشهر (نيسان، أيار، حزيران، تموز، اب، أيلول) وبالأيام.



شكل (3-3) يمثل عدد الاصابات خلال فترة الدراسة

يلاحظ من الشكل اعلاه ان القمم (peaks) تدل على ان اعداد الاصابات بفايروس covid-19 حدثت على شكل تكتلات عنقودية (cluster), وبالتالي يمكن القول ان نمذجة البيانات اعلاه من خلال عملية هوكس هو امر ممكن من خلال خوارزميتي التقدير (DNMLL) و (AEMA)، تم تقدير معالم عملية هوكس المشار اليها بالمعادلة (2-4) وكانت نتائج التقدير لكل من الخوارزميات أعلاه موضحة بالجدول (3-5).

جدول (3-5) يوضح نتائج تقدير الخوارزميتين DNMLL و AEMA

Method	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}/\hat{\beta}$	$E(\lambda)$
DNMLL	0.0586	0.0057	0.0113	0.5044	0.1182
AEMA	0.0793	0.0048	0.0171	0.2807	0.1102

نتائج الجدول (3-5) تبين القيم المقدرة لمعالم عملية هوكس فضلاً عن القيم المقدرة لنسبة التفرع وحساب القيمة المتوقعة لأعداد الاصابات الداخلة لوحدة الطوارئ والتي يقدر معدل عددها (0.1068) بالدقيقة وهو مقارب جدا لنتائج العمود الاخير ($E(\lambda)=0.1182, E(\lambda)=0.1102$) ولكل من الطريقتين. كذلك ومن خلال قيم نسبة التفرع يمكن القول ان عملية هوكس المقدرة هي عملية مستقرة لأن قيم $\hat{\alpha}/\hat{\beta}$ اقل من الواحد الصحيح ولكل من خوارزميتي التقدير.

وبعد تقدير المعالم لعملية هوكس تم حساب البواقي وحساب الدالة التراكمية التجريبية بهدف اجراء حسن المطابقة K-S وكانت نتائج الاختبار كما يأتي:

جدول (3-6) إحصاءات اختبار K-S

إحصاءة K-S	P-Value	القرار
1.560	0.015	قبول فرضية العدم

ومن الجدول (3-6) نستنتج ان القرار كان هو قبول فرضية العدم التي تشير الى ان البيانات متولدة من التوزيع الاسي لمعلمة (1) مما يدل على انها متولدة من عملية هوكس تحت مستوى معنوية 1%.

الفصل الرابع

الاستنتاجات والتوصيات

4-1 الاستنتاجات:

- 1- ان تطبيق عمليات هوكس في البيانات الصحية ولا سيما بيانات فايروس (Covid-19) تعد من المواضيع الحديثة التطبيق، هذه الرسالة تناولت الأساس النظري لعمليات هوكس ذات البعد الواحد والزمن وإعطاء مفهوم العمليات النقطية وعمليات بواسون.
- 2- يمكن القول ان عمليات هوكس تعد أداة اتخاذ قرار للباحثين في معظم المجالات التطبيقية التي تتصف بالتغيرات والفترات بحدوث الحوادث عبر الزمن.
- 3- عمليات هوكس توفر معلومات قيمة للباحثين في تتبع ودراسة سلوك هذه العملية من خلال افتراض قيم مبدئية لمعالم العملية وتحت حجوم عينات مختلفة مستخدمين خوارزميتين للتقدير هي خوارزمية (DNMLL) وخوارزمية (AEMA) واختبار كفاءة هاتين الخوارزميتين في إيجاد أفضل المقدرات لمعالم عملية هوكس مستخدمين معيار التحيز وقيمة الانحراف المعياري.
- 4- تبين تفوق خوارزمية (DNMLL) بشكل طفيف على خوارزمية (AEMA) من خلال المعايير أعلاه، فضلاً عن سرعة التنفيذ.
- 5- تناول الجانب التطبيقي تحليل بيانات (Covid-19) لسنة 2021 للأشهر (نيسان، أيار، حزيران، تموز، اب، ايلول)، إذ تم تقدير معالم عملية هوكس وتم التأكد من ان البواقي تتبع عملية هوكس ذات البعد الواحد مستخدمين اختبار كولموكروف- سمرنوف (K-S).
- 6- تم التأكد من استقرارية عملية هوكس من خلال استخدام القيم المقدره لمعالم دالة الاضمحلال الاسي، إذ قدمت القيمة المقدره لنسبة التفرع دليلاً على استقرار عملية هوكس وبانها تحفز حدوث قفزات بالإصابات، وبالنتيجة من خلال تقدير معالم عملية هوكس للبيانات الحقيقية ومن

اختبار حسن مطابقة البواقي لعملية هوكس يمكن القول ان هذه العملية تصف سلوك بيانات الإصابات لوباء (Covid-19) أفضل وصف.

7- يمكن استخدام هذه العملية في دراسة السلوك المستقبلي لهذا الفيروس من خلال تحليل القفزات التي تحدث بالإصابات مع مرور الزمن.

4-2 التوصيات: -

يوصي الباحث بناءً على ما جاء في الجانب النظري وجانب المحاكاة والتطبيق للبيانات الحقيقية ما يأتي:

1- توسيع العمل بعمليات هوكس ذات البعدين (Two-dimensional Hawkes processes)

نظرياً وتطبيقاً من خلال اعتماد بعد الزمن وبعد المكان للظاهرة المدروسة.

2- استخدام أسلوب بيز في تقدير معالم عملية هوكس ذات البعد الواحد أو البعدين، فضلاً عن تقدير معدل الكثافة البيزي.

3- دمج نموذج عملية هوكس ذات البعدين مع العمليات النقطية المسمى (Marked point

processes) التي تستخدم علامة (Marks) أخرى غير الزمن تصف عملية وصول الحوادث.

4- نوصي وزارة الصحة بالاعتماد على هذا النموذج لكونه دقيقاً في التقدير والتنبؤ بأعداد المصابين بمرض Covid-19 في المستقبل.

المصادر

- [1] Bacry, E., Mastromatteo, I., & Muzy, J. F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01), 1550005.
- [2] Beichelt, F. (2018). *Applied Probability and Stochastic Processes*. CRC Press.
- [3] Bartlett, M. S. (1963). Statistical estimation of density functions. *Sankhyā: The Indian Journal of Statistics, Series A*, 245-254.
- [4] Bartlett, M. S. (1963). The spectral analysis of point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 25(2), 264-281.
- [5] Bartlett, M. S. (1964). The spectral analysis of two-dimensional point processes. *Biometrika*, 51(3/4), 299-311.
- [6] Bonnet, A., Herrera, M. M., & Sangnier, M. (2021). Maximum Likelihood Estimation for Hawkes Processes with self-excitation or inhibition. *Statistics & Probability Letters*, 179, 109214.
- [7] Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2), 876-912.
- [8] Chen, F., & Stindl, T. (2018). Direct likelihood evaluation for the renewal Hawkes process. *Journal of Computational and Graphical Statistics*, 27(1), 119-131.
- [9] Chiang, W. H., Liu, X., & Mohler, G. (2022). Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *International journal of forecasting*, 38(2), 505-520.
- [10] Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society: Series B (Methodological)*, 17(2), 129-157.
- [11] Daley, D. J., & Vere-Jones, D. (2003). *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer New York.
- [12] Dassios, A., & Zhao, H. (2013). Exact simulation of Hawkes process with exponentially decaying intensity. *Electronic Communications in Probability*, 18, 1-13.

- [13] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1), 1-22.
- [14] Engelke, S.; Hautphenne, S.; Hurth T. Notes of Applied Stochastic processes.EPEL. 2018.
- [15] Filimonov, V., & Sornette, D. (2012). Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Physical Review E*, 85(5), 056108.
- [16] Gao, X., Zhou, X., & Zhu, L. (2018). Transform analysis for Hawkes processes with applications in dark pool trading. *Quantitative Finance*, 18(2), 265-282.
- [17] Haghdan, M. (2017). Hawkes Process Models for Unsupervised Learning on Uncertain Event Data (Doctoral dissertation, University of Toledo).
- [18] Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(3), 438-443.
- [19] Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1), 83-90.
- [20] Hawkes, A. G. (2020). Hawkes jump-diffusions and finance: a brief history and review. *The European Journal of Finance*, 1-15.
- [21] Landström, J. (2019). The Hawkes process—a self-exciting Poisson shot noise model.
- [22] Laub, P. J., Taimre, T., & Pollett, P. K. (2015). Hawkes processes. arXiv preprint arXiv:1507.02822.
- [23] Lesage, L., Deaconu, M., Lejay, A., Meira, J. A., & Nichil, G. (2022). Hawkes processes framework with a Gamma density as excitation function: application to natural disasters for insurance. *Methodology and Computing in Applied Probability*, 1-29.
- [24] Lewis, P. A. (1964). A branching Poisson process model for the analysis of computer failure patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(3), 398-441.

- [25] Li, S. (2021). Suicide and Copycat Behavior: An Analysis with Hawkes Process of England and Wales Suicide Data. University of California, Los Angeles.
- [26] Lorenzen, F. (2012). Analysis of order clustering using high frequency data: A point process approach. Swiss Federal Institute of Technology Zurich (ETH Zurich). <http://arno.uvt.nl/show.cgi>.
- [27] Magris, M. (2019). On the simulation of the Hawkes process via Lambert-W functions. arXiv preprint arXiv:1907.09162.
- [28] Møller, J., & Rasmussen, J. G. (2005). Perfect simulation of Hawkes processes. *Advances in applied probability*, 37(3), 629-646.
- [29] Obral, K. (2016). Simulation, estimation and applications of hawkes processes (Master's thesis, University of Minnesota).
- [30] Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30(1), 243-261.
- [31] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401), 9-27.
- [32] Olinde, J., & Short, M. B. (2020, December). A Self-limiting Hawkes Process: Interpretation, Estimation, and Use in Crime Modeling. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 3212-3219). IEEE.
- [33] Olson, J. F., & Carley, K. M. (2013). Exact and approximate em estimation of mutually exciting hawkes processes. *Statistical Inference for Stochastic Processes*, 16(1), 63-80.
- [34] Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1), 145-155.
- [35] Rizoiu, M. A., Lee, Y., Mishra, S., & Xie, L. (2017). A tutorial on hawkes processes for events in social media. arXiv preprint arXiv:1708.06401.

- [36] Rodriguez.A.B.(2019). Hawkes processes in finance. thesis. university of Barcelona. <http://diposit.ub.edu>.
- [37] Simon, G. (2016). Hawkes Processes in Finance: A Review with Simulations.
- [38] Swishchuk, A., Zagst, R., & Zeller, G. (2021). Hawkes processes in insurance: Risk model, application to empirical data and optimal investment. *Insurance: Mathematics and Economics*, 101, 107-124.
- [39] Unwin, H. J. T., Routledge, I., Flaxman, S., Rizoïu, M. A., Lai, S., Cohen, J., ... & Bhatt, S. (2021). Using Hawkes Processes to model imported and local malaria cases in near-elimination settings. *PLoS computational biology*, 17(4), e1008830.
- [40] Veen, A., & Schoenberg, F. P. (2008). Estimation of space–time branching process models in seismology using an em–type algorithm. *Journal of the American Statistical Association*, 103(482), 614-624.
- [41] Vere-Jones, D. (1978). Earthquake prediction-a statistician's view. *Journal of Physics of the Earth*, 26(2), 129-146.
- [42] Wang, H. (2022). Limit theorems for a discrete-time marked Hawkes process. *Statistics & Probability Letters*, 184, 109368.
- [43] Wang.Q.(2015).Applications of Hawkes process in Finance.Master thesis. Tilburg University. <http://arno.uvt.nl>.
- [44] Wehrli, A., Wheatley, S., & Sornette, D. (2021). Scale-, time-and asset-dependence of Hawkes process estimates on high frequency price changes. *Quantitative Finance*, 21(5), 729-752.

ABSTRACT

This thesis dealt with the study of stochastic self-exciting processes called Hawkes processes, where it is usually many accidents as they occurs form data over time, which are called cluster events. That is, the processes of arrival or the occurrence of the event are represented by the cluster samples in which it is the occurrence of each event stimulates the occurrence of another event at an accelerated rate, similar to a cluster. Consequently, we can say that Hawkes processes are stochastic processes depends in its analysis of accidents on the effects resulting from the occurrence of previous accidents and this is called the effect of self-exciting to the event. Simulation experiments were conducted to study the behavior of the Hawkes process, using two methods to estimate the parameters of the Hawkes process, the first method is (DNMLL) and the second method is the (AEMA) method, the Bias criterion and the standard deviation have used to assess the quality of parameters estimates. The results show the superiority of the (DNMLL) method in the estimation. In addition, we have employed the parameters estimation methods that mentioned above for studying the behavior incident data of the number of people infected with COVID-19. The Kilmokrov-Smirnov test has used to test the data fitting to the Hawks process. As well as, the branching ration has obtained to confirm the stationary of Hawkes process.

Republic of Iraq
Ministry of Higher Education & scientific research
University of Al-Qadisiyah
College of Administration and Economics
Department of Statistics
Graduate studies



PARAMETERS ESTIMATION OF HAWKES PROCESS WITH AN APPLICATION

**A thesis submitted to the Council of College of
Administration & Economics\ University of Al-Qadisiya as
partial fulfillment of the requirement for the degree
of Master of Statistics Sciences**

By

Ayman Abbas Najm

Supervised by

Prof.Dr. Muhannad F. Al-Saadony

2023 A.D.

1444 A.H.