**Title**
Suicide and Copycat Behavior: An Analysis with Hawkes Process of England and Wales Suicide Data



**Author**
Li, Sixuan

**Publication Date**
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Suicide and Copycat Behavior:
# An Analysis with Hawkes Process
# of England and Wales Suicide Data

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Statistics

by

Sixuan Li

2021

# ABSTRACT OF THE THESIS

Suicide and Copycat Behavior:
An Analysis with Hawkes Process
of England and Wales Suicide Data

by

Sixuan Li
Master of Science in Statistics
University of California, Los Angeles, 2021
Professor Frederic Paik Schoenberg, Chair

Suicide is a leading cause of death worldwide. This thesis presents the application of Hawkes processes on suicide data from England and Wales since no paper has ever used the Hawkes process model on the analysis of such a topic before. I began by introducing background knowledge and a brief data description. After reviewing the concepts of the Hawkes process and of the general point process, I modified the marked Hawkes process model and applied it, in order to understand the suicide copycat behaviors or even suicide cascade phenomenon based on the data obtained. Two kernel functions were used to compare the predicted results. The Hawkes process model with the power-law kernel seems to provide a relatively more reliable prediction. With predictions, the corresponding preventative interventions were significant as well. Last, I provided an overview of the current analysis and offered further directions and future potential improvements on the explorations of this topic.

The thesis of Sixuan Li is approved.

Jingyi Jessica Li

Hongquan Xu

Frederic Paik Schoenberg, Committee Chair

University of California, Los Angeles

2021

*To my parents*

*who always support and believe in me*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

*Thank my advisor Professor Rick Schoenberg and my friend Jie who give me valuable inspiration and constructive suggestions. Thank my committee members Professor Li and Professor Xu who offer great comments and feedback.*

# CHAPTER 1

# Background

Suicide is a social issue of great concern in many regions. Nearly 800,000 people take their own life each year globally, and even more people try to attempt suicide, a harmful act that influences all ages, genders, and races. Each suicide is a long-lasting tragedy that affects families, communities, and even countries. In order to disrupt the suicide ideation and discontinue the suicidal attempt, many organizations and government institutes, such as the American Association of Suicidology, National Institute of Mental Health(NIMH), etc., make "predictive" prevention efforts and then aim to apply limited resources to high-intensity geographic areas and time spans. Besides, there have been many groups attempting to prevent suicide or intervene in suicidal ideation from various aspects, including studying the effects on the propensity to commit suicide due to race, ethnicity, immigrant status, age, economic hardships, gender, and psychological traits, etc.

In the meantime, a variety of research methods have been applied to the estimation of suicide risk hotspots, including meta-analysis [NFY12], time-series analysis [FSK18], machine learning [CMD20], and artificial neural networks models [HRG21].

On the other hand, a number of studies have been proposed with the use of self-exciting processes, i.e. Hawkes processes, to conduct forecasts with high performances. The epidemic prediction of the incidence of disease [CLM20] and forecasting dynamics of crime [MCR18] are great illustrations.

Since no analysis of using Hawkes process on the suicide data has been presented, this self-exciting process model could be proven useful to improve understanding of the process generating suicides and to predict new events in the near future.

The outline of this thesis is as follows. In section 2, I introduce Hawkes processes, motivations for choosing this model, and dataset search and collection course. The daily suicide data is very critical to this analysis, so I have to give up the United States weekday-level suicide dataset. In Section 3, I present an overview of the dataset I used; it is in England and Wales from 2010 to 2015. The publicized celebrity suicide cases of that time frame are gleaned as well. In Section 4, some key concepts to explain Hawkes process and point processes are discussed. Section 4 includes the definitions of the point process, the conditional event intensity, the definitions of the Hawkes process, kernel selection, the likelihood function and the maximum likelihood estimation for the Hawkes process, and the expected amount of events in the near future. Then section 5 presents the result of modified Hawkes process models with different kernels applied on England and Wales suicide data. I use a stationary Poisson process as a comparison for the Hawkes process. In section 6, I give some preventative suggestions after the suicide prediction obtained from the Hawkes processes. In the end, a conclusion and some discussions are proposed.

# CHAPTER 2

# Introduction

Hawkes process model, as one of the most well-known types of point processes, provides the statistical language to describe the timing and properties of various types of events. Problems from a wide range of areas fit such a setting. Examples [Rei18] include incidence of disease, sightings or births of a species, occurrences of fires, earthquakes, lightning strikes, tsunamis, or volcanic eruptions.

Hawkes process naturally acquires the triggering and clustering behaviors within each context. In the analysis of the suicide issue, Hawkes process has the advantage of capturing the copycat behavior (triggering behavior) or the imitation effect (clustering behavior) within suicide cases along the timeline. According to the descriptions from Wikipedia, a copycat suicide is defined as an emulation of another suicide that the person attempting suicide knows about either from local knowledge or due to accounts or depictions of the original suicide on television and in other media. The publicized suicide works as a trigger, in the absence of protective factors, for the next suicide by a susceptible or suggestible person, also referred to as suicide contagion. Here, each celebrity suicidal case, as the publicized event, works as the immigrant event, which happens at a certain time in a continuous time frame. Each of them has a set of properties, such as celebrity's influence or popularity, specific occupation, gender, connectivity of the surrounding industries, etc. Introduced by Reinhart [Rei18], Hawkes processes model events whose rate depends on the past history of the process. The high-profile celebrity suicide cases serve as the significant past history, and more suicides afterward are assumed to be related, as a result of suicide contagion. Then I explicitly observe the offspring suicide cases as celebrity following or obsession behaviors, and I desire to model those discrete, inter-dependent suicidal events over continuous time via the Hawkes process model.

Usually, suicide is defined as death caused by self-directed injurious behavior with intent to die as a result of the behavior [NIM]. At first, I intended to do such a study on the dataset of the United States suicides. The number of deaths due to intentional self-harm per 100,000 population in the United States keeps rising in the most recent decade (2009-2018) [NIM]. According to the Centers for Disease Control and Prevention (CDC) Leading Causes of Death Reports, shown in Figure 2.1, in the United States, suicide is the top leading cause of death among several age groups [CDCb]. The number of suicides (about 48,334 in 2018) is more than 2.5 times as many as that of homicides (about 18,830 in 2018). The crude rates of suicides in 2018 in geographic areas are presented as well in Figure 2.2 [CDCb]. The rates of some counties are missing, leaving them blank. Plenty of counties are filled with brown color, and they are widely spread over the United States. Hence, the United States is a region of great importance to study suicide rates.

**10 Leading Causes of Death, United States**

2018, Both Sexes, All Ages, All Races

| | <1 | 1-4 | 5-9 | 10-14 | 15-24 | 25-34 | 35-44 | 45-54 | 55-64 | 65+ | All Ages |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Congenital Anomalies 4,473 | Unintentional Injury 1,226 | Unintentional Injury 734 | Unintentional Injury 692 | Unintentional Injury 12,044 | Unintentional Injury 24,614 | Unintentional Injury 22,667 | Malignant Neoplasms 37,301 | Malignant Neoplasms 113,947 | Heart Disease 526,509 | Heart Disease 655,381 |
| 2 | Short Gestation 3,679 | Congenital Anomalies 384 | Malignant Neoplasms 393 | Suicide 596 | Suicide 6,211 | Suicide 8,020 | Malignant Neoplasms 10,640 | Heart Disease 32,220 | Heart Disease 81,042 | Malignant Neoplasms 431,102 | Malignant Neoplasms 599,274 |
| 3 | Maternal Pregnancy Comp. 1,358 | Homicide 353 | Congenital Anomalies 201 | Malignant Neoplasms 450 | Homicide 4,607 | Homicide 5,234 | Heart Disease 10,532 | Unintentional Injury 23,056 | Unintentional Injury 23,693 | Chronic Low. Respiratory Disease 135,560 | Unintentional Injury 167,127 |
| 4 | Sids 1,334 | Malignant Neoplasms 326 | Homicide 121 | Congenital Anomalies 172 | Malignant Neoplasms 1,371 | Malignant Neoplasms 3,684 | Suicide 7,521 | Suicide 8,345 | Chronic Low. Respiratory Disease 18,804 | Cerebrovascular 127,244 | Chronic Low. Respiratory Disease 159,486 |
| 5 | Unintentional Injury 1,168 | Influenza & Pneumonia 122 | Influenza & Pneumonia 71 | Homicide 168 | Heart Disease 905 | Heart Disease 3,561 | Homicide 3,304 | Liver Disease 8,157 | Diabetes Mellitus 14,941 | Alzheimer's Disease 120,658 | Cerebrovascular 147,810 |
| 6 | Placenta Cord Membranes 724 | Heart Disease 115 | Chronic Low. Respiratory Disease | Heart Disease 101 | Congenital Anomalies 354 | Liver Disease 1,008 | Liver Disease 3,108 | Diabetes Mellitus 6,414 | Liver Disease 13,945 | Diabetes Mellitus 60,182 | Alzheimer's Disease 122,019 |
| 7 | Bacterial Sepsis 579 | Perinatal Period 62 | Heart Disease 68 | Chronic Low. Respiratory Disease 64 | Diabetes Mellitus 246 | Diabetes Mellitus 837 | Diabetes Mellitus 2,282 | Cerebrovascular 5,128 | Cerebrovascular 12,789 | Unintentional Injury 57,213 | Diabetes Mellitus 84,946 |
| 8 | Circulatory System Disease 428 | Septicemia 54 | Cerebrovascular | Cerebrovascular 54 | Influenza & Pneumonia 200 | Cerebrovascular 567 | Cerebrovascular 1,704 | Chronic Low. Respiratory Disease 3,807 | Suicide 8,540 | Influenza & Pneumonia 48,888 | Influenza & Pneumonia 59,120 |
| 9 | Respiratory Distress 390 | Chronic Low. Respiratory Disease 50 | Septicemia 34 | Influenza & Pneumonia 51 | Chronic Low. Respiratory Disease 165 | Hiv 482 | Influenza & Pneumonia 956 | Septicemia 2,380 | Septicemia 5,956 | Nephritis 42,232 | Nephritis 51,386 |
| 10 | Neonatal Hemorrhage 375 | Cerebrovascular 43 | Benign Neoplasms 19** | Benign Neoplasms 30 | Complicated Pregnancy 151 | Influenza & Pneumonia 457 | Septicemia 829 | Influenza & Pneumonia 2,339 | Influenza & Pneumonia 5,858 | Parkinson's Disease 32,988 | Suicide 48,344 |

Figure 2.1: Top 10 leading causes of Death for both genders, all age groups in the United States in 2018

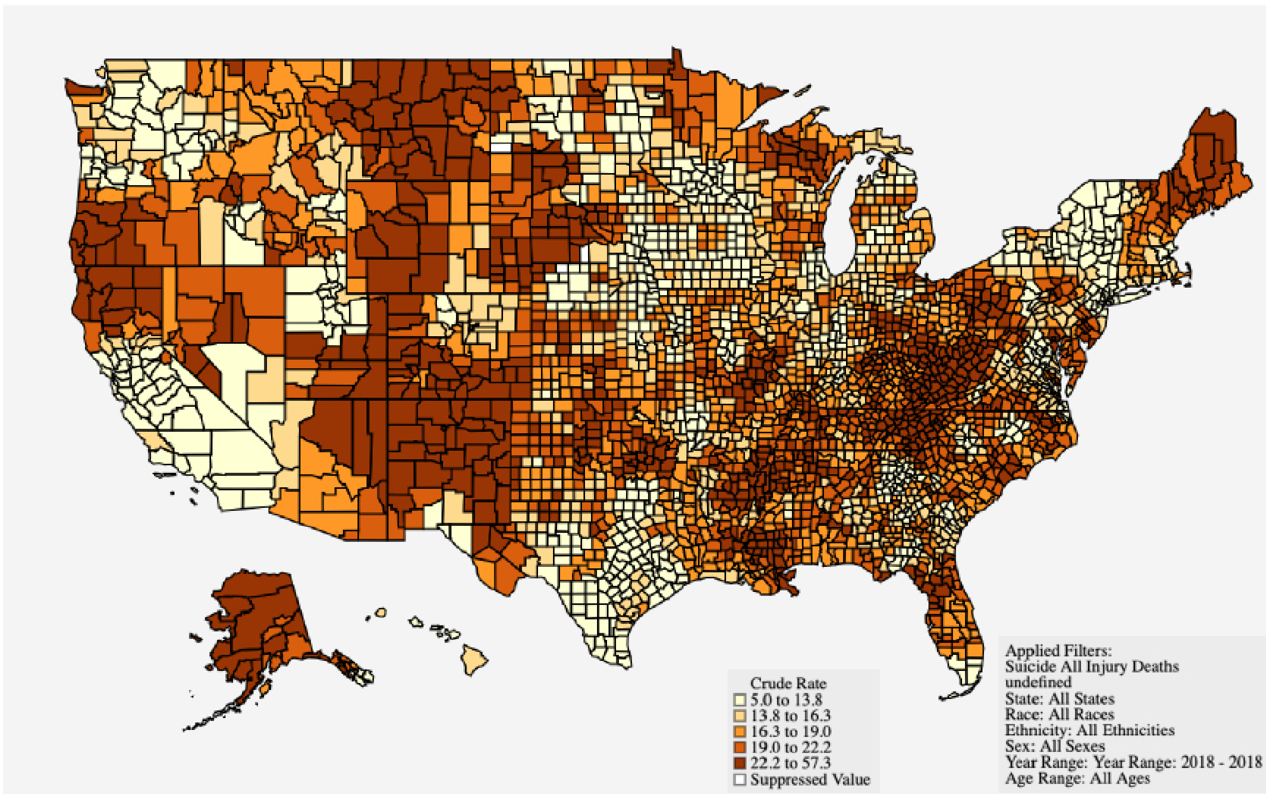


Figure 2.2: The county-level crude rates of suicides in the United States in 2018

Yet, no daily suicide data from the United States is available, neither the weekly data. I tried to use the weekday-level data and the monthly data from CDC WONDER [CDCa], but both did not have enough detailed information for me to explore the clustering or the triggering effect. The red squared points and bold black dots represent the celebrity (publicized) suicide cases in Figure 2.3 and Figure 2.4 respectively. Even though there exists a general upward trend of the number of suicidal deaths in both plots, these special points are not always followed by a spiking number, and some are even at the local maxima, which should not happen. As both datasets are of cumulative numbers, it's not reliable to figure out the underlying or specific relationship between the background points and offspring points through those data. Thus, the daily suicide data is of great importance.

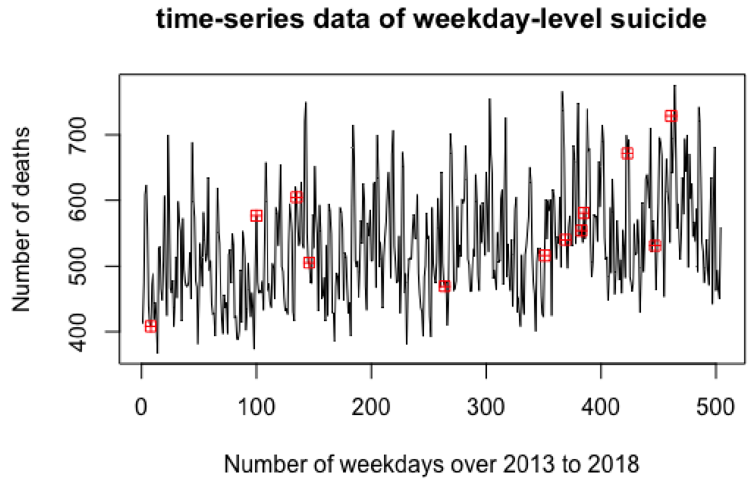


Figure 2.3: The suicide sizes for all weekdays within 2013 to 2018

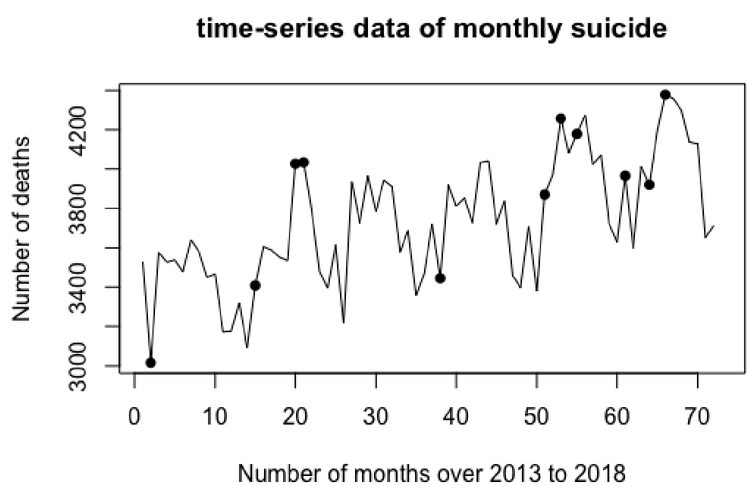


Figure 2.4: The suicide sizes for all months within 2013 to 20188

# CHAPTER 3

## Data Description

Fortunately, the daily suicide data is available on the United Kingdom government website [ONS], Office for National Statistics (ONS). It is about suicides in England and Wales from the years 2001 to 2015. The definition of suicide in England and Wales is a little different from the data collection perspective. The suicide that happened in England and Wales is defined as deaths given an underlying cause of intentional self-harm (for people aged 10 and over) or injury/poisoning of undetermined intent (for people aged 15 and over). The suicide rates continue to have steady slow growth, Iacobucci et al. [Iac20] concluded. Figure 3.1 below shows the age-adjusted suicide rates by regions of England and Wales in 2014 [ONS].

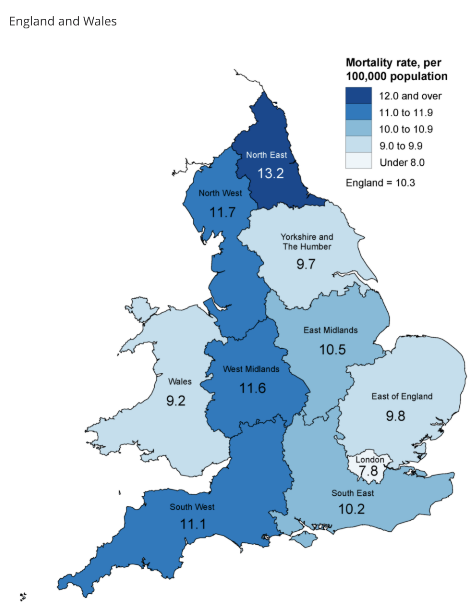


Figure 3.1: The crude suicide rates for regions in England and Wales

Other than that the North East region has the highest mortality rate, over 12, and that London has the lowest rate of 7.8, the rest of the regions basically have similar rates, which means the spread of suicide cases is relatively evenly distributed. I chose the suicides occurring each day between 2010 and 2015, these six years, to study the influence of publicized suicide. Meanwhile, I gathered the well-known suicide cases reported on BBC in the United Kingdom over those years. This dataset includes the celebrity name, death date, occupation, and influence (followers). The information is listed in Table 3.1 on the next page.

Among those well-known celebrity mortalities, 3 out of 10 deaths are related to professional football clubs, and 6 are from the entertainment industry. Their influence is calculated approximately based on the number of followers or fans on social networking platforms such as Instagram, Twitter, etc.

Moreover, the influence of the common people is hard to track down through online social media sites. Also, the dataset obtained does not contain the specific identity of persons who committed suicide, maybe due to privacy protection and public safety concern. Therefore, such information is estimated by the method developed by McCormick et al. [MSZ10], to effectively estimate the personal network size. The log-normal distribution with $\mu = 6.2$ and $\sigma = 0.68$ seems to capture the observed network size data and its posterior simulation pretty well. Then, this method is applied as the influence statistics added to this suicide dataset.

| Name | Death Time | Influence (approximation) | Occupation |
|---|---|---|---|
| Alexander McQueen | 2/11/10 | 1,080,000 | English fashion designer and couturier |
| Charles Haddon | 8/20/10 | 26,600 | Singer, frontman |
| Terry Newton | 9/26/10 | 93,300 | English professional league footballer |
| Dale Roberts | 12/14/10 | 277,300 | English footballer, goalkeeper |
| Angela Scoular | 4/11/11 | 21,700 | Actress |
| Gary Speed | 11/27/11 | 42,900 | Professional footballer and manager |
| Tony Scott | 8/19/12 | 26,700 | Movie director, producer, and screenwriter |
| Paul Bhattacharjee | 7/10/13 | 130,300 | Actor |
| Lil' Chris | 3/23/15 | 29,900 | Singer-songwriter, actor, and television personality |
| Sam Sarpong | 10/26/15 | 15,000 | Actor, supermodel, and musician |

Table 3.1: Publicized suicides in England and Wales over 2010-2015

# CHAPTER 4

# Key Concepts in Hawkes Process (Point Process)

Here I present some informal explanations for the ideas and theories behind the Hawkes processes.

## 4.1 Definition of Point Process

A point process is a random collection of points falling in some metric space. For a spatial-temporal point process, the metric space is a portion of space-time, $S = \mathbb{R}^d \times \mathbb{R}$. Point processes share three main characteristics. First, point processes are stochastic processes on the non-negative real line. That is, a point process can be defined as any non-decreasing, non-negative valued stochastic process. Second, point processes are a list of points, a finite collection of points. Time $t_i \geq 0$ , and $i$ takes integer values $1, 2, \cdots$, and $t_i < t_{i+1}$. Normally, $t_i < \infty$, which is measurable. Let $N(T)$ be the accumulation of the number of points up to time $T$, i.e. the number of events of the point process by time $T$.

$$N(T) := \sum_i \mathbb{1}_{\{t_i < T\}}.$$

$N(T)$ is a piece-wise function, and $N(0) = 0$.

The third feature is the random measure. A real line valued random measure includes a wide range of processes on the line and extends readily to space-time. The measure $N(I)$ represents the number of points falling in the region $I$ of space-time.

## 4.2 Event Intensity

The event intensity $\lambda$ is the limited expected rate of points accumulated over a specific time and spatial interval, given all points before $t$. It resembles the density function.

$$\lambda(t, x, y) = \lim_{t,\delta \to 0} E(N([t, t + \Delta t] \times B[x, y, \delta])|\mathcal{H}_t)/(\Delta t \pi \delta^2), \tag{4.1}$$

where the $B[x, y, \delta]$ is a circle of radius $\delta$ around point $(x, y)$, and $\mathcal{H}_t$ is the history of the process up to but not including time $t$.

Note that $\lambda$ is random, depending on what points have occurred before, and $\lambda$ might be different with every realization. $\lambda$ is predictable, meaning that although sometimes the number of events happening is unknown, the number of events expected to happen could be figured out. Usually, the total expected number of events $E[N(I)]$ would be the summation or integration of event intensity over some fixed spatial-time interval $I$.

$$E[N(I)] = \int_I \lambda(t, x, y) dt dx dy. \tag{4.2}$$

To understand more about the event intensity, it is informative to have a brief description of the Poisson process as an example, the most basic and simple type of point process. If $N$ is a simple point process with conditional event intensity $\lambda$, where $\lambda$ does not depend on what points have occurred previously, then $N$ is the Poisson process. $\lambda$ is deterministic in exhibitions of point processes. Of course, there are other point processes, more complicated point processes, that have conditional intensity with different features. In the stationary Poisson process, $\lambda$ is the constant for every subspace, but events are completely randomly distributed over the fixed space. In the mixed Poisson process, $\lambda$ is equal to some random variable $c$, which $c$ could be an exponential random variable, be half-normal distributed, or be something constrained to be positive.

## 4.3 Hawkes Process

Usually, a Hawkes process, i.e. self-exciting process, has the conditional intensity as

$$\lambda(t, x, y) = \mu(x, y) + \kappa \int_{t'<t} g(t - t', x - x', y - y') dN(t', x', y') \tag{4.3}$$

or

$$\lambda(t, x, y) = \mu(x, y) + \kappa \sum_{(t',x',y':t'<t)} g(t - t', x - x', y - y'), \tag{4.4}$$

where $g$ is called the triggering function or triggering density, and $\kappa$ is the productivity or branching factor.

If $g$ is a density function, then $\kappa$ is the expected number of points triggered directly by each point. Each base point, associated with $\mu(x, y)$, is expected to generate

$$\kappa + \kappa^2 + \kappa^3 + \cdots = \frac{1}{1-\kappa} - 1 \tag{4.5}$$

triggered points, so the expected fraction of background points is $1 - \kappa$.

A suicide copycat behavior, as described before, is an imitation of the suicide mechanism or death method of a publicized suicidal case. A suicide cascade is viewed as numerous suicide cases that occurred after an initial celebrity suicide. Using the language of self-exciting processes, the suicide cascade, as a branching structure, consists of parent events clustered by offspring events. Besides, the dataset collected includes time as a central dimension, prompted by Rizoiu et al. [RLM17], so similarly, the conditional intensity function of my Hawkes process is

$$\lambda(t|\mathcal{H}_t) = \lambda_0(t) + \sum_{t_i<t} \phi_{m_i}(t - t_i), \tag{4.6}$$

where $\lambda_0(t) : \mathbb{R} \to \mathbb{R}_+$ is a conclusive background intensity function, and $\phi : \mathbb{R} \to \mathbb{R}_+$ is called the memory kernel. $\mathcal{H}_t$ is the associated history $t \geq 0$ for the process $N(t)$, and $t_i < t$. Obviously, the Hawkes model is a non-homogeneous Poisson process with stochastic intensity and mainly depends on the kernel function $\phi_{m_i}(\cdot)$ which associates with previous events. $\phi_{m_i}(\cdot)$ represents the kernel function associated with the event $i$ that has the estimated influence of $m_i$.

The base intensity function $\lambda_0(t)$ represents the rate of events arising from external sources, which is estimated by smoothing observed famous suicidal events. Those events are known as background or immigrant events, and they are assumed to be independent of the events that happened earlier within the process. Furthermore, this model is a marked Hawkes process. The mark $m$ is the personal influence or estimated network size for every event.

## 4.4 Choice of Kernel

The kernel function $\phi(t - t_i)$ aim to track how the event arrived at time $t_i$ would affect the intensity function at time $t$. Under most circumstances, the kernel function is considered to be monotonically decreasing along the timeline. The more recent the events are, the more influence they would have on the current case, and the higher the event intensity would be. Thus, the intensity function is assumed to diminish as the events are further away in time scale.

Naturally, the influence of a suicide case decays as time passes. The most common corresponding kernel functions in records are exponential kernel and power-law kernel. $x$ is positive in both kernels.

The exponential kernel is

$$\phi(x) = \alpha e^{-\delta x},$$

where $\alpha \geq 0$, $\delta > 0$, and $\alpha < \delta$.

The power-law kernel is

$$\phi(x) = \frac{\alpha}{(x + \delta)^{\eta+1}},$$

where $\alpha \geq 0$, $\delta > 0$, $\eta > 0$.

For the application to the suicide dataset, the power-law kernel $\phi_m(t_{i+1} - t_i)$ with mark $m$ is constructed as:

$$\phi_m(t_{i+1} - t_i) = \kappa m^{\beta}(t_{i+1} - t_i + c)^{-(1+\theta)}, \tag{4.7}$$

$\kappa$ is the branching factor or the productivity measurement. This quantity describes the Hawkes process through the reproductive perspective. It could be viewed as the expected rate of offspring events triggered by a single parent event or an offspring event. The occurrence of parent events is associated with the base event intensity $\lambda_0(t)$. The productivity scales the subsequent cluster of children events in the process. $\beta$ concludes the wrapping effect for the influence in networks. $(1 + \theta)$ describes the decay process. $c > 0$ is to make sure the term $\phi_m(\cdot)$ is bounded. In general, $\kappa m^{\beta}$ models the overall magnitude of influence, and $(t_{i+1} - t_i + c)^{-(1+\theta)}$ accounts for the memory over time. The personal influence $m$ is assumed to be the number of followers from social media sites or the size of the estimated personal network.

Similarly, the exponential kernel function constructed for this issue:

$$\phi_m(t_{i+1} - t_i) = \kappa m^{\beta}\theta e^{-\theta(t_{i+1} - t_i + c)}. \tag{4.8}$$

Figure 4.1 describes shape of power-law kernel generated by an initialization event of mark $m = 1000$, at time $t = 0$, and parameters defined to be $\kappa = 0.768$, $\beta = 0.68$, $c = 10.45$, $\theta = 0.788$.

Figure 4.2 is the visualization of shape of exponential kernel generated by an initialization event of mark $m = 1000$, at time $t = 0$, with parameters set to be $\kappa = 0.2$, $\beta = 0.3$, $c = 2$, $\theta = 0.5$.

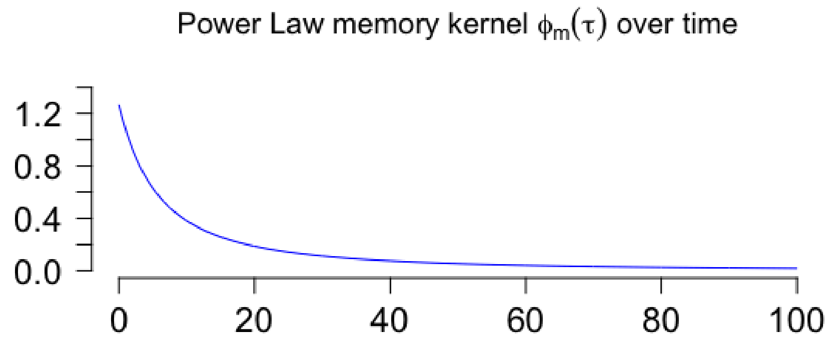


Figure 4.1: Example power-law kernel with parameters $\kappa = 0.768$, $\beta = 0.68$, $c = 10.45$, $\theta = 0.788$

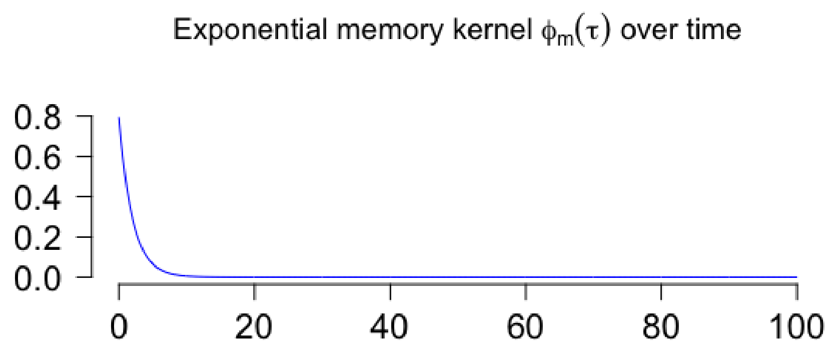


Figure 4.2: Example exponential kernel with parameters $\kappa = 0.2$, $\beta = 0.3$, $c = 2$, $\theta = 0.5$

## 4.5 Likelihood Function and Maximum Likelihood Estimation

For a nonstationary Poisson process with intensity $\lambda(\cdot)$, on $[0, T]$, the likelihood of observing the points $\{t_1, t_2, \cdots, t_n\}$, is simply

$\lambda(t_1) \times \lambda(t_2) \times \cdots \times \lambda(t_n) \times \exp(-A(t_1)) \times \exp(-A(t_2 - t_1)) \times \cdots \times \exp(-A(T - t_n))$

$= \prod_{i=1}^{n} \lambda(t_i) \cdot \exp[-A(T)]$, where $A(T) = \int_0^T \lambda(t)dt$.

Then, in point process, the general log-likelihood function is simply

$$\sum_{i=1}^{n} \log(\lambda(t_i)) - \int \lambda(t)dt.$$

Under the Hawkes process, the likelihood function $L$ with parameter set $\Theta$ is

$$L(\Theta) = \prod_{i=1}^{n} \lambda(t_i) \exp\left(-\int_0^{t_n} \lambda(t)dt\right). \tag{4.9}$$

Accordingly, the log of likelihood function is

$$\ell(\theta) = \log L(\Theta) = -\int_0^{t_n} \lambda(t)dt + \sum_{i=1}^{n} \log \lambda(t_i). \tag{4.10}$$

The marked Hawkes process models have a parameter set of $\Theta = \{\kappa, \beta, c, \theta\}$. By incorporating the intensity function formula (4.6) and kernel function (4.7) into the Hawkes log likelihood function (4.10), the log likelihood function for this marked Hawkes process is

$$\ell(\theta) = \sum_{i=2}^{n} \log \kappa + \sum_{i=2}^{n} \log\left(\sum_{t_i < t_j} \frac{{m_i}^{\beta}}{(t_j - t_i + c)^{-\theta}}\right) - \kappa \sum_{i=1}^{n} (m_i)^{\beta} \left[\frac{1}{\theta c^{\theta}} - \frac{(t + c - t_i)^{-\theta}}{\theta}\right]. \tag{4.11}$$

To make sure the branching ratio to be positive and meaningful, a few constraints on the parameters exist: $\theta > 0$, $\kappa > 0$, $c > 0$, $0 < \beta < \alpha - 1$.

## 4.6 The Expected Number of Future Offspring Events

Equation (4.5) shows the expected portion of the cluster of children events associated with a new background event. Equation (4.2) gives the expected total number of events over spatial-time temporal space $I$. Given those formulas, then, the number of future children events is

$$N = \int_T^{\infty} \lambda(t)dt.$$

In the marked point process, the total number of future events spawned by event $i = 1, 2, \cdots, n$, by introducing its conditional intensity function (Equation (4.6)) and kernel function (Equation (4.7)), would be achieved by the integral of kernel function of events over time $T > t_i$.

$$N = \kappa \sum_{i=1}^{n} \frac{{m_i}^{\beta}}{\theta(T + c - t_i)^{\theta}}. \tag{4.12}$$

# CHAPTER 5

# Hawkes Process Models Fitted to Suicide Data of England and Wales

The trimmed daily suicide dataset of England and Wales consisted of 2150 days, starting on the day of February 11th, 2010. The initial background event is the self-harm death of Alexander McQueen, which is assumed to have occurred at time $t = 0$. The other celebrity suicide cases are treated as immigrant points as well along the time to the end of 2015. Influence is the magnitude of estimated social network size. Equation (4.6) with kernel functions (4.7) and (4.8) are applied to this data to record the suicide cascade as a point process. Figure 5.1 plots the suicide cascade (as defined before, suicide cascade is numerous suicide cases occurred after an initial celebrity suicide; it is observed data) as a sequence of the event for the first 1000 days as a representation. Because of the limitation of the length of plots and the requirement for clarity, 1000 events are used as an example to show the decay and branching structure. The first 7 publicized suicide cases are included over these days, and they are approximately the highest seven points in the figure.
Figure 5.2 is the corresponding event intensity function.

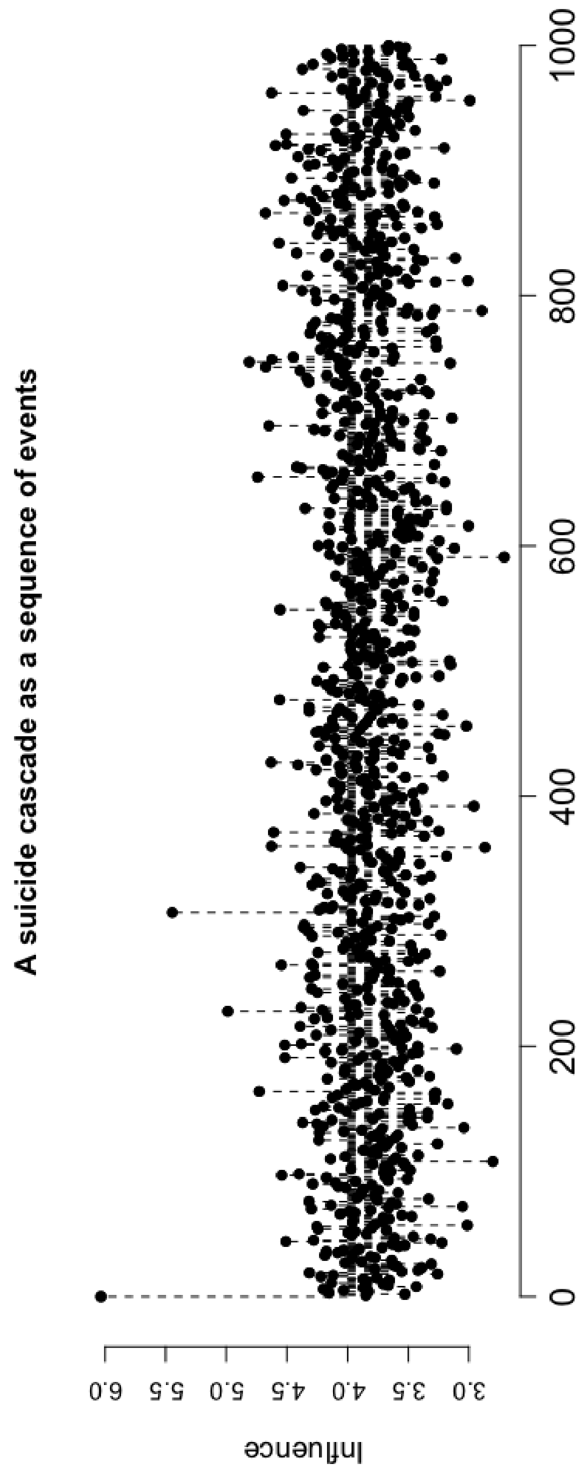


Figure 5.1: The suicide cascade using power-law kernel for the first thousand events since 2/11/2010 in England and Wales

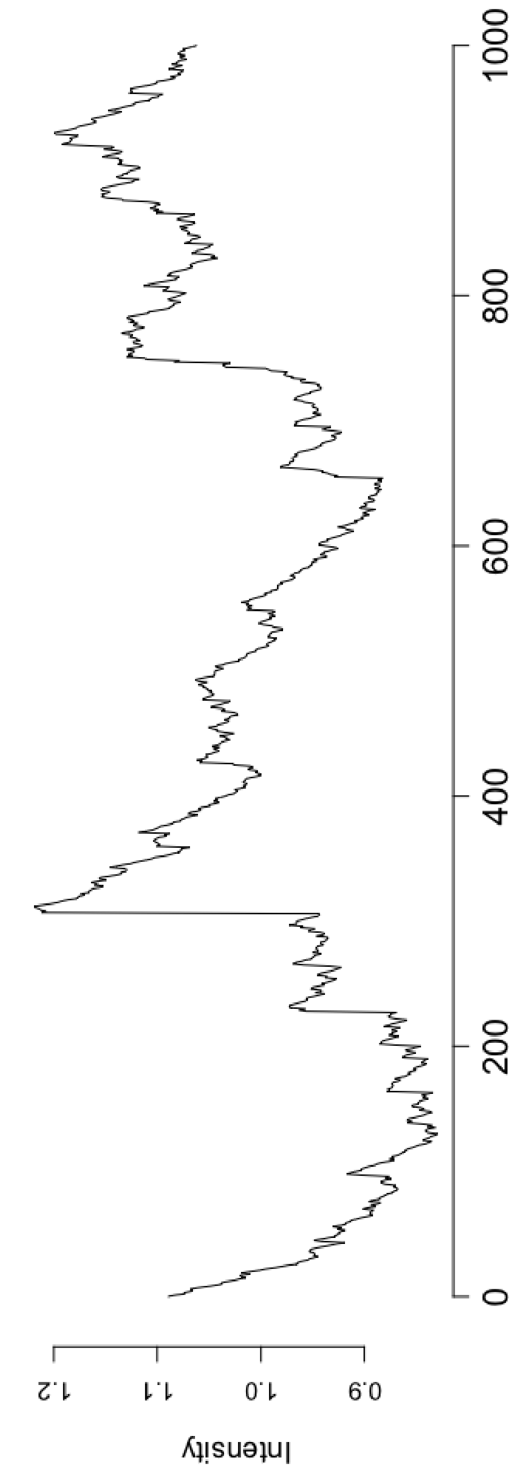


Figure 5.2: The intensity function using power-law kernel for the first thousand events since 2/11/2010 in England and Wales

These two figures appear to be aligned. Each background event seems to lead to a noticeable jump in the intensity function. It is easy to observe 7 local maxima points followed by a relatively rapid decay. Meanwhile, the productivity effect also shows; there exist clusters of large influences after the highest seven points.

There is another decay kernel function, the exponential decay. Plots are Figure 5.3 and Figure 5.4. Those exhibit similar features, but the intensity function shows some different fluctuations.

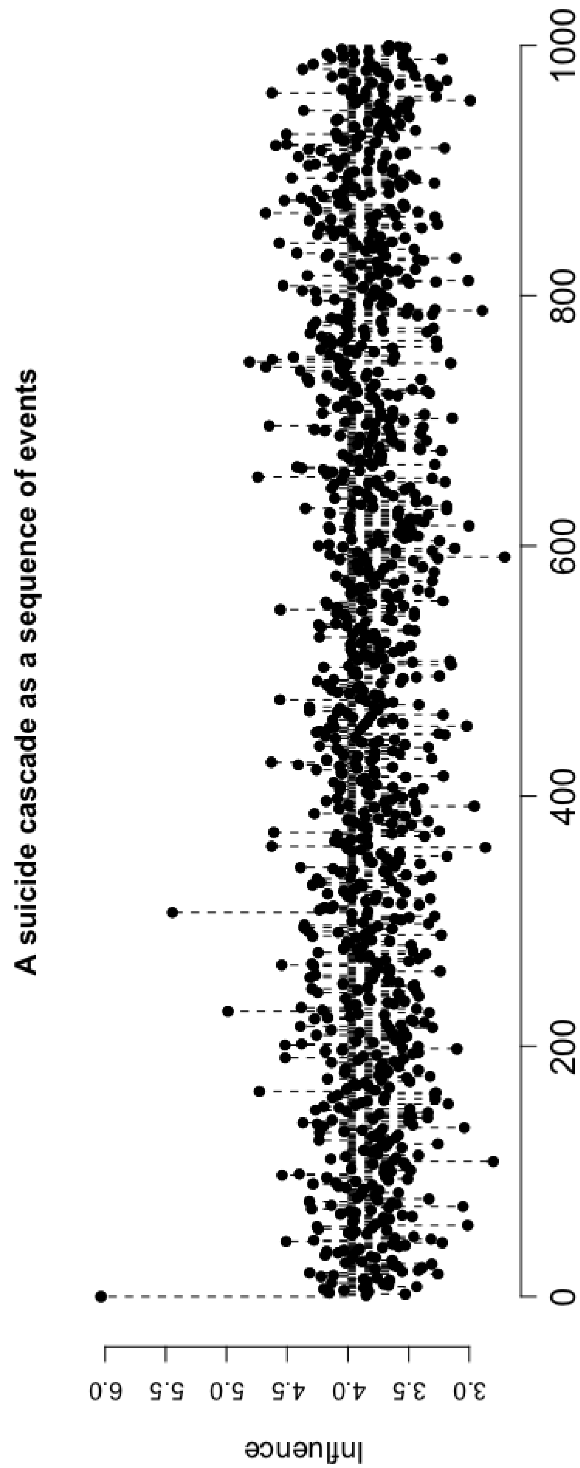


Figure 5.3: The suicide cascade using exponential kernel for the first thousand events since 2/11/2010 in England and Wales

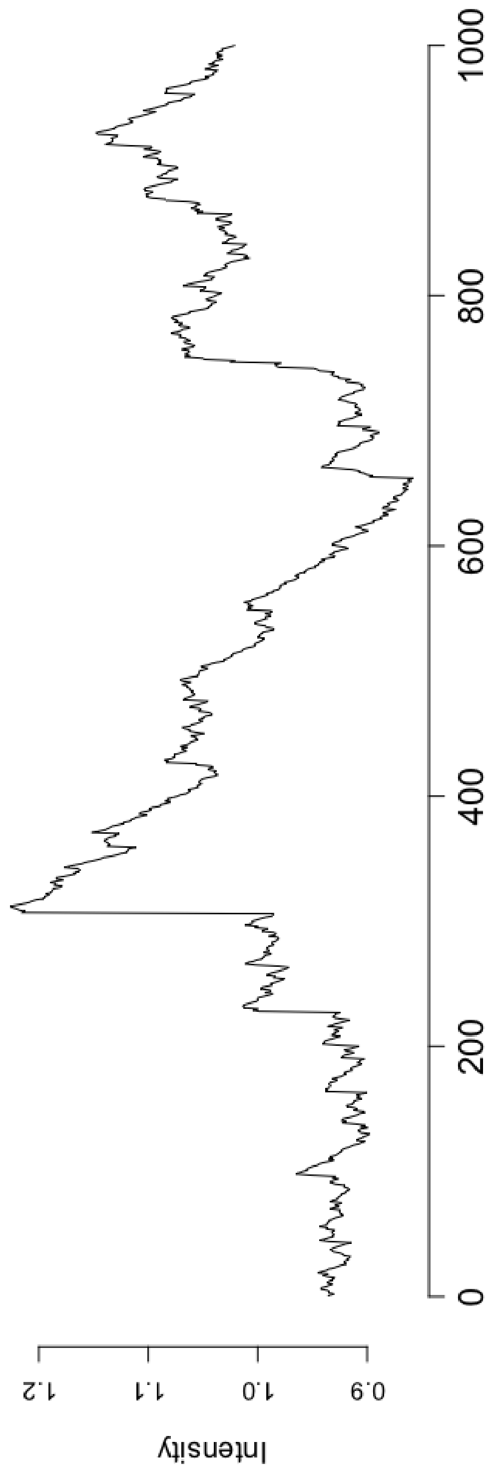


Figure 5.4: The intensity function using exponential kernel for the first thousand events since 2/11/2010 in England and Wales

Figures of using power-law kernel as the social kernel are plotted based on the estimation of maximum likelihood. The maximum likelihood estimates $\Theta_{\text{power-law}}$ with standard errors of Hawkes process model for all the events within the time frame are

| Parameter | $\kappa$ | $\beta$ | $c$ | $\theta$ |
|---|---|---|---|---|
| Estimate | 1 | 1.0158 | 216.4963 | 1.6084 |
| S.E. | 0.7916 | 52.9433 | 12.8610 | 0.2145 |

Table 5.1: Maximum likelihood estimates and their standard errors for power-law kernel marked Hawkes process

The estimated number of future offspring events triggered, given Equation (4.12), is calculated

$$N_{\text{power-law}} = 170.332.$$

The predicted suicide size of 34480, and the real size is 28943. The relative error in percentage is 19.13. The model's AIC value is 349.38, and its log likelihood estimation value is -169.69.

Figure 5.3 and Figure 5.4 are based on results of the maximum likelihood estimation of Hawkes process using exponential kernel as the social kernel. The estimates $\Theta_{\text{exponential}}$ are

| Parameter | $\kappa$ | $\beta$ | $c$ | $\theta$ |
|---|---|---|---|---|
| Estimate | 0.00009 | 1.0158 | 0 | 0.007079 |
| S.E. | 11.9818 | 0.2073 | 49.9383 | 5.8447 |

Table 5.2: Maximum likelihood estimates and their standard errors under the exponential kernel marked Hawkes process

The estimated number of triggered future children events is

$$N_{\text{exponential}} = 144.204.$$

The predicted volume of events is 210713, with a relative error in percentage of 628.03. The AIC value here is 453.96, and the estimated log-likelihood value is -221.98.

The estimated number of triggered future children events is

$$N_{\text{exponential}} = 144.204.$$

The predicted volume of events is 210713, with a relative error in percentage of 628.03. The AIC value here is 453.96, and the estimated log-likelihood value is -221.98.

Additionally, a stationary Poisson process is conducted for comparison. No matter the distribution of the events is continuous or discrete, modeling the inter-event time $u_i = t_i - t_{i-1}$ using a continuous distribution, the Exponential distribution, is applicable under the Poisson process. With the conditional intensity $\lambda$, the inter-arrival time of consecutive events $u_1, u_2, \cdots, u_n$ are i.i.d exponentially distributed with mean $\frac{1}{\lambda}$. It resembles time series.

$$\lambda(u_i) = b_1 + b_s(x_i, y_i) + b_3(u_i),$$

where $b_1$ is the intercept, $b_2$ and $b_3$ conclude the magnitude of effect of location or time respectively.
The AIC for this model is 436.82, and the log-likelihood value is about -215.41.

Hence, compared to the stationary Poisson process model, this marked Hawkes process model with exponential kernel does not appear to be a very good fit for the suicide data of England and Wales, even if it appears to capture the decaying effect and clustering of the triggered offspring points. The marked Hawkes process with power-law kernel seems to be more credible with a much smaller relative error in percentage, smaller AIC value, and a larger log-likelihood estimation value, which may indicate this model is a relatively good fit.

# CHAPTER 6

# Summary

## 6.1 Prevention

With the estimation of spawned future children events, a result of suicide copycat behavior, obsession, or both, Alys Cole-King and Stephen Platt [CP17] suggested the suicide risks can be controlled. This course includes identification, intervention, and mitigation. The assessment of a person after self-harm or suicide attempt matters. Discussions about suicide with patients afterward are potentially life-saving. Empathetic and compassionate health-care professionals encourage is necessary for patients to disclosure their thoughts. Besides, making a safety plan, like articulating reasons for living, creating a safe environment, seeking professional support, etc., could mitigate those risks. Widespread dissemination of mental health services and mental health organizations could encourage people at risk to seek assistance, which would be meaningful. Moreover, there are shreds of evidence collected by Fink et al. [FSK18] about subsequent suicide events as a result of the celebrity suicide and its subsequent media reports. Therefore, the government could take responsibility for establishing more conservative and restrictive regulations on the media coverage of high-profile celebrity mortality cases.

## 6.2 Conclusion and Discussion

Give all the information provided before, this marked Hawkes process model seems to be useful to some extent, to improve understanding of the process generating suicides and to predict the new events in the near future. I discerned that there appears to be some clustering of offspring events (branching effect) in the spatial-temporal sense among incidences of suicide in England and Wales over 2010 to 2015, by observation of both raw data and via the fitted models. When I tried to fit the Hawkes processes with different kernels to the data, I found that the model with the power-law kernel fits the data better than the model with the exponential kernel does. The productivity ratio (reproduction term) is kind of large; it could prove Hawkes process captures triggering behavior and clustering behavior relatively well.

Both marked Hakes models provided close numbers of future offspring events. It appears to be able to explain the phenomena that background events are clustered by offspring points in the long time interval, but are unable to fully account for and explain the subtle gathering of suicides without a high profile celebrity suicide that takes place over a short period, maybe due to the macro-level of the social environment. I explicitly observed the offspring suicide case via celebrity obsession, however, there are other reasons or social mechanisms that are not tracked or recorded.

There are more approaches and aspects to explore this topic that I would plan to further analyze in future exploration. As I only fit two different kernels on marked Hawkes process to the data, I could fit more different models, such as the Hawkes model with covariates, Poisson cluster process, etc, to find out if there are models that fit the data better, especially those considering the nature of the data.

Another notable aspect to think about in the future, given the large $\kappa$ value in the marked Hawkes process, is to analyze a more detailed, valuable suicide dataset for the United States, especially for time recorded in minutes or even seconds. The United States seems to have more complex and severe suicide concerns. The Hawkes process mainly focused on the time dimension that could predict the timing of suicide cascade in near future with high credibility. Then, government institutions and non-profit associations could be prepared for that, make predictive prevention efforts, and maintain public order with efficiency. Last, although it is difficult to achieve because of privacy protection concerns, finding more precise and complete datasets that include more details of each suicide event, like the exact information of location, race, gender, and family history of each suicide committed is also crucial, which would help approximate the covariates or other factors to rule out many noises.

REFERENCES


[CDCa] "Underlying Cause of Death, 1999-2019 Request." `http://wonder.cdc.gov/ucd-icd10.html`. Accessed: 2021-03-22.

[CDCb] "WISQARS Data Visualization EXPLORE FATAL INJURY DATA VISUALIZATION TOOL." `https://wisqars-viz.cdc.gov:8006/lcd/home?lcd=eyJjYXVzZXMiOlsiQ3D`. Accessed: 2021-03-22.

[CLM20] Wen-Hao Chiang, Xueying Liu, and George Mohler. "Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates." *medRxiv*, 2020.

[CMD20] G. Castillo-Sánchez, Gonçalo Marques, E. Dorronzoro, Octavio Rivera-Romero, M. Franco-Martín, and I. de la Torre-Díez. "Suicide Risk Assessment Using Machine Learning and Social Networks: a Scoping Review." *Journal of Medical Systems*, **44**, 2020.

[CP17] Alys Cole-King and Stephen Platt. "Suicide prevention for physicians: identification, intervention and mitigation of risk." *Medicine*, **45**(3):131–134, 2017. Acute Medicine Part 2 of 2.

[FSK18] David S. Fink, Julian Santaella-Tenorio, and Katherine M. Keyes. "Increase in suicides the months after the death of Robin Williams in the US." *PLOS ONE*, **13**(2):1–12, 02 2018.

[HRG21] Ayaan Haque, V. Reddi, and Tyler Giallanza. "Deep Learning for Suicide and Depression Identification with Unsupervised Label Correction." *ArXiv*, **abs/2102.09427**, 2021.

[Iac20] Gareth Iacobucci. "Suicide rates continue to rise in England and Wales." *BMJ*, **370**, 2020.

[MCR18] George Mohler, Jeremy Carter, and Rajeev Raje. "Improving social harm indices with a modulated Hawkes process." *International Journal of Forecasting*, **34**(3):431–439, 2018.

[MSZ10] Tyler H. McCormick, Matthew J. Salganik, and Tian Zheng. "How Many People Do You Know?: Efficiently Estimating Personal Network Size." *Journal of the American Statistical Association*, **105**(489):59–70, 2010. PMID: 23729943.

[NFY12] Thomas Niederkrotenthaler, King-wa Fu, Paul S F Yip, Daniel Y T Fong, Steven Stack, Qijin Cheng, and Jane Pirkis. "Changes in suicide rates following media reports on celebrity suicide: a meta-analysis." *Journal of Epidemiology & Community Health*, **66**(11):1037–1042, 2012.

[NIM] "Suicide Statistics." `https://www.nimh.nih.gov/health/statistics/suicide`. Accessed: 2021-02-26.

[ONS] "Number of suicides by date of death." `https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/adhocs/008141numberofsuicidesbydateofdeath`. Accessed: 2021-04-12.

[Rei18] Alex Reinhart. "A Review of Self-Exciting Spatio-Temporal Point Processes and Their Applications." *Statistical Science*, **33**(3), Aug 2018.

[RLM17] Marian-Andrei Rizoiu, Y. Lee, Swapnil Mishra, and Lexing Xie. "A Tutorial on Hawkes Processes for Events in Social Media." *ArXiv*, **abs/1708.06401**, 2017.