Tourette Tic Time Series Analysis*

Torbjörn Wästerlid torbjorn.wasterlid@nottingham.ac.uk
Georgina Jackson georgina.jackson@nottingham.ac.uk
Elena Nixon elena.nixon@nottingham.ac.uk
Li Bai bai.li@nottingham.ac.uk

Abstract

This paper consists of an exploratory look at a tic time series formed by observing the start time of tics for subjects with Tourette's Syndrome during various activities. It is concluded that the mean waiting time between tics is increased by physical activity and that subsequent waiting times appear independent of each other within these activities. Attempts are made at modelling the waiting times using a Pareto, an Exponential and a Gamma distribution. Of these, the Gamma fits the data the best. This contradicts previous results by Peterson and Leckman suggesting that tics are distributed according to the Pareto distribution.

1 Introduction

Mathematics has been working its way into medicine and psychology for a long time, yet little work has been done to mathematically describe the ticing in Tourette's Syndrome as a stochastic process. This paper presents preliminary work into that effort. The research aims to answer the following questions:

- Are tic characteristics, e.g., tic waiting times, of diagnostic value for assessing clinical severity and monitoring changes during intervention/treatment?
- Are tic characteristics of prognostic value for predicting clinical course, e.g., whether tics will improve or symptoms are likely to worsen with age?
- Can tics be appropriately modelled by some stochastic process, for example, the Poisson process, and can any inferences about tic intensity, or other characteristics be drawn from this model?

*Title design by howtoTeX.com

1.1 Background

Tourette's Syndrome (TS) is a neuropsychiatric disorder characterised by motor and vocal tics in patients [8]. The time course of tics in TS is of particular interest for diagnosing tic severity and measuring the efficacy of treatment. Automated analysis of the time course is desirable but challenging because tics fluctuate in type, frequency, and intensity over time. They are also very sensitive to contextual factors.

Traditionally tic monitoring is reduced to the analysis of mean waiting times over some time intervals. It is hoped, however, that more can be gleaned from the tic time-series by constructing statistical models that accurately describe the series. Such a model might contain parameters which are insensitive to the natural waxing and waning of tics and thus constitute a more suitable measure of the efficacy of treatments or severity of the disorder.

1.2 Previous work

A lot of the recent work regarding Tourette's Syndrome concerns its comorbidity with other neural disorders, such as Attention - Deficit - Hyperactivity disorder (ADHD). Often the comorbid diseases cause more problems for the subject, than Tourette's Syndrome in itself [3]. Work has also been put into how to treat TS [2, 6] and fMRI studies have tried to pin-point areas of brain activity during tics [1].

Little work, however, has gone into analyzing the tics mathematically. Such a study was conducted by Bradley S. Peterson and James F. Leckman [7] in 1998. Their findings suggest "the presence of a fractal, deterministic and possibly chaotic process in the tic time series". They also show that the tic time series of their subjects appears to follow an inverse power law.

1.3 Data collection and processing

Data for this study was provided by the Mental Health Institute at the University of Nottingham. Subjects were filmed whilst engaged in various activities. These included: Tai Chi, a yoga breathing exercise, games, interoception tasks, suppression tasks, a clinical interview and cardio boxing. Thus, several hours of material exist for each subject. Videos were then processed manually, with the types of tic as well as start and end times of each tic noted down with 10ms precision. Due to the long time required for the manual processing, data from only a single subject has been used in this preliminary study. This data was observed during two separate sessions.

The subject studied has 10 different tics: eye blinking, eye movement, facial grimace, mouth movement, head/neck jerk, shoulder shrug/gyrating, hand movement, waist gyrating, throat clearing and grunting. Note that the last two tics are vocal. The main tics, that occur by far the most frequently, are the head-/neck movement and shoulder shrug/gyrating. This study therefore focuses on modelling these tics in particular.

2 Tools for Exploratory Analysis

Let X_t be a random variable with distribution function $F(x)$ and let T be some index set. A time series X is a collection of random variables $X = \{X_t : t \in T\}$. Thus X is the result of a random variable at each t.

2.1 QQ-plot

The QQ-plot is a tool that allows visual inspection and rejection of any hypothesis H one might have regarding the distribution of some random variable Y.

Definition 1 (Quantile set)**.** *The quantile set Q_N for a random variable Y with distribution $F(y)$ is a set of size N:*

$$Q_N = \{y : F(y) = \frac{n}{N+1}, n = 1 \dots N\}$$

The set Q_N thus divides $F(y)$ into $N+1$ parts of equal area. QQ-plots exploit the fact that if two random variables are equal, then the quantile sets of the two random variables are also equal. The quantiles of a random sample can be easily approximated. If a hypothesis H is false then a plot of the approximated versus theoretical quantiles will deviate a lot from the straight line $y = x$. If on the other hand, they lie approximately on the straight line, then H cannot be rejected.

2.2 Auto Correlation Function

The auto correlation function (ACF) is defined in terms of the auto covariance function. It reveals fixed time lag dependecies in a stationary time series [4]. Thus, it provides useful information about whether previous values of the time series can be used to model future values. The formal definition is:

Definition 2 (Auto Covariance Function)**.**

$$\gamma(h) = \mathbf{E}[(X_{t+h} - \mathbf{E}[X])(X_t - \mathbf{E}[X])],$$

where $h > 0$ and $\mathbf{E}$ is the expectation operator. The auto correlation function $\rho(h)$ is obtained from the auto covariance function by normalizing with the variance of the time series.

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} \qquad (1)$$

2.3 Maximum Likelihood Estimation

Let X be a random variable with density $f(x;\theta)$, where θ is some vector of unknown parameters describing the density. Assume that X is sampled N times in an experiment and denote this data $S = x_1, \dots, x_N$. A common question is: given S what values should be assigned to θ? Maximum likelihood estimation answers this question by assigning those values to θ that make the observed values S the most probable. Mathematically this boils down to the following optimization problem:

$$\max_\theta \mathcal{L}(\theta|S) = f(x_1, x_2, \dots; \theta), \qquad (2)$$

where $\mathcal{L}$ is called the likelihood function and $f(y_1, y_2, \dots; \theta)$ is the joint probability for X. Equation 2 may look daunting, but in practice, if observations are independent, the joint distribution becomes:

$$f(y_1, y_2, \ldots; \theta) = \prod_{k=1}^{N} f(y_k; \theta)$$

To simplify further the logarithm is often applied, so that the product reduces to a sum. For common, simple densities such as the exponential, the maximum likelihood leads to an analytical estimator for θ.

2.4 Pareto Distribution

The Pareto distribution is an instance of an inverse power law. It is used frequently to model phenomena with heavy tails. A random variable Y with density $f(y)$ is said to be distributed as a Pareto distribution if:

$$f(y) = \theta(y - y_0)\alpha \frac{y_0^\alpha}{y^{\alpha+1}}, \qquad (3)$$

where θ is the Heaviside function, $y_0 > 0$ is a cut off point to ensure a finite distribution and $\alpha > 0$. By taking the logarithm of both sides of equation (3) an expression is obtained that is linear in $\log(y)$, with slope $-(\alpha + 1)$. This is a simple way to check if a process is Pareto distributed.

2.5 Poisson Process

A Poisson process with intensity λ is a random process $N(t)$ that satisfies the following properties:

1. $N(t)$ is increasing and $N(t) \in \{0, 1, 2, \ldots\}$.
2. For $h > 0$ the probability measure $\mathbf{P}$ satisfies

$$\mathbf{P}(N(t + h) = n + m | N(t) = n) = \begin{cases} \lambda h + O(h) & \text{if } m = 1, \\ O(h) & \text{if } m > 1, \\ 1 - \lambda h + O(h) & \text{if } m = 0. \end{cases}$$

3. The number of counts $N(t) - N(s)$, with $s < t$ is independent of $N(s)$.

From the above definition follows that the Poisson process is distributed according to the Poisson distribution and that the waiting times, the time between each increment of $N(t)$, follows the Exponential distribution. The Poisson process is used to model count data, such as radioactive decays and server queues [5].

2.6 Gamma Distribution

The Gamma distribution $Gamma(\alpha, \beta)$ is a two-parameter distribution of which the Exponential distribution is a special case. We have,

$$\text{Gamma}(x; \alpha, \beta) = \beta^\alpha \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad (4)$$

where $\Gamma(\cdot)$ is the Gamma function and $alpha$, $\beta > 0$. For $\alpha = 1$ we obtain the Exponential distribution. As it has one more parameter, the Gamma distribution is more flexible than the Exponential distribution and may be better suited to model data where the variance is not equal to the mean. The Gamma distribution has been used to model for instance the load on web servers and meteorology [9].

3 Statistical analysis

The waiting times (WTs) studied in this section were formed by subtracting the start time t_i of the ith tic from that of the start time of the $i + 1$th tic, starting with $t_0 = 0$. Only tics of the main type were included in the study. Figure 1 shows the waiting times plotted sequentially as they occurred during the two sessions. As can be seen the suppression task causes significant outliers.

Figure 2 presents a histogram of the data of both sessions bunched together and divided into bins 1 second wide. The number of short WTs vastly outnumber the large WTs. The shape of the histogram looks roughly like that of an exponential or Pareto distribution.

Figure 1 suggests that activities might affect the rate of ticing. The simplest way to quantify that is via the mean waiting time for that activity. Table 1 summarizes this information. As can be seen, physical activities greatly increase the wating time.

3.1 ACF

In Figure 1 it appears that large TIs are often followed by another large TI and vice versa for small TIs. However, this is largely an effect of plotting all the different activities in the same plot. When looking at each activity separately, the ACF, plotted in Figure 3, shows

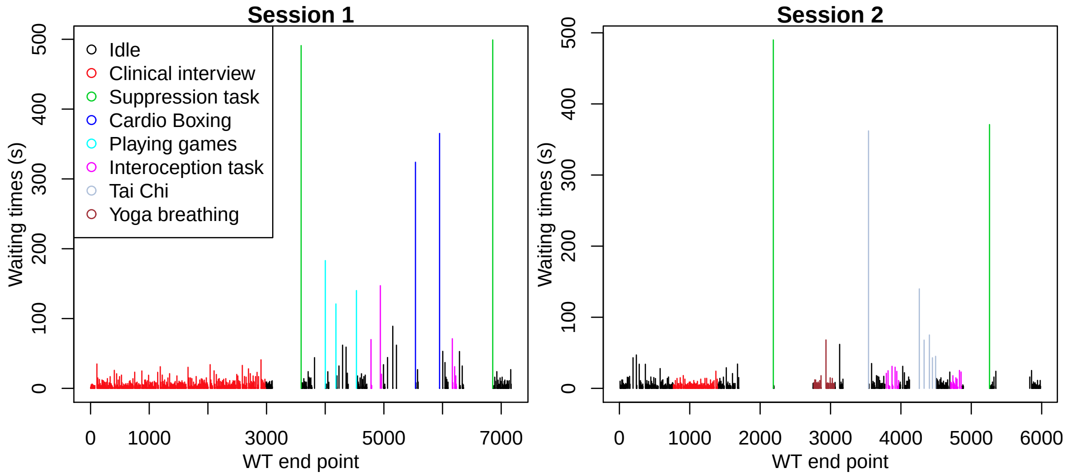

Figure 1: The plot shows waiting times on the y-axis and the time at which the tic at which the wait ended on the x-axis. Note that during certain activities the waiting times are very large.

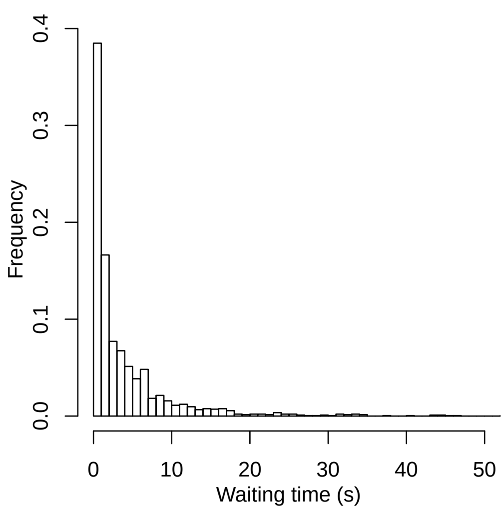

Figure 2: Histogram of data from both sessions clumped together. Data was split into bins one second wide. The X-axis has been cropped at $WT = 50$ for readability. There are bins with positive values up to about $WT = 500$s.

Table 1: Mean waiting time and standard deviation during different activities.

activity	mean	sd
total	6.16	26.86
idle	5.30	8.74
interview	3.22	4.71
suppression	462.76	61.32
cardio boxing	344.52	29.15
games	122.03	58.14
interoception	8.55	19.77
tai chi	46.53	33.80
yoga	6.91	10.54
S_1	4.02	6.64
S_2	8.00	17.2
S_3	155.53	128.11

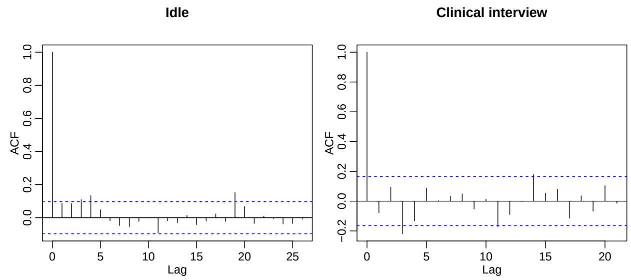

Figure 3: ACF for two different activities. The dotted line is a 95% confidence interval. As can be seen the correlations between waiting times are insignificant.

that waiting times are hardly correlated with previous waiting times at all.

The use of the ACF here is questionable, since it is unclear, and unlikely that the time series is stationary. We argue here, however, that the ACF can be applied to individual activities, since the time span is relatively short and because the time series studied is the differences of another time-series, reducing non-stationarity.

3.2 Inverse Power Law

Peterson and Leckman [7] report that their waiting times follow an inverse power law, that is, they should be distributed according to the Pareto distribution (since WT>0 and is continuous). Since a histogram is an estimate of the density function, one way to visually assess whether data follows an inverse power law or not is to plot the histogram frequencies and bin-values on a loglog scale. The slope of that plot is then equal to the parameter $\alpha - 1$ of the Pareto distribution (see Section 2.4). Figure 4 shows both a log-log plot, suggesting that the data does follow an inverse power law, and a QQ-plot that rejects the idea.

3.3 Poisson Process

A Poisson process has exponentially distributed waiting times. Hence, for the tic process to be adequately modelled as a Poisson process the waiting times should be exponentially distributed, with a parameter λ corresponding to the tic rate. From Table 1 it is apparent that the tic rate is dependant on activity. To take this

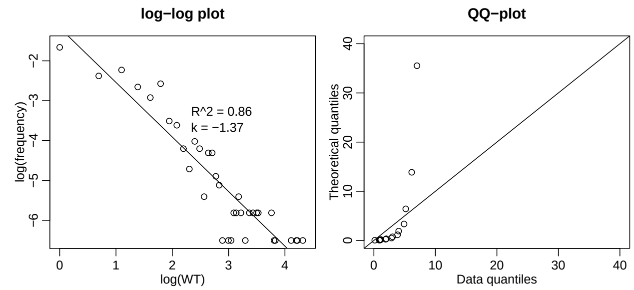

Figure 4: On the left is a log-log plot of histogram bin values and frequencies for all the data collected during Session 2, except for the suppression task. The straight line was fit using least squares. On the right is a QQ-plot of the same data set versus a Pareto distribution with parameters $y_0 = 0.04$ and $\alpha = 0.237$, obtained from maximum likelihood estimation. Note that the plot deviates strongly from the line $y = x$.

into account data was partitioned into three sets:

$$
\begin{aligned}
S_1 &= \{WT : WT \in \{\text{interview, idle}\}\} \\
S_2 &= \{WT : WT \in \{\text{yoga, interoception}\}\} \\
S_3 &= \{WT : WT \in \{\text{tai-chi, cardio boxing and gaming}\}\}.
\end{aligned}
$$

Set three, consisting of the gaming, cardio boxing and tai-chi activities are all considered to be physically demanding tasks and are therefore grouped. However, even so, this group contains only 12 data points meaning results are very uncertain. The sample mean ($\hat{\mu}$) and standard deviation of these sets are available in Table 1. Of further note is that four outliers (out of a total of 128 points) were removed from S_2. S_1 consists of 1825 data points.

Figure 5 shows QQ-plots of the three partitions versus exponential distributions, with $\lambda = 1/\hat{\mu}$. These plots suggest that data is not exponentially distributed.

3.4 Gamma Process

QQ-plots for S_1, S_2 and S_3 against a Gamma distribution can be studied in Figure 6. Parameters for the gamma distribution were estimated from data using the method of maximum likelihood.

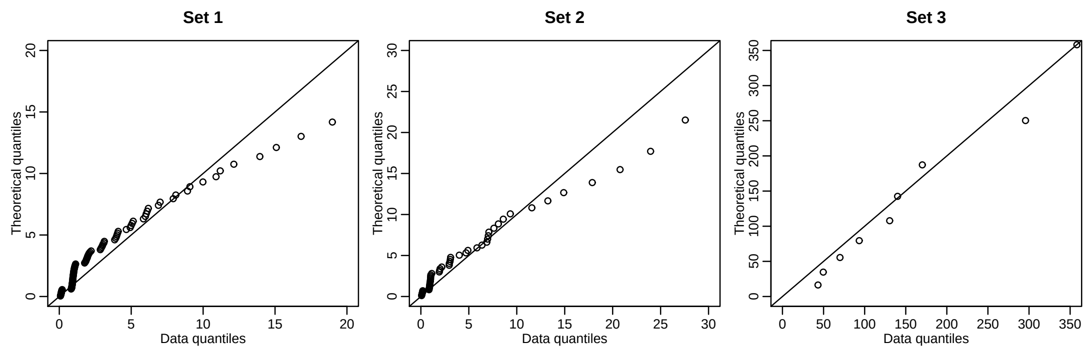

Figure 5: QQ-plots for each of the sets S_1, S_2 and S_3 versus Exponential distributions with rates $\lambda_1 = 1/\mu_1$, $\lambda_2 = 1/\mu_2$ and $\lambda_3 = 1/\mu_3$, respectively. Note that, especially for S_1 the exponential distribution fails to capture the overall shape of the data. The data seems to contain too many small waiting times.

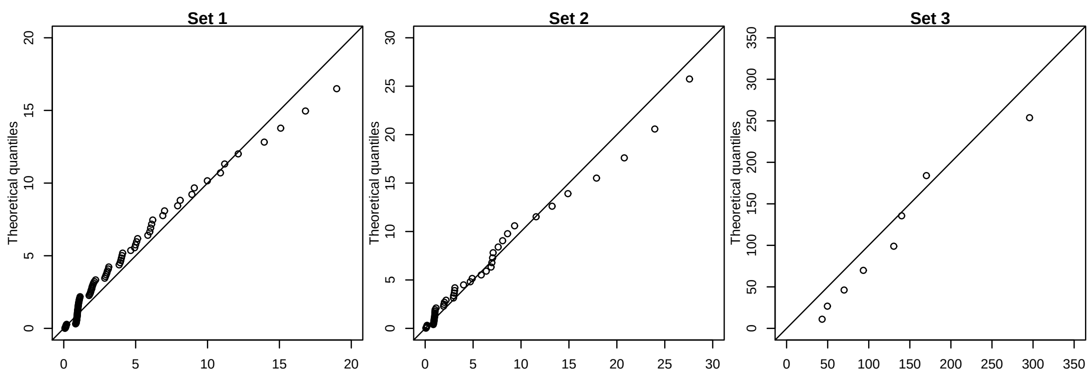

Figure 6: QQ-plots for each of the sets S_1, S_2 and S_3 versus Gamma distributions with parameters ($\alpha = 0.698$, $\beta = 0.173$), ($\alpha = 0.678$, $\beta = 0.173$) and ($\alpha = 0.824$, $\beta = 0.005$).

4 Discussion and Conclusions

It is clear that the tic rate in patients with Tourettes Syndrome is greatly affected by physical activity, both from Figure 1 and Table 1. Furthermore, auto correlation plots found in Figure 3 indicate that within the different activities, individual TIs are not correlated.

Even with this knowledge it has proven to be difficult to find a distribution that models TIs well. The Pareto and Exponential distributions have both been rejected. Figure 4 shows that the data is not sufficiently heavy tailed to warrant a Pareto distribution, and the Exponential distribution is too inflexible, failing to take into account the overdispersion, as evident in Figure 5.

The Gamma distribution initially looks more promising than the exponential, altough the QQ-plot for S_1, seen in Figure 6 suggests that the distribution still fails to fully account for the data. The tail of the Gamma appears not to be heavy enough and the peak not as sharp as needed.

Future work will try fitting the Weibull distribution, commonly used in reliability analysis. An alternative modelling approach, using a Markov chain and focusing on the actual tics, rather than waiting times is also under consideration.

References

[1] S Bohlhlater, A GoldFine, S Matteson, G Garraux, T Hanakawa, K Kansaku, R Wurzman, and M Hallet. Neural correlates of tic generation in Tourette syndrome: an event-related functional mri study. *Brain*, 129:2029–2037, 2006.

[2] Clare M Eddy, Hugh E Rickards, and Andrea Cavanna. Treatment strategies for tics in Tourette syndrome. *Therapeutic Advances in Neurological Disorders*, 4(1):25–45, 2011.

[3] RogerD. Freeman. Tic disorders and ADHD: answers from a world-wide clinical dataset on Tourette syndrome. *European Child & Adolescent Psychiatry*, 16(9):15–23, 2007.

[4] Wayne A. Fuller. *Introduction to Statistical Time Series*. John Wiley & Sons, 1996.

[5] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. Oxford University Press, 2001.

[6] Elana Harris and Steve W Wu. Children with tic disorders: How to match treatment with symptoms. *Current Psychiatry*, 9:29–36, 2010.

[7] Bradley S. Peterson and James F. Leckman. The temporal dynamics of tics in gilles de la Tourette syndrome. *Biological Psychiatry*, 44(12):1337 – 1348, 1998.

[8] Harvey S Singer. Tourette's syndrome: from behaviour to biology. *The Lancet Neurology*, 4(3):149 – 159, 2005.

[9] H. C. S. Thom. A note on the gamma distribution. *Monthly Weather Review*, 86(4):117 – 122, 1958.