A Power-Law Decay Model with Autocorrelation for Posting Data to Social Networking Services

Toshifumi FUJIYAMA, Chihiro MATSUI and Akimichi TAKEMURA

METR 2014–31 November 2014

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.keisu.t.u-tokyo.ac.jp/research/techrep/index.html
The METR technical reports are published as a means to ensure timely dissemination of scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the authors or by other copyright holders, notwithstanding that they have offered their works here electronically. It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may not be reposted without the explicit permission of the copyright holder.
A power-law decay model with autocorrelation for posting data to social networking services

Toshifumi Fujiyama,* Chihiro Matsui* and Akimichi Takemura*

November, 2014

Abstract

We propose a power-law decay model with autocorrelation for posting data to social networking services concerning particular events such as national holidays or major sport events. In these kinds of events we observe people’s interest both before and after the events. In our model the number of postings has a Poisson distribution whose expected value decays as a power law. Our model also incorporates autocorrelations by autoregressive specification of the expected value. We show that our proposed model well fits the data from social networking services.

Keywords and phrases: conditional Poisson autoregressive model, Fisher information, number of postings.

1 Introduction

With the increasing usage of social networking services (SNS), such as many blog services, Facebook or Twitter, it is important to obtain information from postings to these services. We look at the data on the number of postings for various events, such as national holidays or major sport events. By analyzing the data, we gain insights on how a particular topic interests people, how actively it is discussed around the date of the event and how long the memory of the event lasts. Our data is time series data, but the length of the series is usually short, covering several weeks before and after the event. It is far from stationary, because there is a sharp peak of the number of postings on the date of the event.

Many studies on posting data to SNS focus on the network structure of the services (e.g. [2], [9], [12]), such as how certain topic spreads by interaction among users of the services. In this paper we study the pattern of number of postings concerning particular events, such

*Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo

1We are grateful to NTTCom Online Marketing Solutions Corporation for allowing us to use their BuzzFinder service, which provided excellent access to social networking service posting data.
as national holidays or major sport events. The events we study are scheduled events, so that people are aware when the events will happen. This contrasts with unpredictable events, such as the occurrence of large earthquakes ([11]). In the case of unpredictable events, we only observe effects of the event after it happens. On the other hand, in the case of the scheduled events, the number of postings depends on people’s anticipation of the event and the after-effect of the actual outcome of the event. Another similar type of data studied in literature is the registration data until the deadline for events such as academic conferences (e.g. [1], [4]). In this kind of data, we only observe people’s actions before the deadline.

We combine the model of the mean number of postings proposed in [10] and the conditional Poisson autoregressive (AR) models for count data. The conditional Poisson autoregressive models are applied in many problems in econometrics, political science or epidemiology (e.g. [8], [3]). For a theoretical survey of the conditional Poisson autoregressive model see [6], [5] and [7]. Zhu ([13]) generalized the conditional Poisson model to generalized Poisson integer-valued GARCH models. By our model we can better predict how people’s interest on particular events decays with time and our model is useful for example in designing advertisement strategy for web marketing.

In the left graph of Figure 1 we show a typical symmetric pattern of number of postings. Tokyo Marathon 2014 was held on Sunday, February 23 of 2014. It was well anticipated and it ended without unexpected happenings. In these cases, the number of postings shows a symmetric pattern with a sharp peak on the date of the event. Both sides of the peak seems to exhibit a behavior of some negative power in the time difference from the actual date of the event. Also there is some stochastic fluctuation of the number of postings. This is the motivation of our proposed model. A different asymmetric pattern is observed, when there is some surprising element of the event. In the right graph of Figure 1 we show the data on Noriaki Kasai around February 16, 2014, when he won a silver medal in ski jump in 2014 winter Olympics. From the data we see that people did not anticipate the medal before the event.

![Graph 1: Symmetric and asymmetric patterns](image-url)
The organization of this paper is as follows. In Section 2 we propose to model the mean of the number of postings by a power-law decay model with three parameters and we assume independent Poisson distributions. We also study the Fisher information matrix under the proposed model. In Section 3 we incorporate autocorrelation to our model by conditional Poisson autoregression. In Section 4 we apply our model to some social networking service data in Japan. We end the paper with some discussion in Section 5.

2 A power-law decay model for the mean number of postings

Let \( t_0 \) denote the date of the event and let \( y_t \) denote the number of postings on the event on day \( t \). We model the expected value of \( y_t \) by the following power-law decay function.

\[
E(y_t) = \mu_t(\alpha, \beta, \gamma) = \frac{1}{(\alpha | t - t_0| + 1)^\beta}.
\]

This power-law decay model was proposed by [10] without the parameter \( \beta \). They used the least-squares method, while we use the maximum likelihood estimation. They do not consider fitting the data close to the peak, which becomes possible by introducing the parameter \( \alpha \).

The interpretation of the parameters is as follows.

\( \alpha \): steepness of the curve just before and after the event
\( \beta \): longer decay pattern
\( \gamma \): impact of the event (peak level, the maximum number of postings)

In this section we assume that \( y_t, t = t_L, t_L + 1, \ldots, t_U - 1, t_U, t_L \leq t_0 \leq t_U \), are independent Poisson random variables with the mean given by \( \mu_t(\alpha, \beta, \gamma) \) in (1). We call the model “power-law decay independence model”. We denote the probability function of the Poisson distribution with the mean \( \mu \) as

\[
\text{Po}(y \mid \mu) = \frac{\mu^y}{y!} e^{-\mu}.
\]

Then the likelihood function is written as

\[
L(\alpha, \beta, \gamma) = \prod_{t=L}^{t_U} \text{Po}(y_t \mid \mu_t(\alpha, \beta, \gamma)) = \prod_{t=L}^{t_U} \frac{\mu_t(\alpha, \beta, \gamma)^{y_t}}{y_t!} e^{-\mu_t(\alpha, \beta, \gamma)}.
\]

We found that the maximization of the log-likelihood function is numerically very simple. We give a typical example of maximum likelihood estimation (MLE) of (3) in Figure 2. In Figure 2 the solid line is the observed data and the dotted line is the fitted curve of expected values by MLE. We use the same distinction of line types in Figures 3 and 4. The estimates are \( \hat{\alpha} = 1.115, \hat{\beta} = 1.534 \) and \( \hat{\gamma} = 8820.593 \). Our model seems to fit the data well, but there is a slight asymmetry in this data, which is not captured by the symmetric model in (1).

We will discuss many examples of fitting of (3) and its generalizations later in Section 4.
2.1 Fisher information matrix for the independence model

As we already noted in the beginning, the length of our data is not very large and our data is far from stationary. Hence the usual asymptotics for the length of the series is not appropriate. Nevertheless it is of theoretical interest to consider the behavior of MLE of (3) as the length of the series diverges to infinity. On the other hand we have large number of postings $y_{t_0}$ on the date $t_0$ of the event. Under our model of Poisson distribution, we can also think of the asymptotics, where $\gamma = E(y_{t_0})$ diverges to infinity. In the following we calculate the Fisher information matrix

$$I = \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\gamma} \\ I_{\alpha\beta} & I_{\beta\beta} & I_{\beta\gamma} \\ I_{\alpha\gamma} & I_{\beta\gamma} & I_{\gamma\gamma} \end{pmatrix},$$

for our model to gain insights on the behavior of MLE for the model (3).

For notational simplicity let $t_0 = 0$ and $t_L \leq t \leq t_U$. Then the log-likelihood function

$$l(\alpha, \beta, \gamma) = \log L(\alpha, \beta, \gamma)$$

is written as

$$l(\alpha, \beta, \gamma) = \sum_{t=t_L}^{t_U} (y_t \log \mu_t(\alpha, \beta, \gamma) - \mu_t(\alpha, \beta, \gamma) - \log y_t)$$

$$= \sum_{t=t_L}^{t_U} \left( y_t (\log \gamma - \beta \log(\alpha|t| + 1)) - \gamma \frac{1}{(\alpha|t| + 1)^\beta} - \log y_t \right)$$

$$= C + \log \gamma \sum_{t=t_L}^{t_U} y_t - \beta \sum_{t=t_L}^{t_U} y_t \log(\alpha|t| + 1) - \gamma \sum_{t=t_L}^{t_U} \frac{1}{(\alpha|t| + 1)^\beta},$$

Figure 2: Fitting the power-law decay independence model to Tokyo Marathon data
where \( C \) does not depend on parameters. Then

\[
- \frac{\partial^2}{\partial \alpha^2} l = - \beta \sum_{t=1}^{I_u} \frac{y_t}{(\alpha|t| + 1)^{\beta + 1}} + \gamma \beta (\beta + 1) \sum_{t=1}^{I_u} \frac{|t|^2}{(\alpha|t| + 1)^{\beta + 2}}.
\]

The expected value of this second derivative is given as

\[
I_{aa} = E[- \frac{\partial^2}{\partial \alpha^2} l] = - \beta \gamma \sum_{t=1}^{I_u} \frac{y_t}{(\alpha|t| + 1)^{\beta + 1}} + \gamma \beta (\beta + 1) \sum_{t=1}^{I_u} \frac{|t|^2}{(\alpha|t| + 1)^{\beta + 2}}
\]

\[
= \gamma \beta^2 \sum_{t=1}^{I_u} \frac{|t|^2}{(\alpha|t| + 1)^{\beta + 2}}.
\]

Note that \( I_{aa} \to \infty \) as \( \gamma = E(y_0) \to \infty \). However, when \( \gamma \) is fixed

\[
\lim_{\max(t_L, t_U) \to \infty} I_{aa} = \infty \iff \beta \leq 1 \iff \sum_{t=1}^{I_u} E(y_t) \to \infty.
\]

Next consider \( I_{a\beta} \). Noting

\[
(at + 1)^{-\beta} = \exp(-\beta \log(at + 1))
\]

and

\[
\frac{\partial}{\partial \beta} (at + 1)^{-\beta} = - \log(at + 1) \exp(-\beta \log(at + 1)) = - \log(at + 1) \frac{1}{(at + 1)^\beta},
\]

we have

\[
- \frac{\partial^2}{\partial \alpha \partial \beta} l = \frac{\partial}{\partial \alpha} \left[ \sum_{t=1}^{I_u} \frac{y_t \log(a|t| + 1)}{\alpha|t| + 1} - \gamma \sum_{t=1}^{I_u} \frac{\log(a|t| + 1)}{\alpha|t| + 1} \frac{1}{(a|t| + 1)^\beta} \right]
\]

\[
= \sum_{t=1}^{I_u} \frac{y_t |t|}{\alpha|t| + 1} - \gamma \sum_{t=1}^{I_u} \frac{|t|}{(\alpha|t| + 1)^{\beta + 1}} + \gamma \beta \sum_{t=1}^{I_u} \frac{\log(a|t| + 1)}{\alpha|t| + 1} \frac{|t|}{(a|t| + 1)^{\beta + 1}}.
\]

When we take the expected value, the first two terms cancel and

\[
I_{a\beta} = E[- \frac{\partial^2}{\partial \alpha \partial \beta} l] = \beta \gamma \sum_{t=1}^{I_u} \frac{\log(a|t| + 1)}{\alpha|t| + 1} \frac{|t|}{(a|t| + 1)^{\beta + 1}}.
\]

The divergence is the same as in the case of \( I_{aa} \).

Similarly we can evaluate \( I_{a\gamma} \), \( I_{a\gamma} \), \( I_{\beta \gamma} \), \( I_{\gamma \gamma} \) as

\[
I_{a\beta} = \gamma \sum_{t=1}^{I_u} \frac{(\log(a|t| + 1))^2}{(a|t| + 1)^\beta}, \quad I_{a\gamma} = - \beta \sum_{t=1}^{I_u} \frac{|t|}{(a|t| + 1)^{\beta + 1}},
\]

\[
I_{\beta \gamma} = - \gamma \sum_{t=1}^{I_u} \frac{\log(a|t| + 1)}{(a|t| + 1)^\beta}, \quad I_{\gamma \gamma} = - \frac{1}{\gamma} \sum_{t=1}^{I_u} \frac{1}{(a|t| + 1)^\beta}.
\]
Although it is tedious to prove the consistency and the asymptotic normality of MLE based only on the computation of Fisher information matrix, our computation suggests the following results. Since \( y_{t_0} \) has the Poisson distribution with the mean \( y \) and the standard deviation \( \sqrt{y} \), \( y_{t_0}/y \) converges to 1 in probability as \( y \to \infty \). In fact, when we compute MLE, \( y \) is basically estimated by \( y_{t_0} \), since the number of postings has a sharp peak at \( t = t_0 \) (see Figure 2). Furthermore \( I_{\alpha \alpha}, I_{\alpha \beta}, I_{\beta \beta} \) are linear in \( y \). This suggests that MLE is consistent as \( y \to \infty \). For large expected value, Poisson distribution is approximately by normal distribution after normalization. Hence the score functions \( \partial l/\partial \alpha, \partial l/\partial \beta, \partial l/\partial \gamma \) are approximately normally distributed as \( y \to \infty \). The confidence intervals given in Section 4 are based on this approximation. When \( y \) is fixed and \( \max(-t_L, t_U) \to \infty \), the elements of the Fisher information matrix diverge to \( 1 \) if and only if \( \beta \leq 1 \), or equivalently \( \sum_{t=t_0}^{t_U} E(y_t) \to \infty \).

3 Conditional Poisson regression modeling for autocorrelations

In the last section we assumed that the number of postings \( y_t \) are independent. We generalize this model to allow autocorrelations by conditional Poisson regression modeling. As we saw, the estimate of the parameter \( \gamma \) in (1) is very close to \( y_{t_0} \). Hence in this section we replace \( y \) by \( y_{t_0} \). This is the initial value of our autoregressive scheme and we model the number of postings after the event \( y_{t_0}, t > t_0 \).

Concerning the data \( y_t, t < t_0 \), before the date of the event, we can use the model given in (4) by reversing the time axis. This is similar to look at the standard AR(1) process

\[
x_t = \rho x_{t-1} + \epsilon_t
\]

in the reverse time direction by taking the reciprocal of the autoregressive coefficient \( \rho \). However this modeling of the data before the date of the event is somewhat unsatisfactory, in particular for the purpose of predicting \( y_{t_0} \) before the date of the event. We discuss this point again in Section 5.

We replace \( y \) by \( y_{t_0} \) in (1) and regard it as the conditional expected value of \( y_t \) given \( y_{t_0} \)

\[
E(y_t|y_{t_0}) = \frac{1}{(\alpha|t-t_0| + 1)^\beta}, \quad t \geq t_0.
\]

Then \( E(y_t|y_{t_0}) \) is recursively written as

\[
E(y_t|y_{t_0}) = E(y_{t-1}|y_{t_0}) \left( \frac{(|t-t_0| - 1)\alpha + 1}{|t-t_0|\alpha + 1} \right)^\beta
\]

\[
= E(y_{t-2}|y_{t_0}) \left( \frac{(|t-t_0| - 2)\alpha + 1}{|t-t_0|\alpha + 1} \right)^\beta
\]

\[
= \ldots
\]
We propose the following AR(2) type modeling of $y_t, t \geq t_0 + 2$:

$$
p(y_{t_0 + 1} | y_{t_0}) = \text{Po} \left( y_{t_0 + 1} \mid y_{t_0} \frac{1}{(\alpha + 1)\beta} \right)
$$

$$
p(y_{t_0 + 2} | y_{t_0 + 1}, y_{t_0}) = \text{Po} \left( y_{t_0 + 2} \mid s \times y_{t_0 + 1} \left( \frac{\alpha + 1}{2\alpha + 1} \right)^\beta + (1-s) \times y_{t_0} \left( \frac{1}{2\alpha + 1} \right)^\beta \right)
$$

$$
p(y_{t_0 + 3} | y_{t_0 + 2}, y_{t_0 + 1}) = \text{Po} \left( y_{t_0 + 3} \mid s \times y_{t_0 + 2} \left( \frac{2\alpha + 1}{3\alpha + 1} \right)^\beta + (1-s) \times y_{t_0 + 1} \left( \frac{\alpha + 1}{3\alpha + 1} \right)^\beta \right)
$$

$$
\vdots
$$

$$
p(y_{t_0 + m} | y_{t_0 + m-1}, y_{t_0 + m-2}) = \text{Po} \left( y_{t_0 + m} \mid s \times y_{t_0 + m-1} \left( \frac{(m-1)\alpha + 1}{ma + 1} \right)^\beta + (1-s) \times y_{t_0 + m-2} \left( \frac{(m-2)\alpha + 1}{ma + 1} \right)^\beta \right).
$$

(4)

Note that $y_{t_0 + 1}$ is given in an AR(1) form. Then the conditional likelihood function for $\alpha, \beta, s$ given $y_{t_0}$ for the data $y_{t_0 + 1}, \ldots, y_{t_0 + T}$ is written as

$$
L(\alpha, \beta, s) = \text{Po} \left( y_{t_0 + 1} \mid y_{t_0} \frac{1}{(\alpha + 1)\beta} \right)
$$

$$
\times \prod_{m=2}^T \text{Po} \left( y_{t_0 + m} \mid s \times y_{t_0 + m-1} \left( \frac{(m-1)\alpha + 1}{ma + 1} \right)^\beta + (1-s) \times y_{t_0 + m-2} \left( \frac{(m-2)\alpha + 1}{ma + 1} \right)^\beta \right).
$$

(5)

When $s = 1$, we have an AR(1) form. When we estimate $s$ in $L(\alpha, \beta, s)$, we restrict $s \in [0, 1]$, although for some data sets unrestricted MLE of $s$ happened to be larger than 1.

Note that the independence model in (3) and the AR(2) model in (5) are separate models. In the usual AR(1) model of continuous observations $x_t = \rho x_{t-1} + \epsilon_t$, the independence model is a special case of $\rho = 0$. In order to interpolate between the independence model (3) and the AR(2) model (5), we propose the following more generalized and unifying model with the new parameters $u, v \in [0, 1]$ representing the weights of the two models.

$$
p(y_{t_0 + m} | y_{t_0 + m-1}, y_{t_0 + m-2})
$$

$$
= \text{Po} \left( y_{t_0 + m} \mid w \times \frac{y_{t_0}}{(m-1)\alpha + 1} \right)^u \times \left( (m-1)\alpha + 1 \right)^{1-u} \times \left( \frac{y_{t_0 + m-1}}{ma + 1} \right)^\beta
$$

$$
+ (1-w) \times \left( \frac{y_{t_0}}{(m-2)\alpha + 1} \right)^v \times \left( (m-2)\alpha + 1 \right)^{1-v} \times \left( \frac{y_{t_0 + m-2}}{ma + 1} \right)^\beta.
$$

(6)

In this unifying model, $u$ is the weight for the lag one term and $v$ is the weight for the lag two term. We introduced these two parameters separately for flexibility of the model.
3.1 Fisher information matrix for the AR(1) model

Here we evaluate Fisher information matrix for AR(1) model, i.e. the model in (5) with \( s = 1 \). We let \( t_0 = 0 \) for simplicity and assume that \( y_0, \ldots, y_T \) are observed. We also replace \( \gamma \) by \( y_0 \) and consider the conditional likelihood in \( \alpha \) and \( \beta \) given \( y_0 \).

The conditional expected values given \( y_0 \) are evaluated as

\[
E[y_1|y_0] = y_0 \left( \frac{1}{\alpha + 1} \right)^\beta,
\]

\[
E[y_2|y_1] = y_1 \left( \frac{\alpha + 1}{2\alpha + 1} \right)^\beta,
\]

\[
E[y_3|y_2] = y_2 \left( \frac{\alpha + 1}{2\alpha + 1} \right)^\beta
\]

\vdots

\[
E[y_t|y_{t-1}] = y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta
\]

The conditional likelihood function is

\[
L(\alpha, \beta) = \prod_{t=1}^T \frac{\mu_t^y e^{-y_t^2}}{y_t!}, \quad \mu_t = y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta.
\]

Then the conditional log-likelihood function \( l(\alpha, \beta) = \log L(\alpha, \beta) \) is written as

\[
l(\alpha, \beta) = \sum_{t=1}^T (y_t \log \mu_t - \mu_t - \log y_t!)
\]

\[
= \sum_{t=1}^T \left\{ y_t \log y_{t-1} + \beta \log \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right) - y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta - \log y_t! \right\}
\]

\[
= C + \beta \sum_{t=1}^T y_t \log \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right) - \sum_{t=1}^T y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta,
\]

where \( C \) does not depend on \( \alpha, \beta \). The first derivative and the second derivative with respect to \( \alpha \) are evaluated as

\[
-\frac{\partial}{\partial \alpha} l(\alpha, \beta) = -\beta \sum_{t=1}^T y_t \left\{ \frac{t-1}{(t-1)\alpha + 1} - \frac{t}{t\alpha + 1} \right\} - \beta \sum_{t=1}^T y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta \frac{(t-1)^2}{(t\alpha + 1)^{\beta+1}},
\]

\[
-\frac{\partial^2}{\partial \alpha^2} l(\alpha, \beta) = -\beta \sum_{t=1}^T y_t \left\{ -\frac{(t-1)^2}{((t-1)\alpha + 1)^2} + \frac{t^2}{(t\alpha + 1)^2} \right\}
\]

\[
- \beta \sum_{t=1}^T y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta \frac{(t-1)^2}{((t-1)\alpha + 1)^2} - 2r^2 - 2(1-\alpha)t + 1 - \beta.
\]

Taking the expected value we have

\[
I_{\alpha \alpha} = -\beta y_0 \sum_{t=1}^T \frac{1}{(t\alpha + 1)^{\beta}} \left\{ -\frac{(t-1)^2}{((t-1)\alpha + 1)^2} + \frac{r^2}{(t\alpha + 1)^2} \right\}
\]
\[-\beta y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1 - 2}{(t)\alpha + 1 - 2} \left( -2\alpha t^2 - 2(1 - \alpha)t + 1 - \beta \right) \]

\[= \beta y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1 - 2}{(t)\alpha + 1 - 2} \left( -2\alpha t^2 - 2(1 - \alpha)t + 1 \right) \]

\[+ \beta y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1 - 2}{(t)\alpha + 1 - 2} \left( 2\alpha t^2 + 2(1 - \alpha)t - (1 - \beta) \right) \]

\[= \beta^2 y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1 - 2}{(t)\alpha + 1 - 2} . \]

The mixed derivative with respect to \(\alpha\) and \(\beta\) and its expected value are evaluated as

\[-\frac{\partial^2}{\partial \beta \partial \alpha} l(\alpha, \beta) = -\sum_{t=1}^{T} y_t \left\{ \frac{t-1}{(t-1)\alpha + 1} - \frac{t}{t\alpha + 1} \right\} \]

\[= -\sum_{t=1}^{T} y_{t-1} \frac{(t-1)\alpha + 1}{(t\alpha + 1)^{\beta+1}} \left\{ 1 + \beta \log \frac{(t-1)\alpha + 1}{t\alpha + 1} \right\} , \]

\[I_{\alpha\beta} = -y_0 \sum_{t=1}^{T} \frac{1}{(t\alpha + 1)^\beta} \left\{ \frac{t-1}{(t-1)\alpha + 1} - \frac{t}{t\alpha + 1} \right\} \]

\[= -y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1}{(t\alpha + 1)^{\beta+1}} \left\{ 1 + \beta \log \frac{(t-1)\alpha + 1}{t\alpha + 1} \right\} \]

\[= y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1}{(t\alpha + 1)^{\beta+1}} \]

\[= -y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1}{(t\alpha + 1)^{\beta+1}} \left\{ 1 + \beta \log \frac{(t-1)\alpha + 1}{t\alpha + 1} \right\} \]

\[= -\beta y_0 \sum_{t=1}^{T} \frac{(t-1)\alpha + 1}{(t\alpha + 1)^{\beta+1}} \log \frac{(t-1)\alpha + 1}{t\alpha + 1} . \]

Similarly, the second derivative with respect to \(\beta\) and its expected values are evaluated as

\[-\frac{\partial^2}{\partial \beta^2} l(\alpha, \beta) = \sum_{t=1}^{T} y_{t-1} \left( \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^\beta \left( \log \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^2 . \]

\[I_{\beta\beta} = y_0 \sum_{t=1}^{T} \frac{1}{(t\alpha + 1)^\beta} \left( \log \frac{(t-1)\alpha + 1}{t\alpha + 1} \right)^2 . \]

Note that the elements of the Fisher information matrix are proportional to \(y_0\). Also the relevant series diverges if and only if \(\beta \leq 1\). This is the same as in the power-law decay independence model.

For more complicated models of this section, the evaluation of Fisher information matrix is difficult, mainly because we can not separate \(y_{t-1}\) and \(y_{t-2}\) in \(\log \mu_t\).
4 Data analysis of some Japanese social networking data

We apply our models to some social networking service data in Japan. The data which we used are summarized in Table 1. In Table 1, “Date” is the date of the event in the format month/data in 2014. “ID” is our identifier of the events used in later tables. “Searchword” is the word we used in BuzzFinder service to search for the postings related to the events. “Remarks” are the explanations of the events.

4.1 Parameter estimation of the power-law decay independence model

In Table 2 we show fits of the power-law decay independence model to data. Because many events showed asymmetry before and after the date of the event, we estimated the before-event parameters \( \alpha_b, \beta_b \) and the peak level \( \gamma \) for one week before the event, and then estimated the after-event parameters \( \alpha_a, \beta_a \) separately with the same \( \gamma \) as the before-event parameter. We also computed 95% confidence intervals.

In the graph of Figure 3 we show the data of Valentine’s day around February 14, 2014. The graph looks almost symmetric at first sight, but the estimated before-event parameters and after-event parameters were different. Indeed in the graph of Figure 3 the slope just before the date of the event is steeper than after the event and the number of postings decrease to zero faster after the event than before the event. Our estimated parameters reflect these facts.

\[
\begin{array}{|c|c|c|}
\hline
\text{Log-likelihood} & \text{95\% confidence interval} \\
\hline
\alpha_b = 3.498 & 3.267 < \alpha_b < 3.728 \\
\beta_b = 0.951 & 0.929 < \beta_b < 0.973 \\
\gamma = 55538.8 & 54728.4 < \gamma < 56349.3 \\
\alpha_a = 0.742 & 0.707 < \alpha_a < 0.777 \\
\beta_a = 1.991 & 1.935 < \beta_a < 2.046 \\
\hline
\end{array}
\]

Figure 3: Valentine’s Day 2014 for power-law decay independence model

Based on the power-law decay independence model we considered predicting the after-event parameters based on the data before the event. However this was difficult, because of the asymmetry found in many events. To confirm this phenomenon we performed multiple regression analysis, where the before-event parameters \( \alpha_b, \beta_b, \gamma \) are explanatory variables and the after-event parameters \( \alpha_a, \beta_a \) are objective variables. But we did not find significant correlation.
4.2 Parameter estimation of the AR(2) model

In Table 3 we show the fit of AR(2) model to our data. In Table 3 “log-lik.” stands for the log-likelihood of the estimated model. Figure 4 shows the fit of the AR(2) model for “Kodomo no Hi” and for “Yuko Oshima” as typical examples. In Figure 4 $s$ is estimated as $s = 1$ for “Kodomo no Hi”, whereas $s$ is estimated as $s = 0.576$ for “Yuko Oshima”. It seems that the parameter $s$ reflects the property of the event. The parameter $s$ tends to be close to 1 for events with faster decay patterns, but tends to be less than one for events with long-lasting interest after the events. This is reasonable, because $1 - s$ represents the effect of two days ago and $s = 1$ means that the autocorrelation is fully explained only by the number of postings one day ago. We compared AIC (Akaike’s Information Criterion) for AR(2) model and AIC for AR(2) model with $s = 1$. For many data sets AIC was smaller when $s$ is estimated to be less than 1.

4.3 Parameter estimation of the unifying model

In Table 4 we apply the unifying model (6) to data. In the Table 5 we compare the unifying model and other models based on AIC. In many cases the values of the parameters $u, v$ are close to 1 in this model. This suggests that the unifying model is over-parameterized for many events and the maximum likelihood estimation is not very stable. Indeed when we compare AIC for various models, often other models have smaller AIC than the unifying model.

5 Summary and discussion

In this paper we proposed a Poisson autoregression model with the power-law decay of the mean parameter for the number of postings data. Our model shows a good fit to various Japanese social networking data. Also the parameters of our model are easy to interpret and our model is useful in describing patterns of the events. Since the length of the data considered in this paper
is fairly short, covering only about one month, the unifying model in (6) with five parameters is probably flexible enough. In fact for many events, we found that smaller models than the unifying model showed better fits.

Our model assumes that there is a single date $t_0$ of an event. Some events such as the Olympic games have longer duration. The pattern of number of postings during the event with longer duration seems to be more complicated, although the patterns before the beginning of the event and the after the event seem to be similar to single-day events. It is not clear how to generalize our model to events with a longer duration.

From practical viewpoint, it is important to predict the peak level $\gamma$ of the number of postings and the after-event parameters $\alpha_a, \beta_a$ before the event. However we found that this prediction was difficult for our data. Therefore in data analysis in Section 4, we separately estimated the before-event parameters and the after-event parameters, although for the modeling purpose this is somewhat unsatisfactory. As discussed at the beginning of Section 3 our conditional Poisson regression model is not suited for the prediction before the event. In addition, if the event has some surprising element on the date of the event, then it is naturally difficult to predict it before the event. We could use some characteristic of a particular event for the purpose of prediction. For example, national holiday has a fixed date every year and we can analyze stability of the pattern from year to year.

References


Table 1: Social networking data in Japan, 2014

<table>
<thead>
<tr>
<th>Date</th>
<th>ID</th>
<th>Searchword</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/13</td>
<td>Seijin</td>
<td>Seijin no Hi</td>
<td>Coming-of-Age Day</td>
</tr>
<tr>
<td>1/26</td>
<td>O-Marathon</td>
<td>Osaka International Ladies Marathon</td>
<td></td>
</tr>
<tr>
<td>2/3</td>
<td>Setsubun</td>
<td>Setsubun</td>
<td>Bean Throwing Night</td>
</tr>
<tr>
<td>2/8</td>
<td>Sochi</td>
<td>Sochi Olympics</td>
<td>The opening ceremony took place February 7, local time.</td>
</tr>
<tr>
<td>2/9</td>
<td>Uemura</td>
<td>Aiko Uemura</td>
<td>The women’s moguls final</td>
</tr>
<tr>
<td>2/11</td>
<td>Kenkoku</td>
<td>Kenkokukinenbi</td>
<td>National Foundation Day</td>
</tr>
<tr>
<td>2/14</td>
<td>Valentine</td>
<td>Valentine’s Day</td>
<td></td>
</tr>
<tr>
<td>2/15</td>
<td>Hanyu</td>
<td>Yuzuru Hanyu</td>
<td>Figure skating men’s singles free</td>
</tr>
<tr>
<td>2/16</td>
<td>Kasai</td>
<td>Noriaki Kasai</td>
<td>Large hill individual men final</td>
</tr>
<tr>
<td>2/21</td>
<td>Asada</td>
<td>Mao Asada</td>
<td>Figure skating ladies’s singles free</td>
</tr>
<tr>
<td>2/23</td>
<td>T-Marathon</td>
<td>Tokyo Marathon</td>
<td></td>
</tr>
<tr>
<td>3/3</td>
<td>Academy</td>
<td>Academy Awards</td>
<td>The Oscar ceremonies took place March 2, local time.</td>
</tr>
<tr>
<td>3/11</td>
<td>Earthquake</td>
<td>The Great East Japan Earthquake</td>
<td></td>
</tr>
<tr>
<td>3/14</td>
<td>White</td>
<td>White Day</td>
<td></td>
</tr>
<tr>
<td>3/21</td>
<td>Shunbun</td>
<td>Shunbun no Hi</td>
<td>Vernal Equinox Day</td>
</tr>
<tr>
<td>3/24</td>
<td>mayoral</td>
<td>Osaka's mayoral elections</td>
<td></td>
</tr>
<tr>
<td>5/5</td>
<td>Kodomo</td>
<td>Kodomo no Hi</td>
<td>Children’s Day</td>
</tr>
<tr>
<td>6/9</td>
<td>Oshima</td>
<td>Yuko Oshima</td>
<td>AKB graduation performance</td>
</tr>
<tr>
<td>6/15</td>
<td>W-cup</td>
<td>The World Cup</td>
<td>In Japan’s opening match against Cote d’Ivoire</td>
</tr>
<tr>
<td>6/21</td>
<td>Geshi</td>
<td>Geshi</td>
<td>the summer solstice</td>
</tr>
</tbody>
</table>
Table 2: Parameter estimation of the power-law decay independence model

<table>
<thead>
<tr>
<th>ID</th>
<th>$\alpha_b$</th>
<th>$\beta_b$</th>
<th>$\gamma$</th>
<th>$\alpha_u$</th>
<th>$\beta_u$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seijin</td>
<td>0.889</td>
<td>1.493</td>
<td>53664</td>
<td>0.908</td>
<td>1.463</td>
</tr>
<tr>
<td>O-Marathon</td>
<td>7.902</td>
<td>0.858</td>
<td>993</td>
<td>0.207</td>
<td>4.337</td>
</tr>
<tr>
<td>Setsubun</td>
<td>2.694</td>
<td>1.318</td>
<td>105334</td>
<td>0.397</td>
<td>2.921</td>
</tr>
<tr>
<td>Sochi</td>
<td>0.179</td>
<td>5.499</td>
<td>5479</td>
<td>1.089</td>
<td>1.65</td>
</tr>
<tr>
<td>Uemura</td>
<td>0.346</td>
<td>3.256</td>
<td>4992</td>
<td>0.231</td>
<td>3.436</td>
</tr>
<tr>
<td>Kenkoku</td>
<td>1.405</td>
<td>2.012</td>
<td>7708</td>
<td>1.021</td>
<td>2.122</td>
</tr>
<tr>
<td>Valentine</td>
<td>3.498</td>
<td>0.951</td>
<td>55539</td>
<td>0.742</td>
<td>1.991</td>
</tr>
<tr>
<td>Hanyu</td>
<td>0.898</td>
<td>1.823</td>
<td>10317</td>
<td>2.353</td>
<td>0.891</td>
</tr>
<tr>
<td>Kasai</td>
<td>599789.772</td>
<td>0.221</td>
<td>3368</td>
<td>0.131</td>
<td>4.584</td>
</tr>
<tr>
<td>Asada</td>
<td>0.447</td>
<td>2.755</td>
<td>18898</td>
<td>1.289</td>
<td>1.047</td>
</tr>
<tr>
<td>T-Marathon</td>
<td>2.327</td>
<td>1.186</td>
<td>8702</td>
<td>0.619</td>
<td>1.948</td>
</tr>
<tr>
<td>Academy</td>
<td>39.489</td>
<td>0.423</td>
<td>2923</td>
<td>3.97</td>
<td>0.302</td>
</tr>
<tr>
<td>Earthquake</td>
<td>10.695</td>
<td>0.822</td>
<td>41580</td>
<td>2.611</td>
<td>1.185</td>
</tr>
<tr>
<td>White</td>
<td>6.285</td>
<td>0.691</td>
<td>45488</td>
<td>0.309</td>
<td>2.483</td>
</tr>
<tr>
<td>Shunbun</td>
<td>2.99</td>
<td>1.259</td>
<td>14379</td>
<td>0.902</td>
<td>1.969</td>
</tr>
<tr>
<td>Mayoral</td>
<td>0.096</td>
<td>7.803</td>
<td>1632</td>
<td>1.217</td>
<td>1.375</td>
</tr>
<tr>
<td>Kodomo</td>
<td>14.37</td>
<td>0.783</td>
<td>29200</td>
<td>0.637</td>
<td>2.316</td>
</tr>
<tr>
<td>Oshima</td>
<td>1.083</td>
<td>1.136</td>
<td>4456</td>
<td>0.673</td>
<td>1.23</td>
</tr>
<tr>
<td>Wcup</td>
<td>0.013</td>
<td>34.707</td>
<td>45925</td>
<td>1.091</td>
<td>0.787</td>
</tr>
<tr>
<td>Geshi</td>
<td>8.096</td>
<td>0.746</td>
<td>8183</td>
<td>0.766</td>
<td>1.498</td>
</tr>
</tbody>
</table>
### Table 3: Parameter estimation of the AR(2) model

<table>
<thead>
<tr>
<th>ID</th>
<th>AR(2) model</th>
<th>( s = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \alpha )</td>
<td>( \beta )</td>
</tr>
<tr>
<td>Seijin</td>
<td>0.612</td>
<td>1.834</td>
</tr>
<tr>
<td>O-Marathon</td>
<td>0.105</td>
<td>7.517</td>
</tr>
<tr>
<td>Setsubun</td>
<td>0.265</td>
<td>3.944</td>
</tr>
<tr>
<td>Sochi</td>
<td>0.946</td>
<td>1.774</td>
</tr>
<tr>
<td>Uemura</td>
<td>0.276</td>
<td>3.173</td>
</tr>
<tr>
<td>Kenkoku</td>
<td>0.688</td>
<td>2.720</td>
</tr>
<tr>
<td>Valentine</td>
<td>0.722</td>
<td>2.040</td>
</tr>
<tr>
<td>Hanyu</td>
<td>1.919</td>
<td>0.951</td>
</tr>
<tr>
<td>Kasai</td>
<td>0.097</td>
<td>5.667</td>
</tr>
<tr>
<td>Asada</td>
<td>1.486</td>
<td>0.984</td>
</tr>
<tr>
<td>T-Marathon</td>
<td>0.331</td>
<td>2.934</td>
</tr>
<tr>
<td>Academy</td>
<td>0.692</td>
<td>0.743</td>
</tr>
<tr>
<td>Earthquake</td>
<td>2.615</td>
<td>1.207</td>
</tr>
<tr>
<td>White</td>
<td>0.313</td>
<td>2.485</td>
</tr>
<tr>
<td>Shunbun</td>
<td>1.030</td>
<td>1.821</td>
</tr>
<tr>
<td>Mayoral</td>
<td>1.175</td>
<td>1.385</td>
</tr>
<tr>
<td>Kodomo</td>
<td>0.493</td>
<td>2.733</td>
</tr>
<tr>
<td>Oshima</td>
<td>0.480</td>
<td>1.530</td>
</tr>
<tr>
<td>Wcup</td>
<td>0.930</td>
<td>0.859</td>
</tr>
<tr>
<td>Geshi</td>
<td>0.552</td>
<td>1.789</td>
</tr>
</tbody>
</table>
Table 4: Parameter estimation of the unifying model

<table>
<thead>
<tr>
<th>ID</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>$w$</th>
<th>$u$</th>
<th>$v$</th>
<th>log-lik.</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>seijin</td>
<td>0.735</td>
<td>1.618</td>
<td>53188</td>
<td>0.643</td>
<td>0</td>
<td>1</td>
<td>-329.506</td>
<td>669.013</td>
</tr>
<tr>
<td>O-Marathon</td>
<td>0.177</td>
<td>4.857</td>
<td>992</td>
<td>0.287</td>
<td>0.005</td>
<td>0.996</td>
<td>-53.204</td>
<td>116.407</td>
</tr>
<tr>
<td>Setsubun</td>
<td>0.398</td>
<td>2.919</td>
<td>105410</td>
<td>0.998</td>
<td>0.999</td>
<td>0.999</td>
<td>-2092.256</td>
<td>4194.512</td>
</tr>
<tr>
<td>Sochi</td>
<td>1.053</td>
<td>1.665</td>
<td>5374</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-37.463</td>
<td>84.927</td>
</tr>
<tr>
<td>Uemura</td>
<td>0.286</td>
<td>3.055</td>
<td>5348</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-68.389</td>
<td>146.779</td>
</tr>
<tr>
<td>Kenkoku</td>
<td>0.817</td>
<td>2.404</td>
<td>7679</td>
<td>0.670</td>
<td>0</td>
<td>1</td>
<td>-58.269</td>
<td>126.538</td>
</tr>
<tr>
<td>Valentine</td>
<td>0.741</td>
<td>1.993</td>
<td>55456</td>
<td>0.007</td>
<td>0.988</td>
<td>0.561</td>
<td>-145.248</td>
<td>300.495</td>
</tr>
<tr>
<td>Hanyu</td>
<td>2.168</td>
<td>0.901</td>
<td>9969</td>
<td>0.993</td>
<td>0.702</td>
<td>0.996</td>
<td>-71.639</td>
<td>153.279</td>
</tr>
<tr>
<td>Kasai</td>
<td>0.135</td>
<td>4.488</td>
<td>3381</td>
<td>0</td>
<td>0.966</td>
<td>0.999</td>
<td>-138.561</td>
<td>287.121</td>
</tr>
<tr>
<td>Asada</td>
<td>1.296</td>
<td>1.045</td>
<td>18939</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-109.559</td>
<td>229.119</td>
</tr>
<tr>
<td>T-Marathon</td>
<td>0.433</td>
<td>2.388</td>
<td>8688</td>
<td>0.674</td>
<td>0.002</td>
<td>0.995</td>
<td>-180.467</td>
<td>370.934</td>
</tr>
<tr>
<td>Academy</td>
<td>4.524</td>
<td>0.288</td>
<td>2923</td>
<td>0.977</td>
<td>0.920</td>
<td>0.990</td>
<td>-604.371</td>
<td>1218.743</td>
</tr>
<tr>
<td>Earthquake</td>
<td>2.710</td>
<td>1.173</td>
<td>41566</td>
<td>0.429</td>
<td>1</td>
<td>0</td>
<td>-277.531</td>
<td>565.061</td>
</tr>
<tr>
<td>White</td>
<td>0.310</td>
<td>2.475</td>
<td>45438</td>
<td>0.010</td>
<td>0.997</td>
<td>0.873</td>
<td>-220.389</td>
<td>450.779</td>
</tr>
<tr>
<td>Shunbun</td>
<td>0.901</td>
<td>1.969</td>
<td>14369</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-59.885</td>
<td>129.770</td>
</tr>
<tr>
<td>Mayoral</td>
<td>1.145</td>
<td>1.396</td>
<td>1581</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-37.753</td>
<td>85.506</td>
</tr>
<tr>
<td>Kodomo</td>
<td>0.493</td>
<td>2.733</td>
<td>29201</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>-104.649</td>
<td>219.298</td>
</tr>
<tr>
<td>Oshima</td>
<td>0.674</td>
<td>1.229</td>
<td>4458</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-120.807</td>
<td>251.614</td>
</tr>
<tr>
<td>cup</td>
<td>0.966</td>
<td>0.815</td>
<td>44480</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-1249.637</td>
<td>2509.273</td>
</tr>
<tr>
<td>Geshi</td>
<td>0.632</td>
<td>1.619</td>
<td>8174</td>
<td>0.814</td>
<td>0</td>
<td>1</td>
<td>-137.969</td>
<td>285.939</td>
</tr>
</tbody>
</table>
Table 5: Comparison of models based on AIC

<table>
<thead>
<tr>
<th>ID</th>
<th>AIC</th>
<th>$w = 1$</th>
<th>$u = 1$</th>
<th>$u = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$u$ AIC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seijin</td>
<td>669.013</td>
<td>0.39</td>
<td>667.378</td>
<td>732.505</td>
</tr>
<tr>
<td>O-Marathon</td>
<td>116.407</td>
<td>0.764</td>
<td>112.277</td>
<td>111.576</td>
</tr>
<tr>
<td>Setsubun</td>
<td>4194.512</td>
<td>1</td>
<td>4190.483</td>
<td>4188.482</td>
</tr>
<tr>
<td>Sochi</td>
<td>84.927</td>
<td>1</td>
<td>80.927</td>
<td>78.927</td>
</tr>
<tr>
<td>Uemura</td>
<td>146.779</td>
<td>1</td>
<td>142.777</td>
<td>140.777</td>
</tr>
<tr>
<td>Kenkoku</td>
<td>126.538</td>
<td>0.334</td>
<td>122.515</td>
<td>127.391</td>
</tr>
<tr>
<td>Valentaine</td>
<td>300.495</td>
<td>1</td>
<td>311.419</td>
<td>309.419</td>
</tr>
<tr>
<td>Hanyu</td>
<td>153.279</td>
<td>0.705</td>
<td>149.28</td>
<td>150.843</td>
</tr>
<tr>
<td>Kasai</td>
<td>287.121</td>
<td>1</td>
<td>283.097</td>
<td>281.095</td>
</tr>
<tr>
<td>Asada</td>
<td>229.119</td>
<td>1</td>
<td>225.119</td>
<td>223.119</td>
</tr>
<tr>
<td>T-Marathon</td>
<td>370.934</td>
<td>0.41</td>
<td>364.891</td>
<td>394.555</td>
</tr>
<tr>
<td>Academy</td>
<td>1218.743</td>
<td>0.905</td>
<td>1212.032</td>
<td>1217.415</td>
</tr>
<tr>
<td>Earthquake</td>
<td>565.061</td>
<td>1</td>
<td>644.564</td>
<td>642.564</td>
</tr>
<tr>
<td>White</td>
<td>450.779</td>
<td>1</td>
<td>449.023</td>
<td>447.023</td>
</tr>
<tr>
<td>Shunbun</td>
<td>129.77</td>
<td>1</td>
<td>125.77</td>
<td>123.77</td>
</tr>
<tr>
<td>Mayoral</td>
<td>85.506</td>
<td>1</td>
<td>81.506</td>
<td>79.506</td>
</tr>
<tr>
<td>Kodomo</td>
<td>219.298</td>
<td>0</td>
<td>215.298</td>
<td>273.882</td>
</tr>
<tr>
<td>Oshima</td>
<td>251.614</td>
<td>1</td>
<td>247.613</td>
<td>245.613</td>
</tr>
<tr>
<td>cup</td>
<td>2509.273</td>
<td>1</td>
<td>2505.273</td>
<td>2503.273</td>
</tr>
<tr>
<td>Geshi</td>
<td>285.939</td>
<td>0.211</td>
<td>282.618</td>
<td>357.08</td>
</tr>
</tbody>
</table>