利用布朗運動模型估計急診滯留人數

Using Brownian motion model to estimate the number of patients staying in the emergency department

學生：翁梓瑄　　指導老師：王彥雯 老師

Introduction

- The number of patients in the emergency department (ED) is influenced by two factors: the length of stay of each patient in ED and the hospitalization rate of emergency patients.
- Increased length of stay in the ED **can lead to adverse effects**, higher mortality rate, and longer overall length of stays for patients admitted from ED.
- Understanding how the number of patients staying in ED fluctuates may help hospitals **identify trends** and be better prepared for busier periods.
- Our objective is to use a **Brownian motion model** to estimate the number of in-ED patients at a given time, to predict busy hours and help hospitals **allocate medical resources and staff members** more efficiently.

Methodology

1. Brownian Motion Model

- A stochastic process $\{X(t), t \geq 0\}$ is called Brownian motion if :
  1. X(0) = 0, and X(t) is continuous
  2. increments are normally distributed
  3. increments that do not overlap in time are independent, i.e. if $0 \leq t_1 < t_2 < \cdots < t_n$, then $X(t_k) - X(t_{k-1})$, $1 \leq k \leq n$ are independent.
- When describing an event that is always positive, exponential transformation is performed to get $Y(t) = e^{X(t)}$, which we call **geometric Brownian motion**.
- When estimating the number of patients staying in the emergency department, it is possible for the number of patients to **reach zero** at some point, and for the number of patients to rise rapidly due to major events
- We consider the model:

$$Y(t) = ce^{aB(t)+bt} + de^{-aB(t)-bt} - 2\sqrt{cd}$$

Where a, b, c, d are parameters that require estimation, t is the time point of each data point and $B(t)$ is a random Brownian motion.

2. Parameter Estimation

2.1 Method 1 – Estimation Formula:

To estimate the parameters of $Y(t)$, Itô's formula is used to present $Y(t)$ as an Itô process. From the **approximation of quadratic variance** and the model's initial value, the estimation formula for each parameter can be derived.

2.2 Method 2 – Dirichlet Process:

Estimate the parameters of $Y(t)$ using **Dirichlet process** to get the **posterior distribution** of the parameters. The model is:

$$Y(t) \sim f(y|\theta)$$

$$B(t) \sim N(0, t)$$

$$\theta = (a, b, c, d) \sim \pi_\theta$$

$$\pi_\theta \sim \mathrm{DP}(\alpha, G_0)$$

$B(t)$ is a Brownian motion with parameters $(\mu, \sigma^2) = (0, t)$. Mixing distribution π_θ is assigned a Dirichlet process prior with a base measure G_0 and concentration parameter α. We let $\theta = (a, b, c, d)$ with **prior density** $\boldsymbol{\pi_\theta}$. Parameter b mainly controls how $Y(t)$ changes with time. The prior distribution of θ is:

$$a \sim Gamma(\alpha_1, \beta_1)$$

$$c \sim Gamma(\alpha_2, \beta_2)$$

$$d \sim Gamma(\alpha_3, \beta_3)$$

$$b \sim N(\mu, \sigma^2)$$

3. Emergency Department Data

- Data was obtained from the Taipei Municipal Wanfang Hospital
- This study included all emergency department patients seen by ED physicians between January 1, 2018 and December 31, 2019
- Excluded patients seen by gynecologists, pediatricians, or dentists. Additionally, patients who were revisited, expired, transferred, or eloped and those lacking complete data were also excluded.

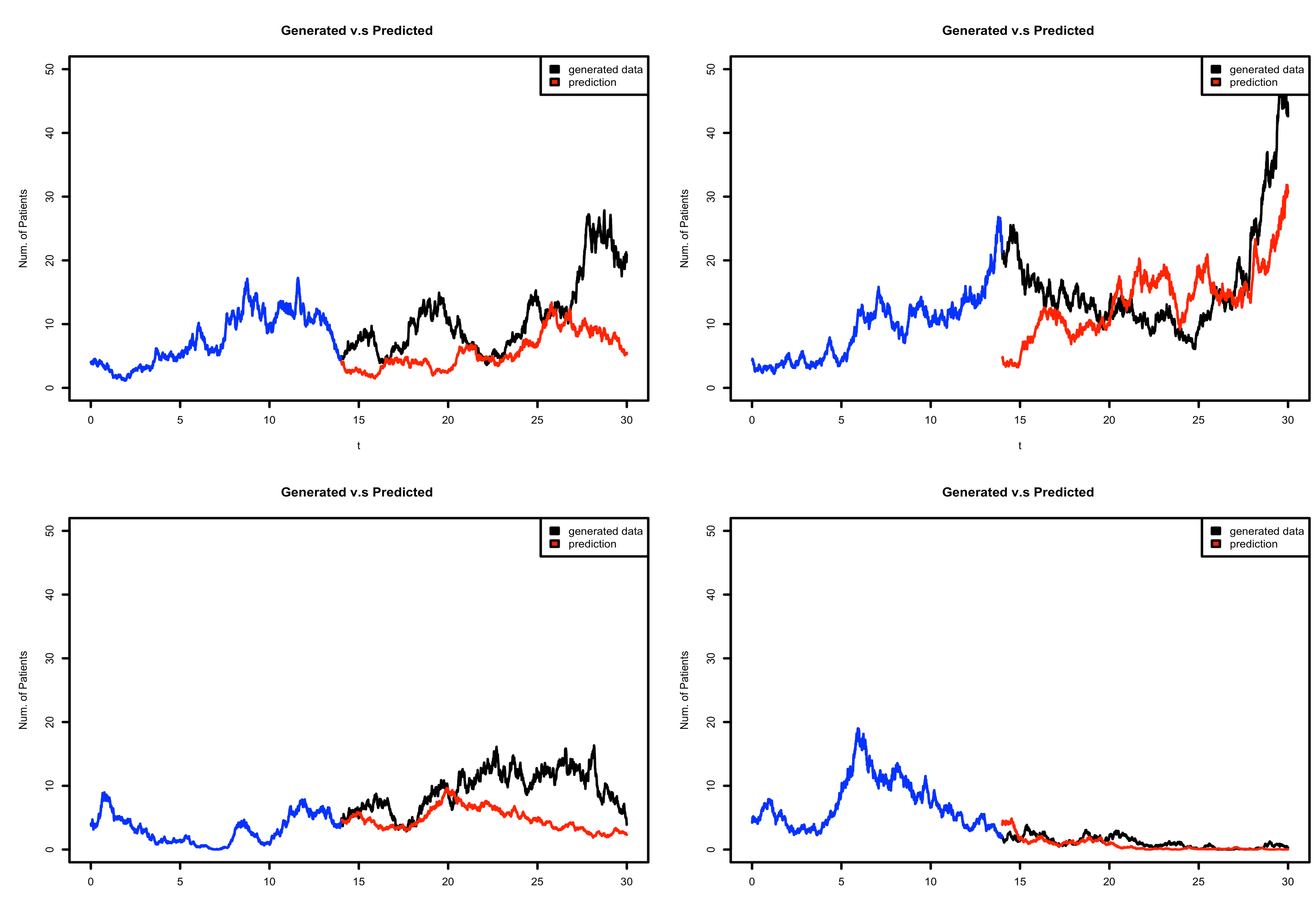
Results

1.1 Simulation Results Using Method 1 (Estimation Formula)

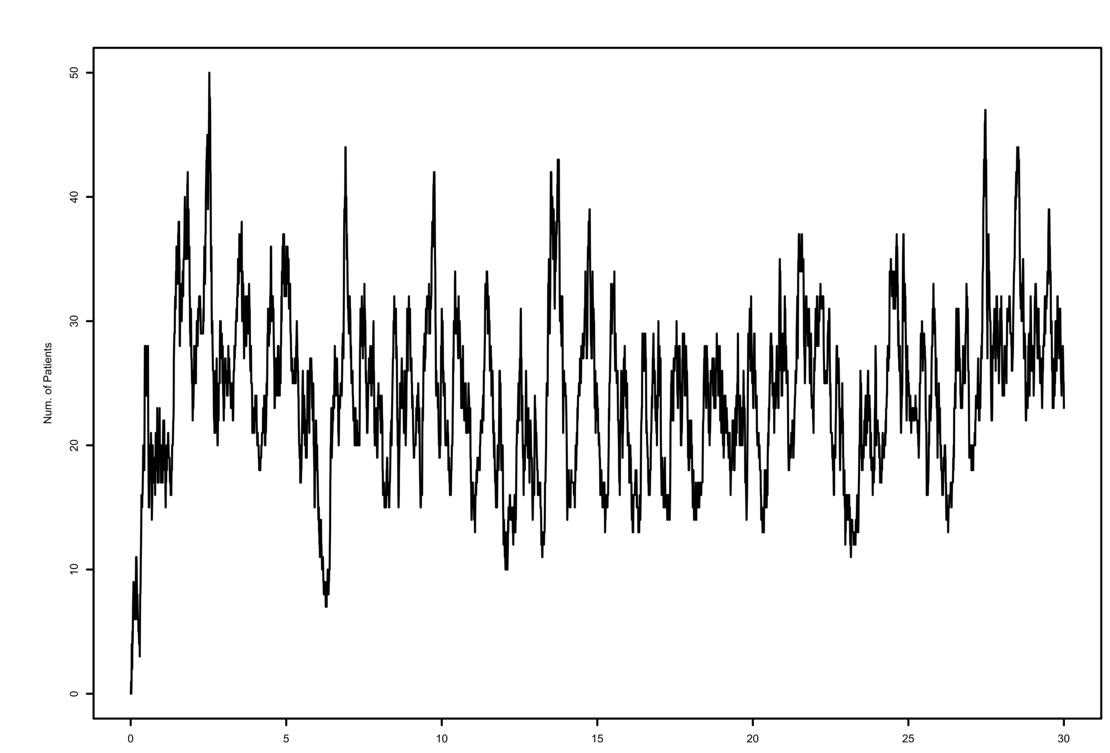
- Use data generated with random Brownian motion with given parameters
- Generate 1 month of data (black), then using 2 weeks of data (blue) to predict the remaining 2 weeks (red)
- Data points with 10 minute intervals, using 1 day as a unit ($\Delta t = 1/144$)
- Simulation Estimation Results (100 runs):

	Mean	Median	SD	Real Value
a	0.1956	0.1957	0.0520	0.2
b	0.00062	0.00133	0.0012	0.01
c	8.470	8.094	1.7868	12
d	0.7945	0.6480	0.7078	2

- Plots of Generated Data v.s Model Prediction:



- Real Data from the Emergency Department



- ✓ Different fluctuation patterns compared to generated data
- ✓ Fluctuates more within a shorter time frame (within a single day)
- ✓ Remains relatively steady across the entire month
- ✓ May impact the estimation of model parameters

1.2 Issues & Findings of Method 1

- Encounters runs with negative a^2 estimates (decreases when Δt is smaller)
- Compared to parameters given to generate data, b, c, d are usually underestimated using this method.
- Highly inaccurate estimation/negative a^2 estimate when used on real data

Discussion

- This model performs better with smaller time increments ($\downarrow \Delta t$)
- Optimal method for parameter estimation may vary with data charachteristics

  → Estimation using method 1 works better when there are significant changes in trends over a longer period of time (days) as seen in the simulation data
- As seen with the simulation results, this model has the ability to detect trends and fluctuations, but parameter estimation remains an issue when applied to real data.