Project 2
Computer Intensive Statistical Methods

Erling Fause Steen og Christian Oppegård Moen

08 03 2022

Contents

```r
library(ggplot2)
```

Problem 1

We will look at a portion of the Tokyo rainfall dataset. The response is whether the amount of rainfall exceed 1 mm over a time period, and is given by

$$y_t|\tau_t \sim \text{Bin}(n_t, \pi(\tau_t)), \quad \pi(\tau_t) = \frac{\exp(\tau_t)}{1 + \exp(\tau_t)} = \frac{1}{1 + \exp(-\tau_t)}. \tag{1}$$

a)

We start by downloading the Tokyo rainfall dataset and plot the response as a function of t.

```r
# Read the rain data. Might have to change wd. Should knit without.
# setwd(./Exe2bsm)
load("./rain.rda")
```

We start by plotting the amount of rainfall against each day.

```r
## Plotting the data
head(rain)
```

```
##   day n.years n.rain
## 1   1      39      8
## 2   2      39      7
## 3   3      39      8
## 4   4      39     11
## 5   5      39      8
## 6   6      39      6
```

```
ggplot(data = rain, mapping = aes(x = day, y = n.rain)) + geom_line() + xlab("Day") +
    ylab("Number of days with more than 1 mm rain")
```
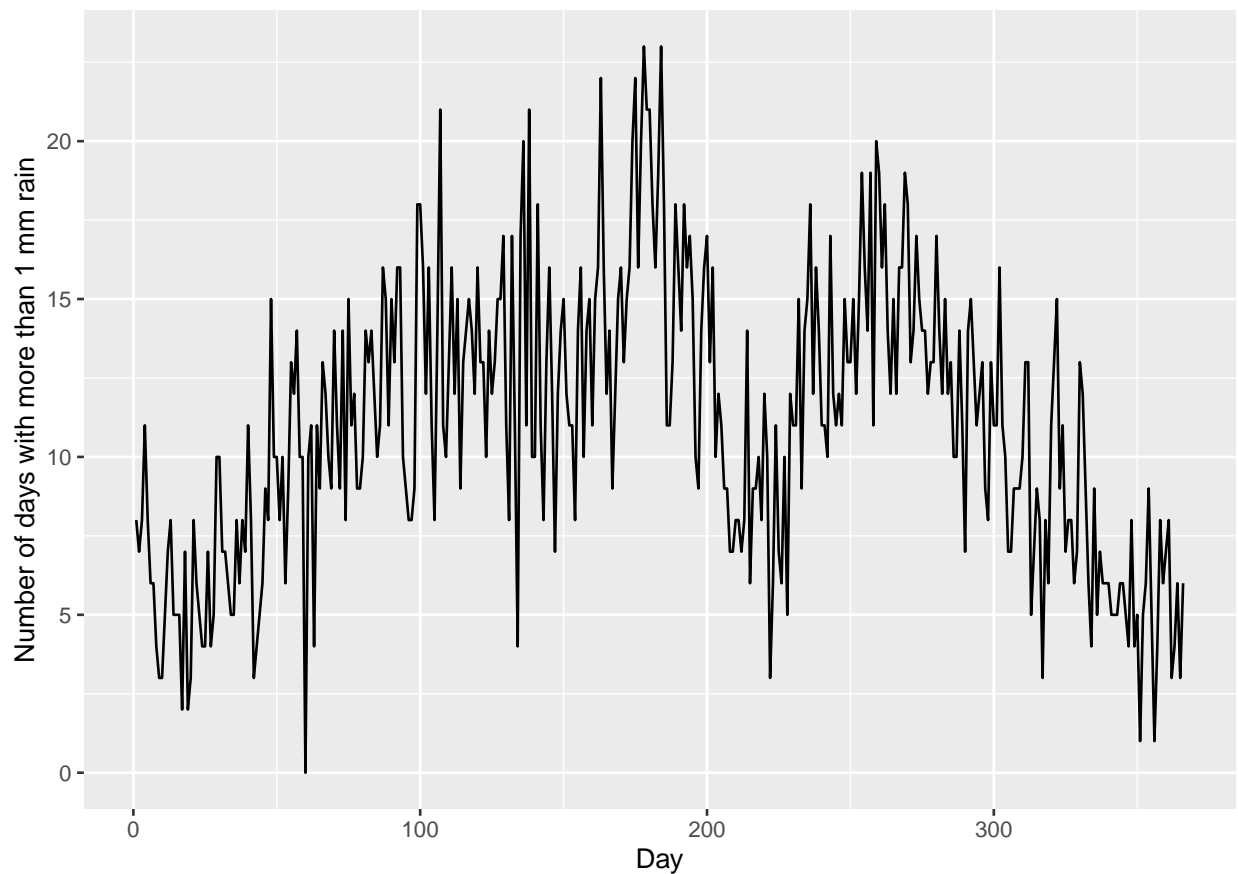

From the plot, we can see that there are fewer days in the start of the year and in the end of the year with an amount of rainfall over 1 mm. This is in January and December. The number of days steadily increases until the beggining of the summer which seems to be the period with the most days with an amount of rainfall over 1 mm. Then, the amount of days decreases during july and august before increasing during the autumn. This is somewhat consistent with the results we get from googling the amount of days with precipitation in Tokyo, where we can see that June and September are the months with the most days with rainfall and December and January are the month with the most.

b) Likelihod

The likelihood of Equation (1) is given by

$$
\begin{aligned}
L(\pi(\tau_t)) &= \prod_{i=1}^{T} \binom{n_t}{y_t} \pi(\tau_t)^{y_t}(1-\pi(\tau_t))^{n_t-y_t} \\
&\propto \prod_{i=1}^{T} \pi(\tau_t)^{y_t}(1-\pi(\tau_t))^{n_t-y_t} \\
&= \prod_{t=1}^{T} \left(\frac{\exp(\tau_t)}{1+\exp(\tau_t)}\right)^{y_t}\left(1-\frac{\exp(\tau_t)}{1+\exp(\tau_t)}\right)^{n_t-y_t},
\end{aligned}
$$

where $y_t = 1, 2, ..., 39$ and $n_t = 39$ for $t \neq 60$, and $y_t = 1, 2, ..., 10$ and $n_t = 10$ for $t \neq 60$.

c) Posterior

$$
\begin{aligned}
P(\sigma^2|\tau, y) &= \frac{P(\sigma_u^2, \tau, y)}{P(\tau, y)} \\
&\propto P(y|\sigma_u^2, \tau)P(\sigma_u^2, \tau) \\
&= P(y|\sigma_u^2, \tau)P(\tau|\sigma_u^2)P(\sigma_u^2) \\
&= \underbrace{\prod_{t=1}^{T}\left(\frac{\exp(\tau_t)}{1+\exp(\tau_t)}\right)^{y_t}\left(1-\frac{\exp(\tau_t)}{1+\exp(\tau_t)}\right)^{n_t-y_t}}_{\text{Constant wrt } \sigma^2}.
\end{aligned}
$$

$$
\prod_{t=1}^{T}\frac{1}{\sigma_u}\exp\left\{\frac{1}{2\sigma_u^2}(\tau_t-\tau_{t-1})^2\right\}\cdot\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma_u^2}\right)^{\alpha+1}exp\left\{-\frac{\beta}{\sigma_u^2}\right\}
$$

$$
\propto \prod_{t=1}^{T}\frac{1}{\sigma_u}\exp\left\{\frac{1}{2\sigma_u^2}(\tau_t-\tau_{t-1})^2\right\}\cdot\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma_u^2}\right)^{\alpha+1}exp\left\{-\frac{\beta}{\sigma_u^2}\right\}
$$

$$
= \frac{1}{\sigma_u^{T-1}}exp\left\{\frac{T-1}{2\sigma_u^2}\boldsymbol{\tau Q\tau}\right\}\cdot\frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\sigma_u^2}\right)^{\alpha+1}exp\left\{-\frac{\beta}{\sigma_u^2}\right\}
$$

which should be the core of an inverse gamma.

d) Acceptance probability

Let $\mathcal{I} \subseteq \{1, 2, ..., 366\}$ be a set of time indices, and let $-\mathcal{I} = \{1, 2, ..., 366\} \setminus \mathcal{I}$. Furthermore, let $\boldsymbol{\tau}'$ denote the proposed values for $\boldsymbol{\tau}$. Then, by using iterative conditioning, the acceptance probability is given by

$$
\alpha(\boldsymbol{\tau}_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y}) = \min\left(1, \frac{P(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y})}{P(\boldsymbol{\tau}_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y})}\frac{Q(\boldsymbol{\tau}_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y})}{Q(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y})}\right),
$$

where our prior proposal distribution is $Q(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y}) = P(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)$. By considering

$$
\begin{aligned}
P(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y}) &= \frac{P(\boldsymbol{\tau}'_\mathcal{I}, \boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y})}{P(\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y})} \\
&= \frac{P(\boldsymbol{y}|\boldsymbol{\tau}'_\mathcal{I}, \boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)P(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)P(\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}{P(\boldsymbol{y}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)P(\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)} \\
&\quad\quad\quad \overset{\text{Conditionally independent}}{} \\
&= \frac{\overbrace{P(\boldsymbol{y}|\boldsymbol{\tau}'_\mathcal{I}, \boldsymbol{\tau}_{-\mathcal{I}})}\ \ P(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}{P(\boldsymbol{y}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)} \\
&= \frac{P(\boldsymbol{y}_\mathcal{I}|\boldsymbol{\tau}'_\mathcal{I})P(\boldsymbol{y}_{-\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}})P(\boldsymbol{\tau}'_\mathcal{I}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}{P(\boldsymbol{y}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}
\end{aligned}
$$

and equally
$$P(\boldsymbol{\tau}_{\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y}) = \frac{P(\boldsymbol{y}_{\mathcal{I}}|\boldsymbol{\tau}_{\mathcal{I}})P(\boldsymbol{y}_{-\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}})P(\boldsymbol{\tau}_{\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}{P(\boldsymbol{y}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}$$

the acceptance probability becomes

$$\alpha(\boldsymbol{\tau}_{\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2, \boldsymbol{y}) = \min\left(1, \frac{P(\boldsymbol{y}_{\mathcal{I}}|\boldsymbol{\tau}_{\mathcal{I}}')P(\boldsymbol{y}_{-\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}})P(\boldsymbol{\tau}_{\mathcal{I}}'|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)/P(\boldsymbol{y}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}{P(\boldsymbol{y}_{\mathcal{I}}|\boldsymbol{\tau}_{\mathcal{I}})P(\boldsymbol{y}_{-\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}})P(\boldsymbol{\tau}_{\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)/P(\boldsymbol{y}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)} \frac{P(\boldsymbol{\tau}_{\mathcal{I}}|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}{P(\boldsymbol{\tau}_{\mathcal{I}}'|\boldsymbol{\tau}_{-\mathcal{I}}, \sigma_u^2)}\right)$$
$$= \min\left(1, \frac{P(\boldsymbol{y}_{\mathcal{I}}|\boldsymbol{\tau}_{\mathcal{I}}')}{P(\boldsymbol{y}_{\mathcal{I}}|\boldsymbol{\tau}_{\mathcal{I}})}\right)$$