Homework 6 - Statistical modelling and inference
Group 7

November 29, 2015

Students	**Student IDs**
Max van Esso	73539
Marco Fayet	125593
Felix Gutmann	125604

1

1.1

To show: Likelihood function of the Laplace distribution for μ and Λ:

$$
\begin{aligned}
P(t|\mu, \Lambda) &= \prod_{i=1}^{N} \frac{1}{2}\Lambda \exp\{-\Lambda|t_i - \mu|\} \\
&= (\frac{1}{2}\Lambda)^N \exp\{\sum_{i=N}^{N} -\Lambda|t_i - \mu|\}
\end{aligned}
$$

1.2

To show: $\mu_{MLE} = median(t_1, \ldots, t_n)$

Prerequisite information:

Definition of the sgn function:

$$sgn(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Prerequisite information:

In the first step we compute the log-likelihood function:

$$P(t|\mu, \Lambda) = (\frac{1}{2}\Lambda)^N \exp\{\sum_{i=N}^{N} -\Lambda|t_i - \mu|\}$$

$$\log P(t|\mu, \Lambda) = N \log \Lambda - N \log 2 + \sum_{i=1}^{N} -\Lambda|t_i - \mu|$$

In the next step we take the derivative with respect to μ:

$$\frac{\partial}{\partial \mu} = -\Lambda \sum_{i=1}^{N} \frac{t_i - \mu}{|t_i - \mu|} = 0$$

$$= \sum_{i=1}^{N} \frac{t_i - \mu}{|t_i - \mu|} = 0$$

Since the some in the last expression is either $(-)1$ or $(+)1$, it can can be expressed as the sgn function:

$$= \sum_{i=1}^{N} sgn(t_i - \mu) = 0$$

We assume the number of datapoints to be odd. Using the definition of the sgn function, we need one data point to be exaclty equal to μ and half of the data points to be greater and lower than μ to fullfil the last equation. From that it is obvious that μ has to be the median.

1.3

To show: $\Lambda_{MLE} = (\frac{1}{N}\sum_{i=1}^{N}|t_i - \mu|)^{-1}$

We start by taking the derivative of the log-likelihood function from the last exercise with respect to Λ and simplyfie it to get the result:

$$\frac{\partial}{\partial\Lambda} = \frac{N}{\Lambda} - \sum_{i=1}^{N}|t_i - \mu| = 0$$

$$\frac{N}{\Lambda} = \sum_{i=1}^{N}|t_i - \mu|$$

$$\Lambda_{MLE} = (\frac{1}{N}\sum_{i=1}^{N}|t_i - \mu|)^{-1}$$

1.4

To show: Under the equivariance property of the MLE: $\Sigma_{MLE} = 2(\frac{1}{N}|t_i - \mu|)2$

Equivariance property:

We know for any Laplace distribution $L(x|a,b)$ the variance is $2b^2$. The equivariance theorem basically states that the transformationen of a true parameter also applies to the estimated parameter (Wasserman, Theorem 9.14).

Solution:

Following the interpretation we can express the variance as:

$$\Sigma = var[t] = 2\frac{1}{\Lambda^2} = 2(\frac{1}{N}|t_i - \mu|)^2$$

2

2.1

To show: $\eta_n | t_n, \boldsymbol{x}_n, \boldsymbol{w}, q \sim Gam((\frac{v+1}{2}, \frac{v+qe_n^2}{2-1}))$

Prerequisite information:

- $t_n | \boldsymbol{x}_n, \boldsymbol{w}, \eta_n, q \sim N(\phi(\boldsymbol{x}_n)^T \boldsymbol{w}, (\eta_n, q)^{-1}\boldsymbol{I})$
- $\eta_n \sim Gam(\frac{v}{2}, \frac{v}{2-1})$
- $e_n := t_n - \phi(\boldsymbol{x}_n)^T \boldsymbol{w}$

Solution:

We define the two distributions explicitaly:

$$\begin{aligned}
N(t_n|\phi(\boldsymbol{x}_n)^T \boldsymbol{w}, \eta_n q) &= \frac{1}{\sqrt(2\pi)} \eta_n q \exp\{-\frac{1}{2}\eta_n q (t_\phi(\boldsymbol{x}_n)^T \boldsymbol{w})^T (t_n - \phi(\boldsymbol{x}_n)^T \boldsymbol{w})\} \\
Gam(\eta_n | \frac{v}{2}, \frac{v}{2-1}) &= \frac{(\frac{v}{2-1})^{\frac{v}{2}}}{(\frac{v}{2}-1)!} \eta_n^{\frac{v}{2}-1} e^{-(\frac{v-1}{2})\eta_n}
\end{aligned}$$

In the next step we compute the conditional probability:

$$\begin{aligned}
p(\eta_n | t_n, \boldsymbol{x}_n, \boldsymbol{w}, q) &\propto p(\eta_n)(t_n | \boldsymbol{x}_n, \boldsymbol{w}_n, q, \eta_n) \\
&= C * \eta_n^{\frac{v}{2}+\frac{1}{2}} \exp\{-(\frac{v}{2}-1)\eta_n - \eta_n q \frac{1}{2} t_\phi(\boldsymbol{x}_n)^T \boldsymbol{w})^T (t_n - \phi(\boldsymbol{x}_n)^T \boldsymbol{w})\} \\
&= C * \eta_n^{\frac{v+1}{2}-1} \exp\{-(\frac{v}{2}-1) - q\frac{1}{2}e_n^2)\eta_n\} \\
&= C * \eta_n^{\frac{v+1}{2}-1} \exp\{(1 - \frac{v+qe_n^2}{2})\eta_n\} \\
&= C * \eta_n^{\frac{v+1}{2}-1} \exp\{-(\frac{v+qe_n^2}{2}-1)\eta_n\} \\
&\sim Gam(\frac{v+1}{2}, \frac{v+qe_n^2}{2}-1)
\end{aligned}$$

2.2

To show: $Q(\boldsymbol{\theta}, \boldsymbol{\theta}') = \frac{N}{2} \log q - \frac{q}{2} (\boldsymbol{t} - \boldsymbol{\Phi w})^T diag(\mathbb{E}[\boldsymbol{\eta} | \boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\theta}](\boldsymbol{t} - \boldsymbol{\Phi w}) + C$

Solution:

We start with the definition of $Q(\cdot)$:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}') = \int \log p(\boldsymbol{t}, \boldsymbol{\eta} | \boldsymbol{\theta}) p(\boldsymbol{\eta} | \boldsymbol{t}, \boldsymbol{\theta}') d\boldsymbol{\eta}$$

Since we integrating out $\boldsymbol{\eta}$, we look for the expected value of it.

$$= \mathbb{E}[\log p(\boldsymbol{t}, \boldsymbol{\eta} | \boldsymbol{\theta})]$$

We re-express the complete log-likelihood and expand the log:

$$\begin{aligned} &= \mathbb{E}[\log(p(\boldsymbol{t} | \boldsymbol{\theta}) p(\boldsymbol{\eta} | \boldsymbol{t}, \boldsymbol{\theta}')] \\ &= \mathbb{E}[\log(p(\boldsymbol{t} | \boldsymbol{\theta}) + \log p(\boldsymbol{\eta} | \boldsymbol{t}, \boldsymbol{\theta}')] \end{aligned}$$

We notice that the last term goes to a constant. By pluggin in the result from the last exercise and vectorize it we get:

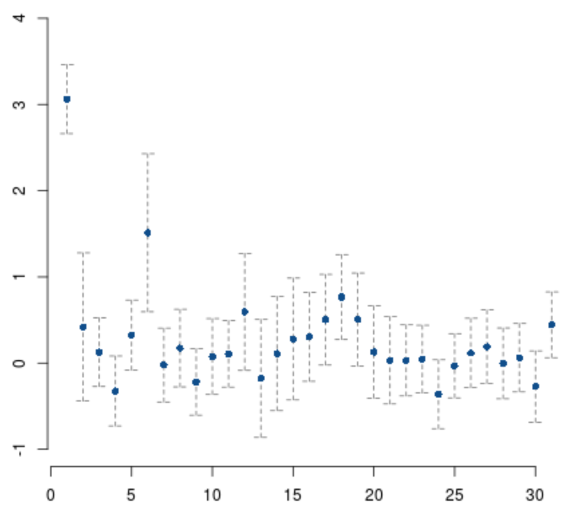
$$= \mathbb{E}[\frac{N}{2} \log q + \frac{q}{2} (\boldsymbol{t} - \boldsymbol{\Phi w})^T diag(\boldsymbol{\eta})(\boldsymbol{t} - \boldsymbol{\Phi w})] + C$$

Finally we can pull out non eta terms from the expected value to obtain the result:

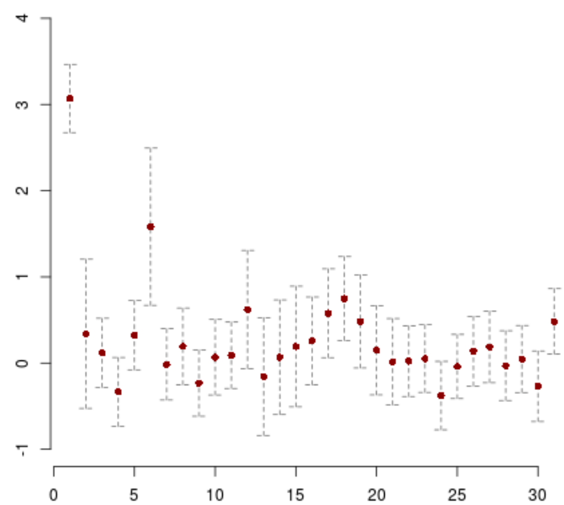
$$= \frac{N}{2} \log q + \frac{q}{2} (\boldsymbol{t} - \boldsymbol{\Phi w})^T diag(\mathbb{E}[\boldsymbol{\eta}])(\boldsymbol{t\Phi} - \boldsymbol{w})] + C$$

3 Exercises in R

3.1

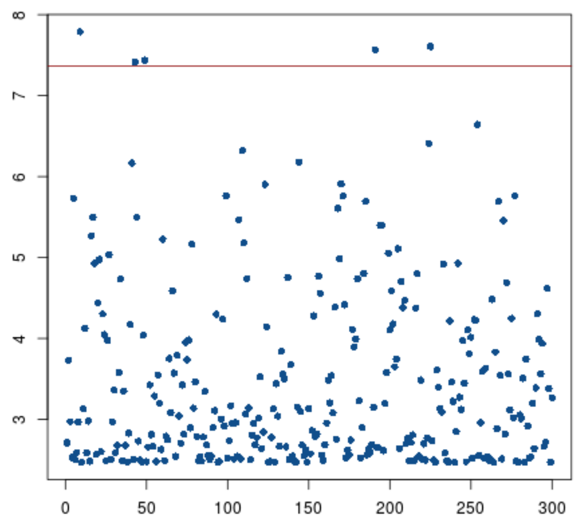


(a) Robust regression estimators (with 1.96 SE)

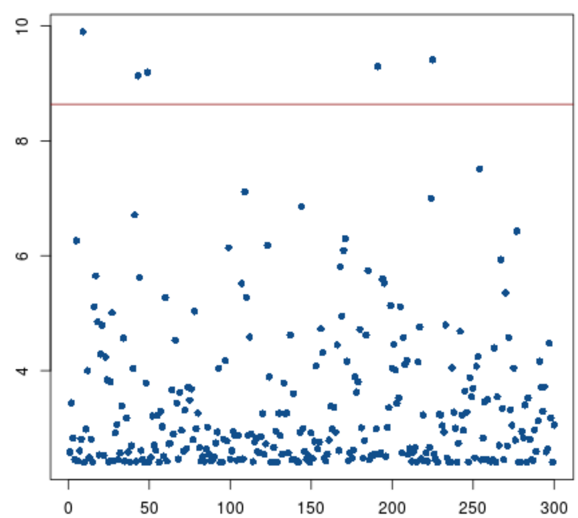


(b) MLE regression estimators (with 1.96 SE)

3.2

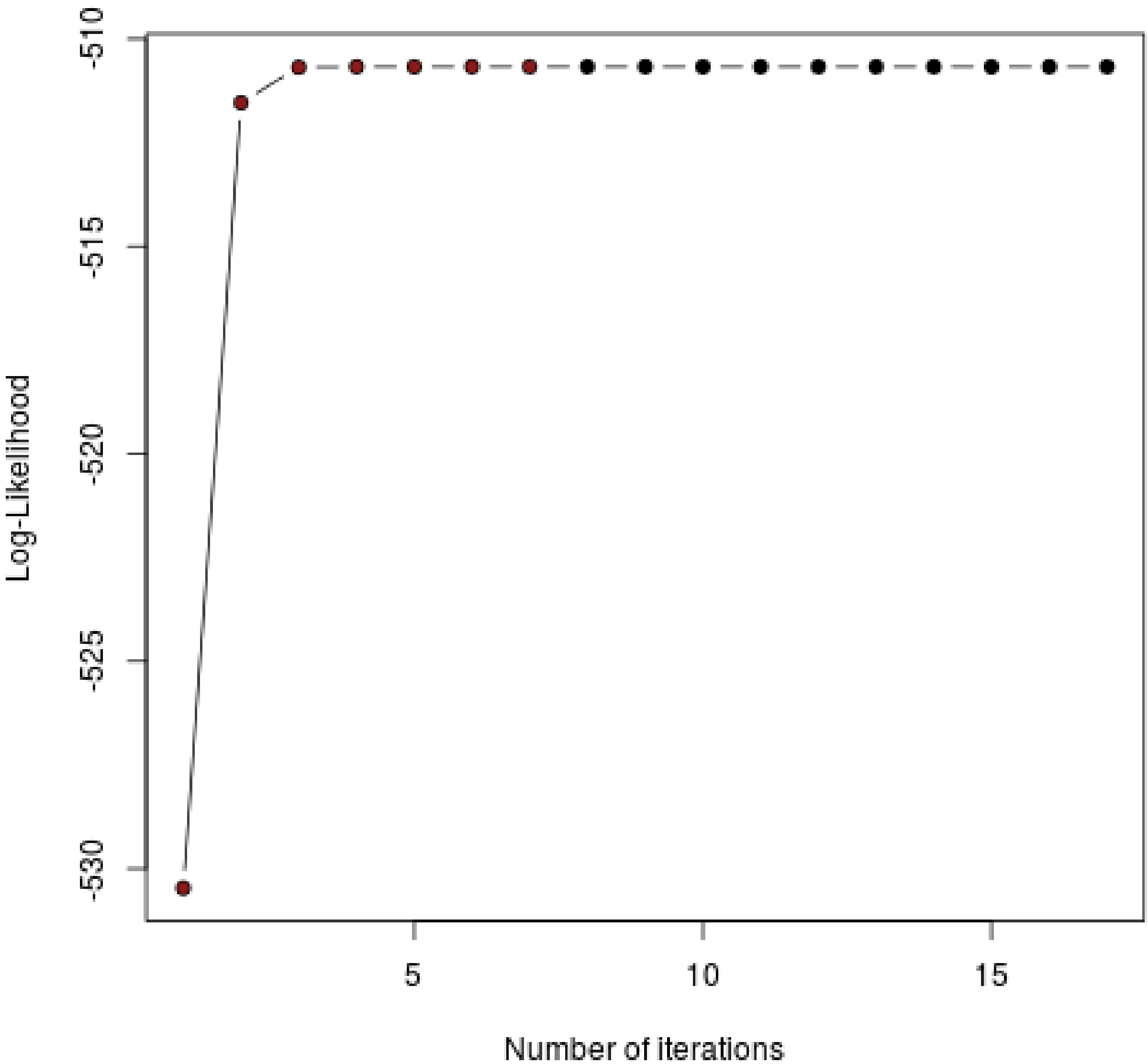


(a) Deviance residuals vs. simulation RR



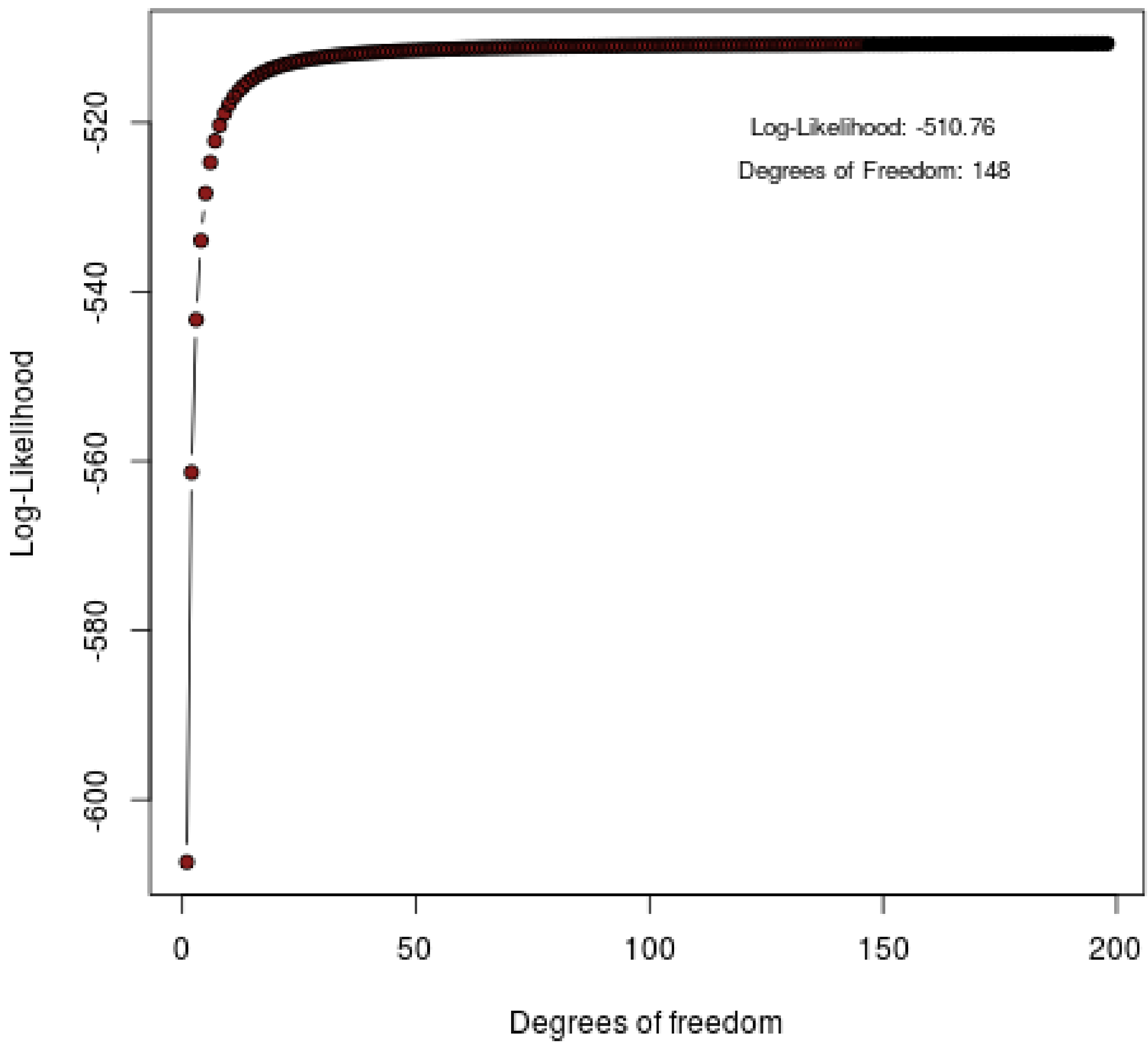
(b) Deviance residuals vs. simulation MLE

3.3



-510
-515
-520
-525
-530
Log-Likelihood
5
10
15
Number of iterations

3.4


Log-Likelihood: -510.76
Degrees of Freedom: 148
-520
-540
-560
-580
-600
Log-Likelihood
0
50
100
150
200
Degrees of freedom