Université catholique de Louvain

Louvain School of Statistics

LSTAT2130 - Bayesian Statistics

Project - Group Q

Lionel Lamy - 1294-1700
Adrien Kinart
Simon Lengendre

May 8, 2021

Contents

1 Introduction **2**

2 Question 1 **2**

2.1 (a) Theoretical probability 3

2.2 (b) Theoretical expression for the likelihood 3

3 Question 2 : Priors **4**

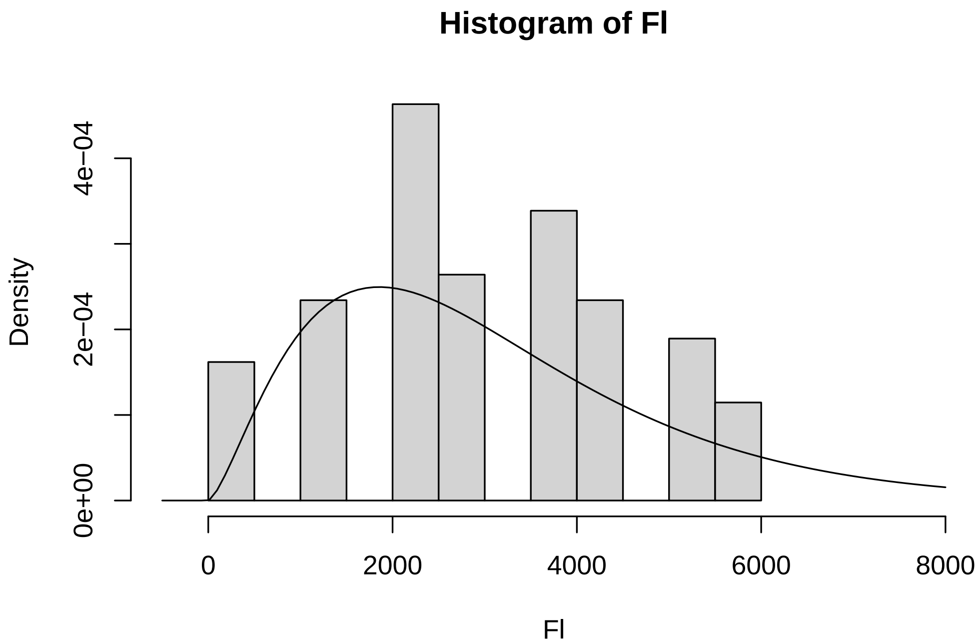
A Appendix **4**

A.1 Figures 4

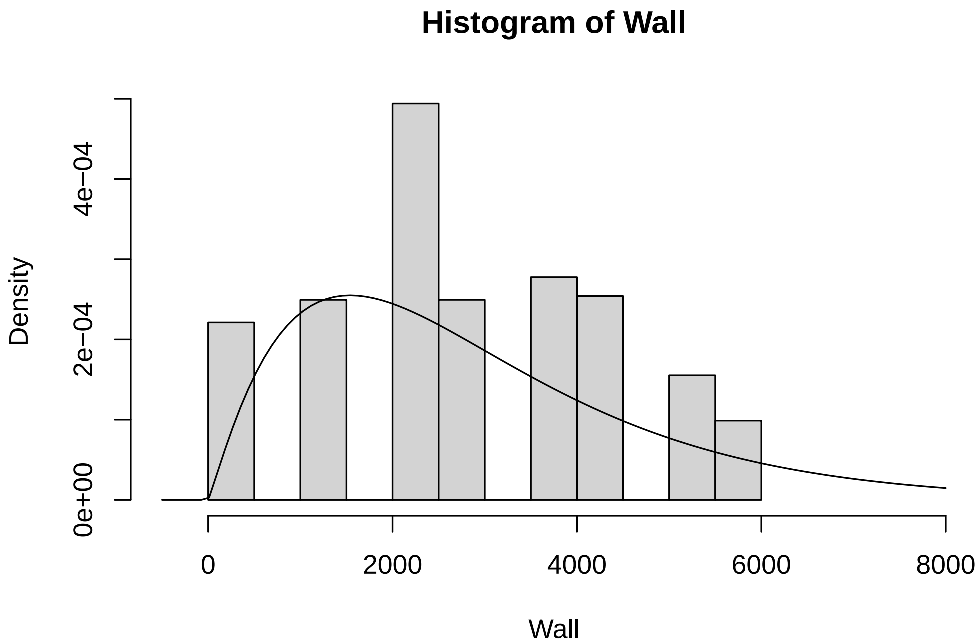
A.2 Code 4

1 Introduction

```
##    shape
## 3089.944
```




```
##    shape
## 2914.882
```



2 Question 1

Let $\theta_k := (\mu_k, \phi_k)$ be the set of parameters for a HNI with respect to region k.

2.1 (a) Theoretical probability

Let X be the monthly net income of 1123 Belgian households net income (HNI) older than 30 years. Regardless the 2 regions ($k = \{1, 2\}$ wrt Flanders and Wallonnia, respectively), is assumed it follows a Gamma distribution. It can be reparametrised in terms of its mean μ and dispersion parameter ϕ with the following trick:

$$\text{shape: } \kappa = \frac{1}{\phi}$$
$$\text{rate: } \lambda = \frac{1}{\phi\ \mu}$$

For both regions $k = \{1, 2\}$: This gives

$$f(x_k) = \frac{1}{\Gamma(\phi_k^{-1})} \big(\phi_k \mu\big)^{\phi_k} x_k^{\frac{1}{\phi_k}-1} \exp\big(\frac{-x_k}{\phi_k \mu}\big)$$

Then, the probability to fall into a certain HNI interval is:

$$P(x_{k,j} < x_k < x_{k,j+1}) = \int_{x_j}^{x_{j+1}} \frac{1}{\Gamma(\phi_k^{-1})} \big(\phi_k \mu\big)^{\phi_k} x_k^{\frac{1}{\phi_k}-1} \exp\big(\frac{-x_k}{\phi_k \mu}\big)$$

2.2 (b) Theoretical expression for the likelihood

On behalf of writing simplicity, the region index is removed. Since the frequency distribution in a given region is multinomial, i.e.

We have, writing $P := (p1, .., p_{10})$ and $X := (X_1, ...X_{10})$:

$$X|P \sim \text{Mul}(N, P) = \frac{x!}{x_1! \ldots x_{10}!} p_1^{x_1} \times \ldots \times p_{10}^{x_{10}} \text{ when} \sum_{j=1}^{10} p_j = 1$$
$$= 0 \text{ otherwise}$$

Up to a multiplicative constant, the likelihood can be written as follow:

$$L(\theta, D) = P(D|\mu, \phi) \propto \prod_{j=1}^{10} p_j$$

p_j corresponds to the area in the j^{th} interval. One can take the approximation mean the mean,e.g. \$x_{Flanders,3} = (1500+1800)/2 = 1650 \$? On can approximate that with $f(x_i)\Delta$ where Δ is the unit of measurement.

$$
\begin{aligned}
p_j = P(x_j - \frac{\Delta_j}{2} < x_j < x_j + \frac{\Delta_j}{2}) &\approx f(x_j)\Delta_j \\
&\approx \frac{1}{\Gamma(\phi_k^{-1})}\big(\phi_k\mu\big)^{\phi_k} x_j^{\frac{1}{\phi_k}-1} \exp\big(\frac{-x_j}{\phi_k\mu}\big)\Delta_j
\end{aligned}
$$

This gives for the likelihood:

$$
\begin{aligned}
P(D|\mu,\phi) &\propto \prod_{j=1}^{10} x_j^{\frac{1}{\phi_k}-1} \exp\big(\frac{-x_j}{\phi_k\mu}\big)\Delta_j \\
&\propto \exp\big(\frac{-\sum x_j}{\phi_k\mu}\big) \prod_{j=1}^{10} x_j^{\frac{1}{\phi_k}-1}\Delta_j
\end{aligned}
$$

3 Question 2 : Priors

A Appendix

A.1 Figures

A.2 Code

Note

For reproducibility purposes, the complete R project containing the source code and the results is available on https://github.com/.