IMS Inleveropdracht 2

2023-10-18

Opdracht (later te verwijderen)

We want to check whether the gas extraction from the Groningen gas field had some influence on the number of earthquakes in the Netherlands. The dataset Earthquakes.Rdata contains the number of yearly earthquakeswith a magnitude of 3.0 or larger on the Richter scale. It is based on the recordings of the KMNI: https://dataplatform.knmi.nl/dataset/preview/ aardbevingen-catalogus-1. [In contrast to the KMNI, I counted earth- quakes which happened within 3 days as one.] We want to model the num-ber of yearly earthquakes before the start of the gas extraction (1900-1962) and during the gas extraction (1963-2022) individually. Find statistical models for both time periods. Try to use the same family of distributions1, but allow for different parameters. [In this case we can easily compare the estimated models by comparing the estimated parameters, see Example 3.8] Calculate the maximum likelihood estimator for your model and com- pute the mean squared error of your estimator. Construct also two-sided asymptotic 0.95 confidence intervals for your parameters. For most rea- sonable distributions you can calculate ML-estimator, MSE and CI's the- oretically. If this seems to difficult, you can numerically maximize your log-likelihood (see Example 1.16), simulate the MSE for some values of the parameters by a Monte Carlo simulation (see Example 1.12) and simulate bootstrap confidence intervals (see Example 1.16). What is your conclusion about the influence of the gas extraction?

Introduction

Earthquakes are .....

Organizing the data

We have a data set available to us that ......

```
load("Earthquakes2.Rdata")
sep_data <- stack(Data) #seperates the data from our data file
all_values <- sep_data$values #seperates just the values into a vector
prior_gas_values <- all_values[1:63] #data from the years prior to gas extraction
post_gas_values <- all_values[64:123] #data from the years after gas extraction began
```

The histograms:

```
hist1 <- hist(prior_gas_values, freq = FALSE, right = FALSE, breaks = c(0:8), main = "Histogram of our p
     ,xlab = "amount of earthquakes", xlim = range(c(prior_gas_values,post_gas_values+1)), ylim = c(0,1
hist2 <- hist(post_gas_values, freq = FALSE, right = FALSE, breaks = c(0:8), main = "Histogram of our po
     ,xlab = "amount of earthquakes", xlim = range(cq(prior_gas_values,post_gas_values+1)), ylim = c(0,
```

The point graphs

We want to make a point graph for both data sets with the x-axis being all the possible values (in our case 0-7) and the y-axis being the density of every point.

```
par(mfrow=c(1, 2))
#First we need to know some information, like the maximum amount of earthquakes to plot
```

```r
max_amount = max(prior_gas_values, post_gas_values)
#we generate a table from the prior_gas_values vector
prior_freqs_table <- table(factor(prior_gas_values, levels = 0:max(prior_gas_values)))/length(prior_gas
prior_freqs_table
```

```
##
##          0          1          2          3
## 0.69841270 0.19047619 0.09523810 0.01587302
```

```r
prior_freqs <- stack(prior_freqs_table)$values #this will be our y-axis vector
prior_x <- c(0:(length(prior_freqs)-1)) #this will be our x-axis vector
#we generate a table from the post_gas_values vector
post_freqs_table <- table(factor(post_gas_values, levels = 0:max(post_gas_values)))/length(post_gas_valu
stack(post_freqs_table)
```

```
##       values ind
## 1 0.10000000   0
## 2 0.35000000   1
## 3 0.25000000   2
## 4 0.20000000   3
## 5 0.06666667   4
## 6 0.01666667   5
## 7 0.00000000   6
## 8 0.01666667   7
```

```r
post_freqs <- stack(post_freqs_table)$values #this will be our y-axis vector
post_x <- c(0:(length(post_freqs)-1)) #this will be our x-axis vector
length(prior_freqs)
```

```
## [1] 4
```

```r
length(prior_x)
```

```
## [1] 4
```

```r
length(post_freqs)
```

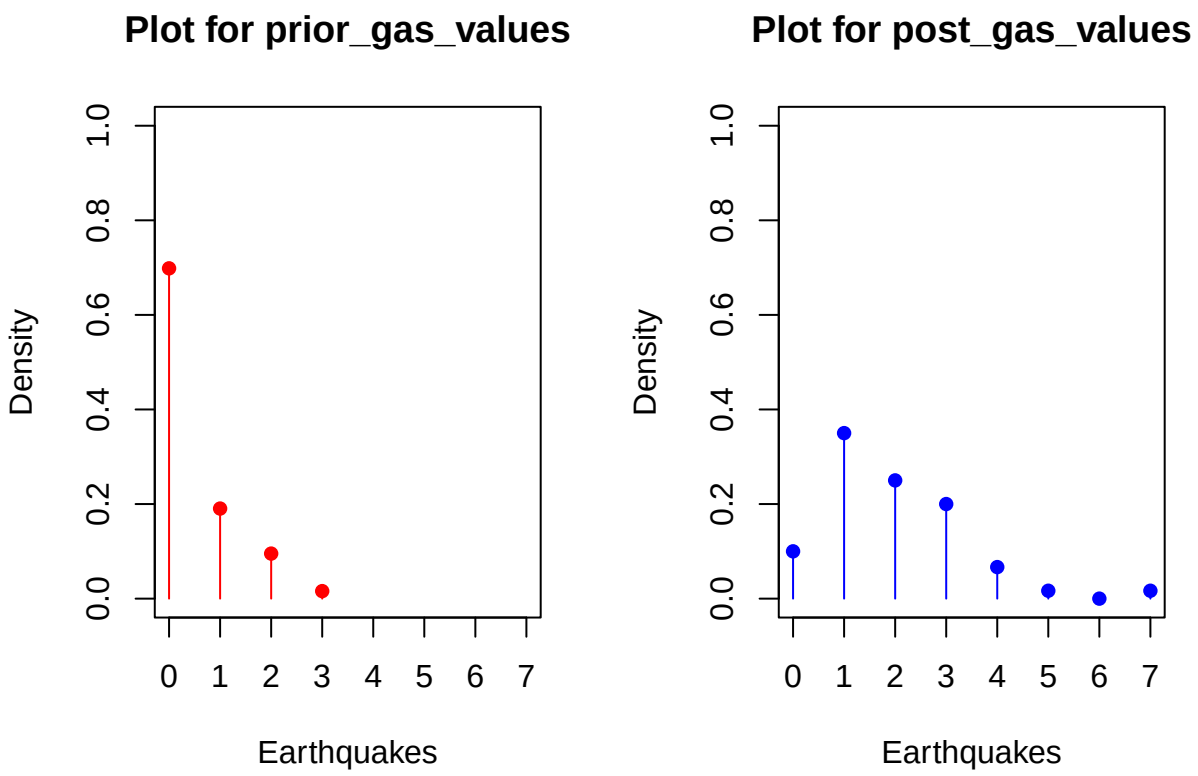
```
## [1] 8
```

```r
length(post_x)
```

```
## [1] 8
```

```r
# Create a blank plot
plot1 <- plot(x = 1,
     type = "n",
     xlim = c(0, max_amount),
     ylim = c(0, 1),
     pch = 16,
     xlab = "Earthquakes",
     ylab = "Density",
     main = "Plot for prior_gas_values")
points1 <- points(prior_x, prior_freqs, pch = 16, lwd = 1, col = "red")
segments1 <- segments(x0 = prior_x, y0 = 0, x1 = prior_x, y1 = prior_freqs, col = "red")
plot2 <- plot(x = 1,
     type = "n",
     xlim = c(0, max_amount),
     ylim = c(0, 1),
```

```
     pch = 16,
     xlab = "Earthquakes",
     ylab = "Density",
     main = "Plot for post_gas_values")
points2 <- points(post_x, post_freqs, pch = 16, lwd = 1, col = "blue")
segments2 <- segments(x0 = post_x, y0 = 0, x1 = post_x, y1 = post_freqs, col = "blue")
```



Calculating the mean:

```
mean_prior <- sum(prior_gas_values)/63
mean_post <- sum(post_gas_values)/63
print(mean_prior)
```

```
## [1] 0.4285714
```

```
print(mean_post)
```

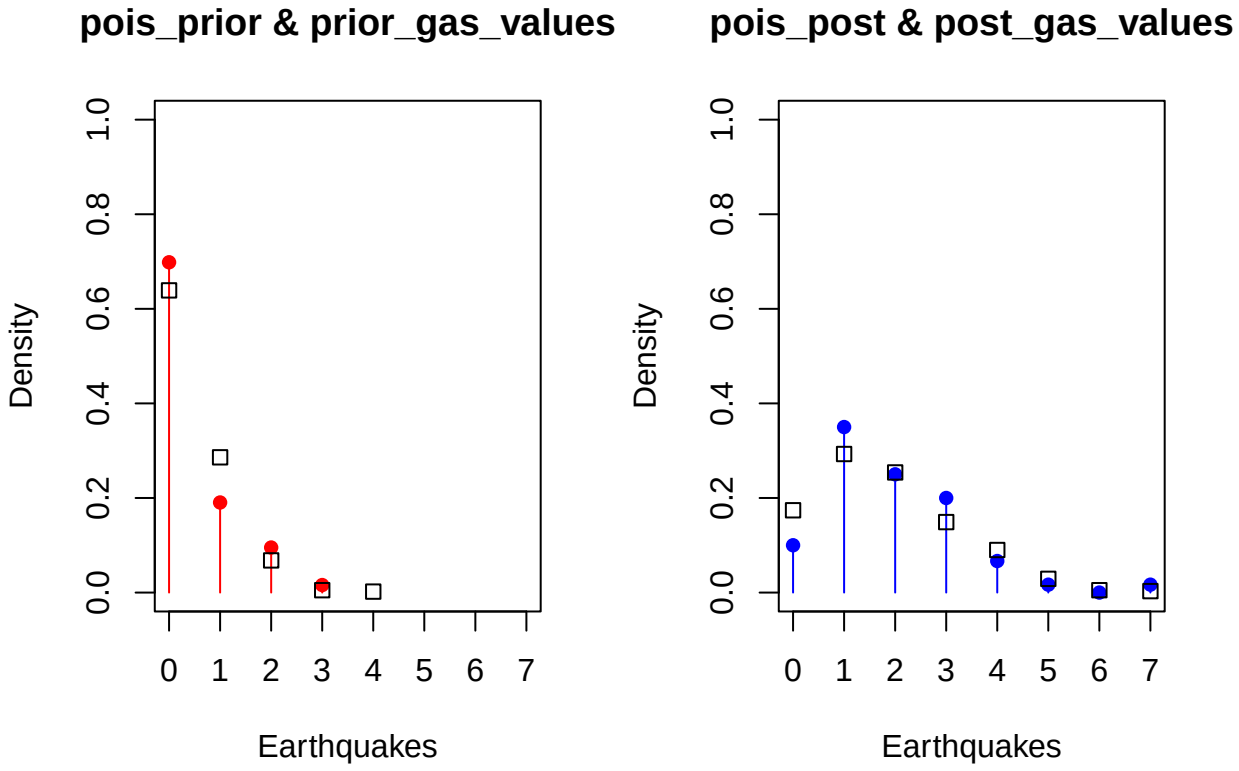
```
## [1] 1.825397
```

```
par(mfrow=c(1, 2))
?points
#we plot the Poisson distribution over the previous plots.
pois_prior_values <- rpois(1000, 0.4285714)
pois_prior_freqs_table <- table(factor(pois_prior_values, levels = 0:max(pois_prior_values)))/length(poi
pois_prior_freqs <- stack(pois_prior_freqs_table)$values
pois_prior_x <- c(0:(length(pois_prior_freqs)-1))

pois_post_values <- rpois(1000, 1.825397)
pois_post_freqs_table <- table(factor(pois_post_values, levels = 0:max(pois_post_values)))/length(pois_
pois_post_freqs <- stack(pois_post_freqs_table)$values
pois_post_x <- c(0:(length(pois_post_freqs)-1))
```

```r
plot3 <- plot(x = 1,
     type = "n",
     xlim = c(0, max_amount),
     ylim = c(0, 1),
     pch = 16,
     xlab = "Earthquakes",
     ylab = "Density",
     main = "pois_prior & prior_gas_values")
points1 <- points(prior_x, prior_freqs, pch = 16, lwd = 1, col = "red")
segments1 <- segments(x0 = prior_x, y0 = 0, x1 = prior_x, y1 = prior_freqs, col = "red")
points3 <- points(pois_prior_x, pois_prior_freqs, pch = 0, lwd = 1, col = "black")

plot2 <- plot(x = 1,
     type = "n",
     xlim = c(0, max_amount),
     ylim = c(0, 1),
     pch = 16,
     xlab = "Earthquakes",
     ylab = "Density",
     main = "pois_post & post_gas_values")
points2 <- points(post_x, post_freqs, pch = 16, lwd = 1, col = "blue")
segments2 <- segments(x0 = post_x, y0 = 0, x1 = post_x, y1 = post_freqs, col = "blue")
points4 <- points(pois_post_x, pois_post_freqs, pch = 0, lwd = 1, col = "black")
```



some text after