Question4

Pawan Kumar

2022-11-16

Problem 4

Part -A: We want to predict the Claims as function of Holders. So we want to fit the following models:

Claims $i = \beta^0 + \beta^1\ Holders_i + \varepsilon_i$, $i = 1, 2, \cdots, n$ Assume : ε_i N(0,σ^2). Note that β^0, β^1 R and $\sigma \in R^+$.

The above model can alse be re-expressed as, $Claims_i \sim N(_i, \sigma^2)$, where $_i = \beta^0 + \beta^1 Holders_i + \varepsilon_i$, $i = 1, 2, \cdots, n$

(i) Clearly write down the negative-log-likelihood function in R. Then use optim function to estimate MLE of $\theta = (\beta^0, \beta^1, \sigma)$
(ii) Calculate Bayesian Information Criterion (BIC) for the model.

```
library(SciViews)
```

```
## Warning: package 'SciViews' was built under R version 4.0.5
```

```
library(MASS)
H=Insurance$Holders
C=Insurance$Claims

k=length(H)
Loglikelihood <- function(pars,y)
{
  lf= 0
  for(i in 1:k)
  {
    lf=lf+dnorm(y[,1][i],mean = pars[1]+(pars[2]*y[,2][i]),sd=exp(pars[3]),log = TRUE)
  }
  return(-lf)
}

d = optim(c(0,1,50),Loglikelihood,y=data.frame(cbind(C,H)))
d
```

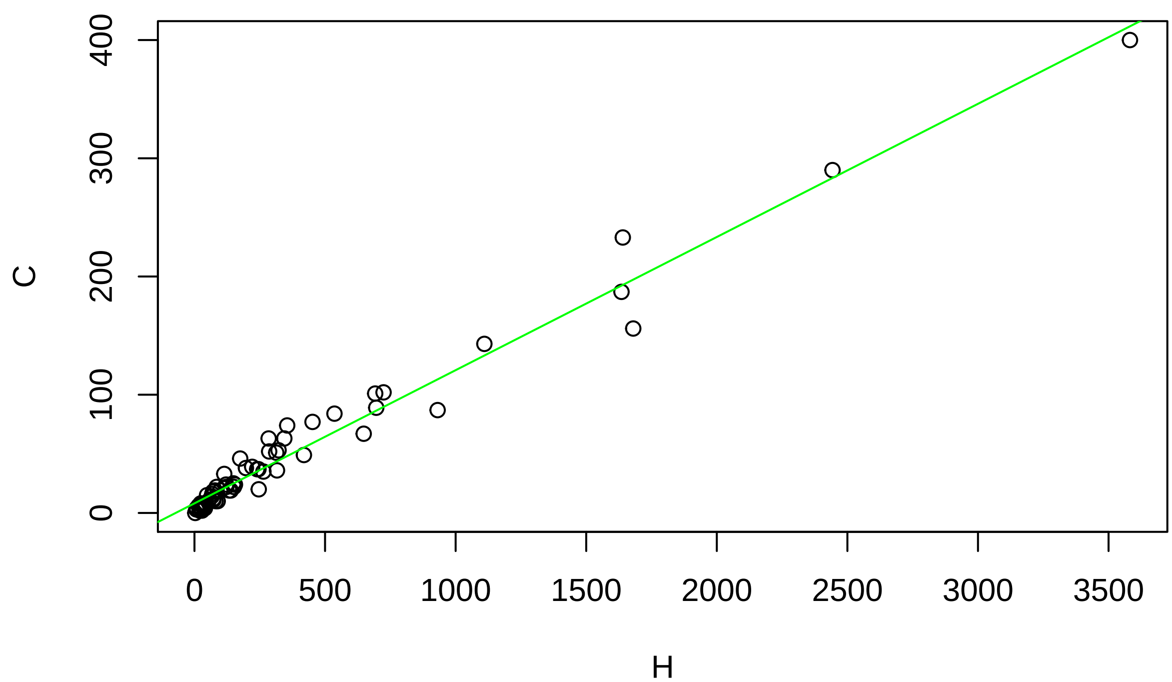
```
## $par
## [1] 8.1197142 0.1126451 2.4738755
##
## $value
```

```
## [1] 249.141
##
## $counts
## function gradient
##      280       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

```
BIC=ln(k)*(length(d$par))+2*d$value
BIC
```

```
## [1] 510.7587
```

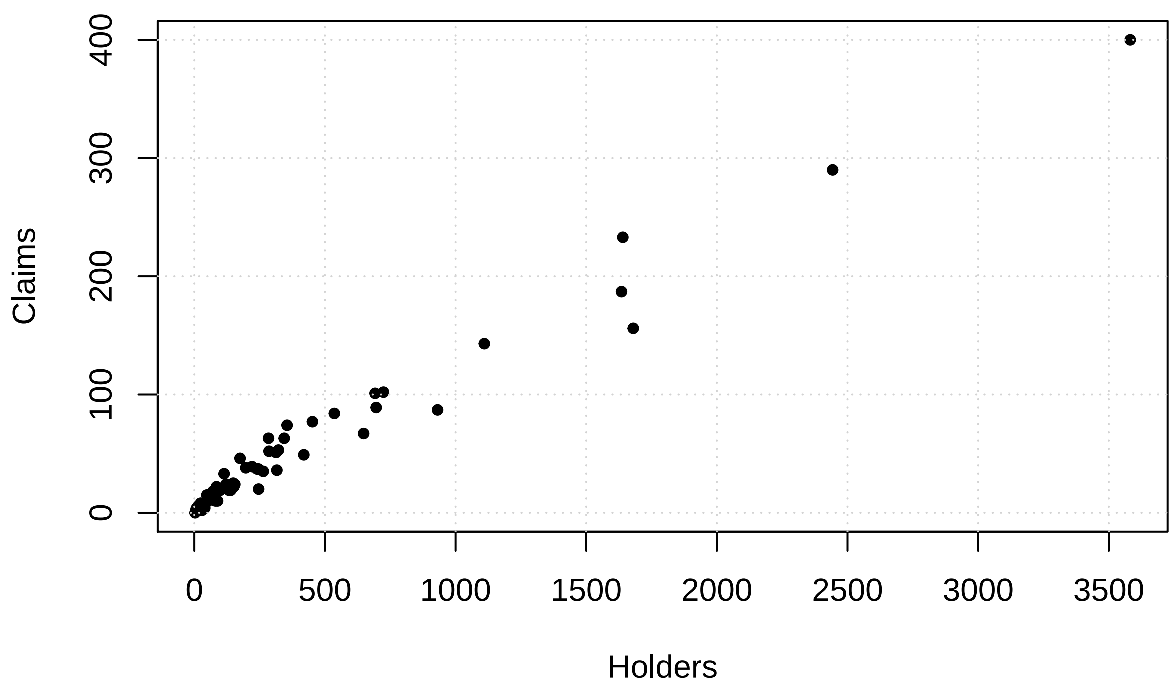
```
plot(C~H)
abline(d$par[1],d$par[2],col='green')
```



Part -B: Now we want to fit the same model with change in distribution:

Claims $i = \beta^0 + \beta^1 Holders_i + \varepsilon^i$, $i = 1, 2, \cdots, n$ Assume : $\varepsilon_i \sim Laplace(0, \sigma^2)$. Note that $\beta^0, \beta^1 \in R$ and $\sigma \in R^+$. (i) Clearly write down the negative-log-likelihood function in R. Then use optim function to estimate MLE of $= (0 , 1 ,)$ (ii) Calculate Bayesian Information Criterion (BIC) for the model.

```
rm(list=ls())
library(MASS)
plot(Insurance$Holders,Insurance$Claims
      ,xlab =
        'Holders'
      ,ylab=
        'Claims'
      ,pch=
        20
)
grid()
```




```
attach(Insurance)
data=Insurance
data=data.frame(cbind(Claims,Holders))



library(SciViews)
library(jmuOutlier)
```

```
## Warning: package 'jmuOutlier' was built under R version 4.0.5
```

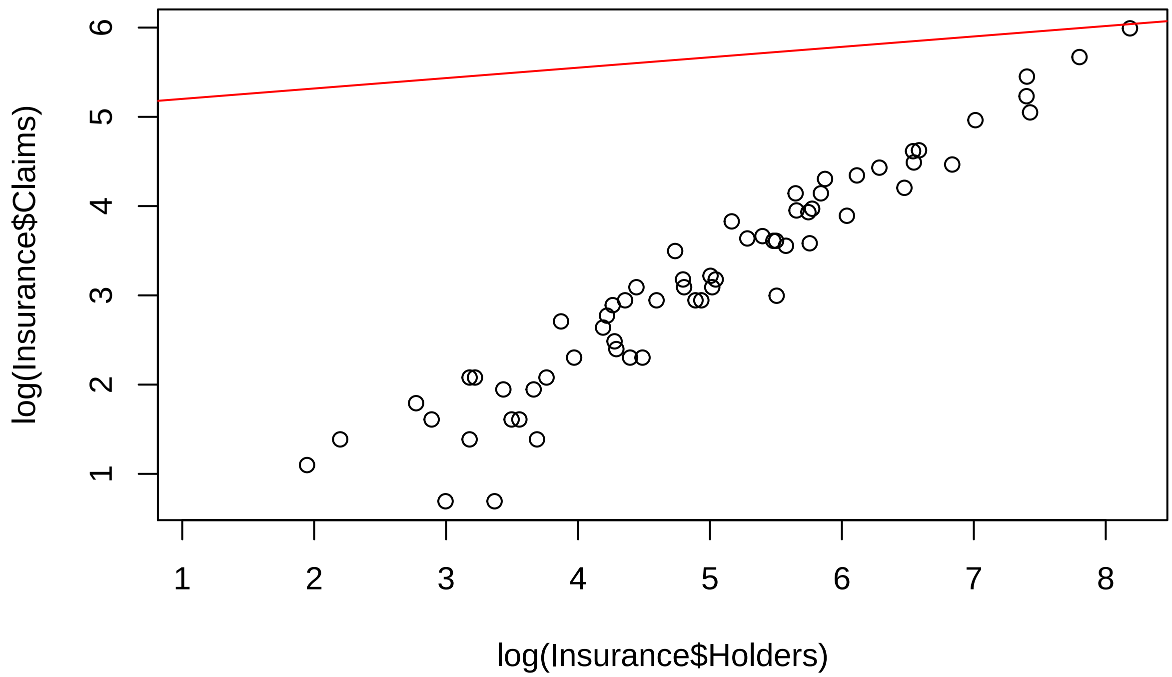
```r
library(MASS)
H=Insurance$Holders
C=Insurance$Claims

Loglaplace <- function(theta,data){
  lf=0
  n=nrow(data)
  for(i in 1:n){
    lf=lf+log(dlaplace(data[,1][i]-theta[1]-theta[2]*data[,2][i],0,(theta[3])))

  }
  return(-lf)
}

b2=optim(c(9,0.1,50),Loglaplace,data=data)
b2
```

```
## $par
## [1]  5.0843690  0.1166253 11.6118219
##
## $value
## [1] 243.1051
##
## $counts
## function gradient
##      158       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

```r
n=nrow(data)
BIC2=log(n)*(3)-2*(-1*b2$value)
BIC2
```

```
## [1] 498.6869
```

```r
plot(log(Insurance$Claims)~log(Insurance$Holders))
abline(b2$par[1],b2$par[2],col='red')
```



Part - C: We want to fit the following models

$Claims_i \sim LogNormal({}_i, \sigma^2)$,where ${}_i = \beta^0 + \beta^1 log(Holders_i)$, $i = 1, 2, ..., n$ Note that $\beta^0, \beta^1 \in R$ and $\sigma \in R^+$.

```
library(SciViews)
library(jmuOutlier)
library(MASS)


data=data.frame(cbind(Insurance$Claims,Insurance$Holders))
data = data[-61,]
n=length(Insurance$Holders)-1

LogNormal=function(pars,data)
{
  lf=0
  for(i in 1:n)
  {
    lf=lf+log(dlnorm(data[,1][i] , pars[1]+pars[2]*log(data[,2][i]), pars[3]))
  }
  return(-lf)
}
```

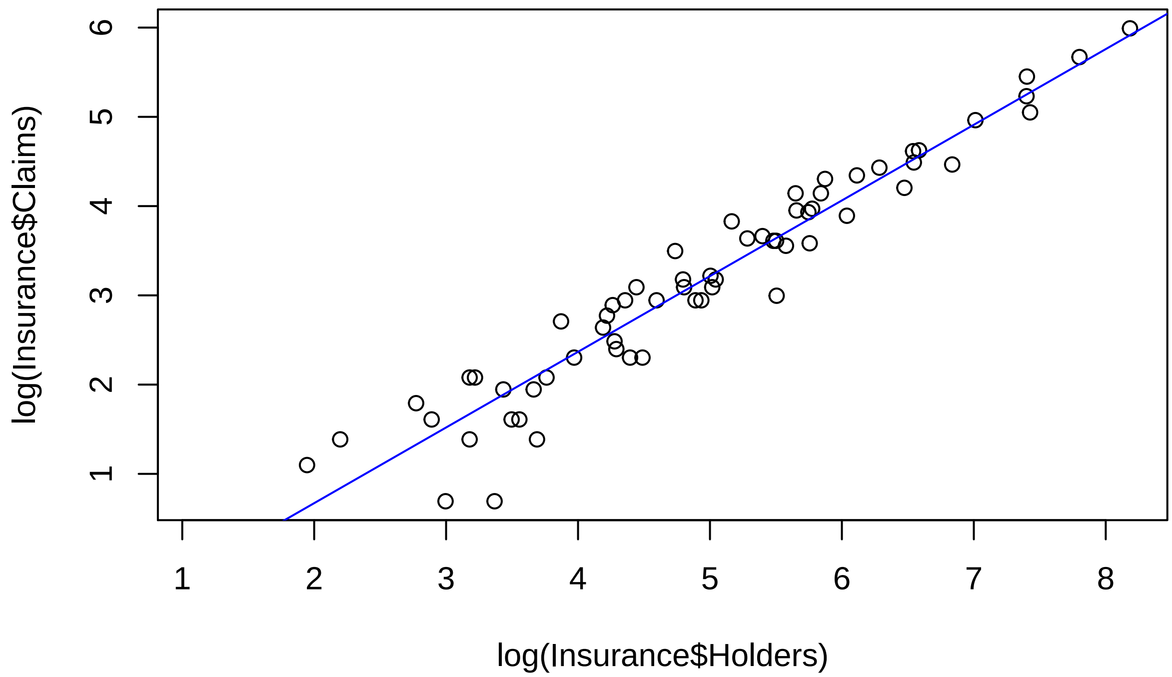
```r
d3=optim(c(1,0,1),LogNormal,data=data)
d3
```

```
## $par
## [1] -1.0241463  0.8478830  0.3293378
##
## $value
## [1] 220.087
##
## $counts
## function gradient
##      196       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

```r
BIC3=ln(n)*(length(d3$par))-2*ln(d3$value)
BIC3
```

```
## [1] 1.641358
```

```r
plot(log(Insurance$Claims)~log(Insurance$Holders))
abline(d3$par[1],d3$par[2],col= 'blue')
```



Part - D: We want to fit the following models:

$Claims_i \sim Gamma(\alpha_i, \sigma)$,where $log(\alpha_i) = \beta_0 + \beta_1 log(Holders_i)$, $i = 1, 2, ..., n$ (iii) Compare the BIC of all three models

```
library(SciViews)
library(jmuOutlier)
library(MASS)

data=data.frame(cbind(Insurance$Claims,Insurance$Holders))
data = data[-61,]
n=length(Holders)-1

LogGamma=function(pars,data)
{
  lf=0
  for(i in 1:n)
  {
    lf=lf+log(dgamma(data[,1][i] , exp(pars[1]+pars[2]*log(data[,2][i])), scale = pars[3]))

  }
  return(-lf)
}
```

```r
d4=optim(c(-3,0.11,1),LogGamma,data=data)
d4
```

```
## $par
## [1] -1.6412901  0.8369852  2.0564055
##
## $value
## [1] 212.4544
##
## $counts
## function gradient
##      216       NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

```r
BIC4=ln(n)*(length(d4$par))-2*ln(d4$value)
BIC4
```

```
## [1] 1.71195
```

```r
plot(log(Insurance$Claims)~log(Insurance$Holders))
abline(d4$par[1],d4$par[2],col="violet")
```