Project Regression Methods - 2022

Céline Guex - sciper n°310673, Candice Baud - sciper n°359523

06 décembre 2022

1 Introduction - Summary

This report describes the analysis of three databases on events in French motorway tunnels.
The data shows ... to do after our analysis.

2 Initial data analysis

2.1 Description of Fires dataset

This data set lists fires that occurred in French tunnels. The data set gathers 92 tunnels ran by 25 companies. Many parameters are also registered, like traffic, proportion of trucks among tunnel's users (in %), length of the tunnel, maximum speed allowed,slope inclination in the tunnel (in %), shape of the tunnel (slope type in the tunnel) and situation of the tunnel (if it lies in urban area or not). Some tunnels are unidirectional and some are bidirectional. We will use this data set to determine the factors that are the most likely to yields to fires in tunnel. We will also look for factors that might reduce the risk of fires.

2.2 Description of Accidents dataset

The accidents dataset gathers contains 92 tunnels ran by 25 companies. The variables from the dataset Fires are also present in the Accidents one. But the latter also includes the Year in which the accidents occurred, the width of lanes in the tunnel in metres (variable Width) and the number of lanes in the tunnel (variable Lanes).

3 Model fitting

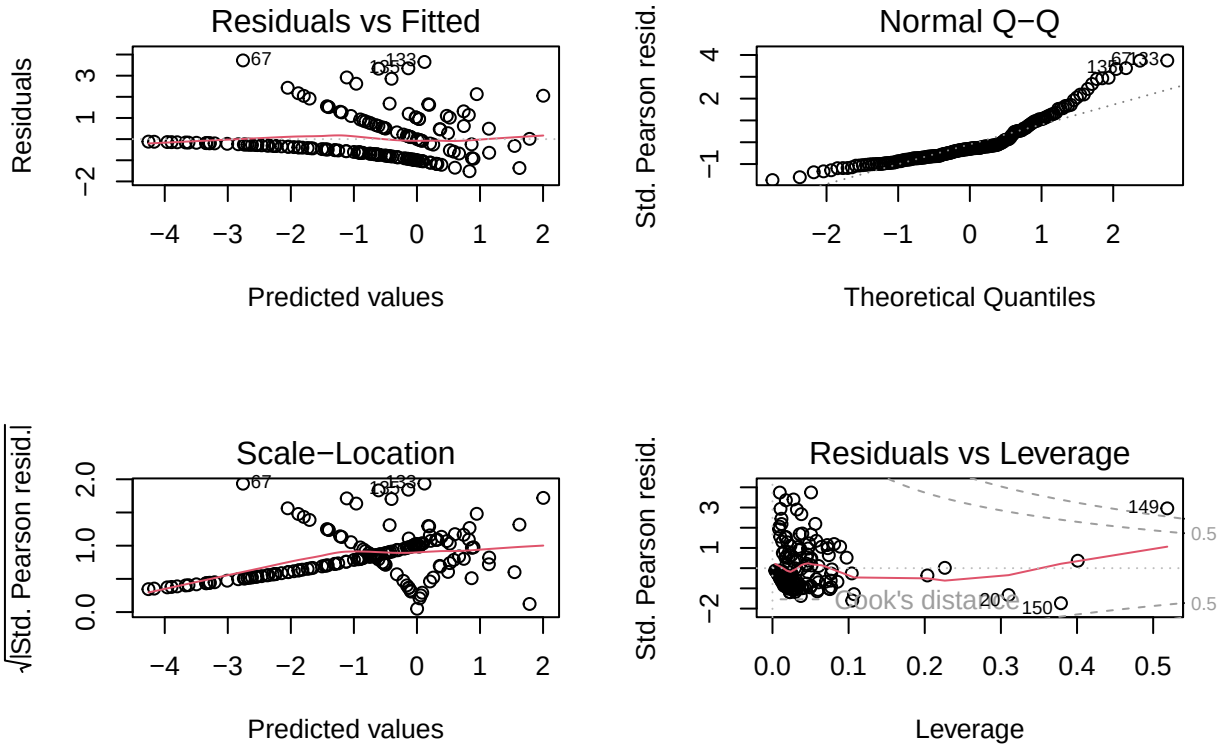
Explanation of technique to justify our choice of model : The principle is quite simple : we use parametric bootstrap to test if the law F of a random sample comes from a certain parametric family $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$, namely we test $H_0 : F \in \mathcal{F}$ versus $H_1 : F \notin \mathcal{F}$. We proceed by measuring $T = \sup_x |\hat{F}_N(x) - F_{\hat{\lambda}}(x)|$ and we hope to have a small T under H_0. Details of the algorithm can be found under this link : https://htmlpreview.github.io/?https://raw.githubusercontent.com/TMasak/StatComp/master/Notes/10_Bootstrap.html TO DO : change place of link into references

3.1 Fires

We try to fit to the data a Poisson model.

Poisson model

```
##
## Call:
## glm(formula = Fires ~ log(Traffic) + HGV + log(Length), family = "poisson",
##     data = Fires)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.7271  -1.0384  -0.5034   0.4395   2.8400
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -20.74042    2.13772  -9.702  < 2e-16 ***
## log(Traffic)   0.81806    0.09049   9.040  < 2e-16 ***
## HGV            5.49984    1.00758   5.458  4.8e-08 ***
## log(Length)    0.64439    0.12372   5.208  1.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 322.16  on 169  degrees of freedom
## Residual deviance: 191.59  on 166  degrees of freedom
## AIC: 368.67
##
## Number of Fisher Scoring iterations: 5
```



Poisson mixed model With tunnel as the random effect

```
## Warning: le package 'lme4' a été compilé avec la version R 4.2.2

## Le chargement a nécessité le package : Matrix

## Warning: le package 'Matrix' a été compilé avec la version R 4.2.2

## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: Fires ~ log(Traffic) + HGV + log(Length) + (1 | Tunnel)
##    Data: Fires
##
##      AIC      BIC   logLik deviance df.resid
##    362.6    378.3   -176.3    352.6      165
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8249 -0.6091 -0.3277  0.3921  2.5358
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  Tunnel (Intercept) 0.322    0.5675
## Number of obs: 170, groups:  Tunnel, 92
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -21.6162     3.0211  -7.155 8.36e-13 ***
## log(Traffic)   0.8663     0.1304   6.641 3.11e-11 ***
## HGV            5.2058     1.4834   3.509 0.000449 ***
## log(Length)    0.6222     0.1603   3.882 0.000103 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) lg(Tr) HGV
## log(Traffc) -0.941
## HGV         -0.238  0.387
## log(Length) -0.568  0.261 -0.352
```

Then with company

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: Fires ~ log(Traffic) + HGV + log(Length) + (1 | Company)
##    Data: Fires
##
##      AIC      BIC   logLik deviance df.resid
##    367.1    382.8   -178.6    357.1      165
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.8326 -0.6663 -0.3322  0.3232  3.6168
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  Company (Intercept) 0.09815  0.3133
## Number of obs: 170, groups:  Company, 25
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -21.9744     2.6301  -8.355  < 2e-16 ***
## log(Traffic)   0.8491     0.1116   7.610 2.75e-14 ***
## HGV            5.1012     1.2166   4.193 2.75e-05 ***
## log(Length)    0.7334     0.1485   4.940 7.80e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) lg(Tr) HGV
## log(Traffc) -0.926
## HGV         -0.186  0.400
## log(Length) -0.589  0.248 -0.463
```

3.2 Accidents

4 Discussion

5 References