Série Temporelle

Master 1 MAS Rennes - Série Temporelle

Dylan

15 mai 2020

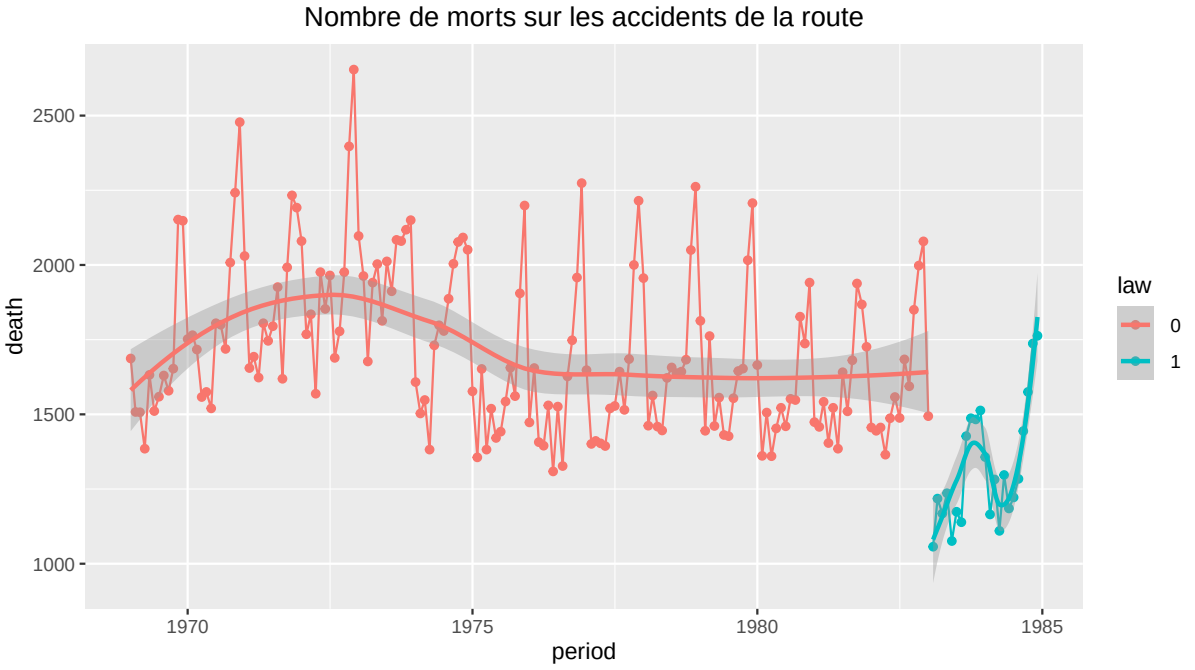
Table des matières

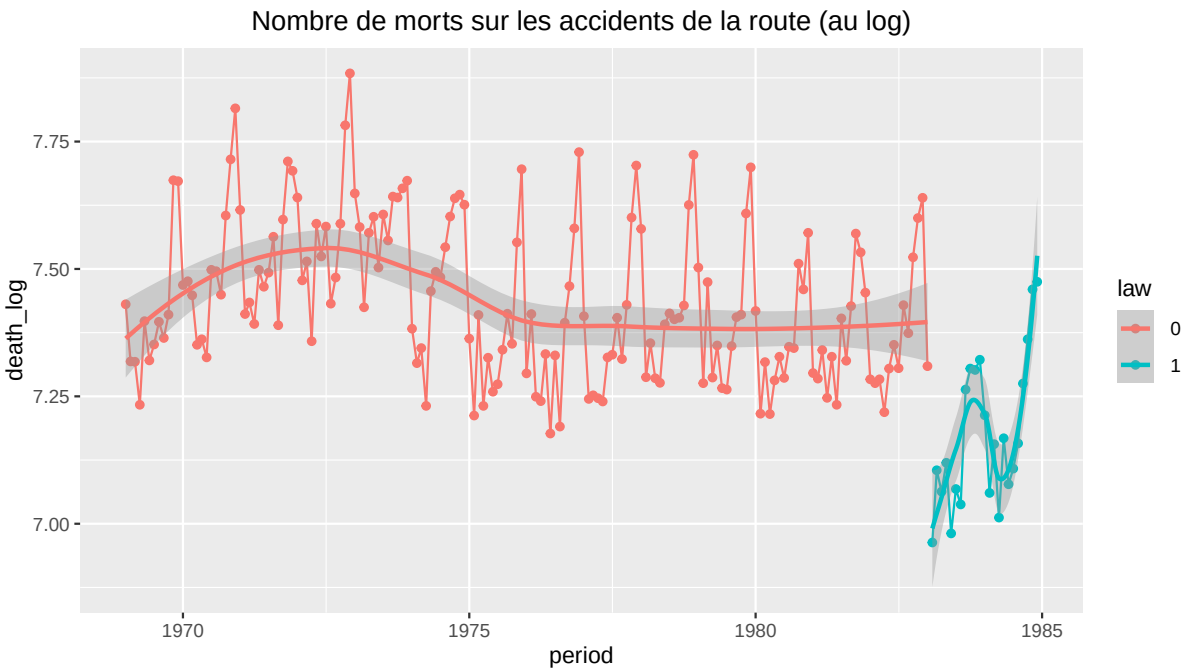
1 Import des données

```
period <-
  seq(as.Date('1969-01-01'), as.Date('1984-12-31'), by = "month")

ukdeath <-
  read_delim("../data.txt", delim = "  ", col_types = "if") %>%
  mutate(death_log = log(death),
         period = period)
```

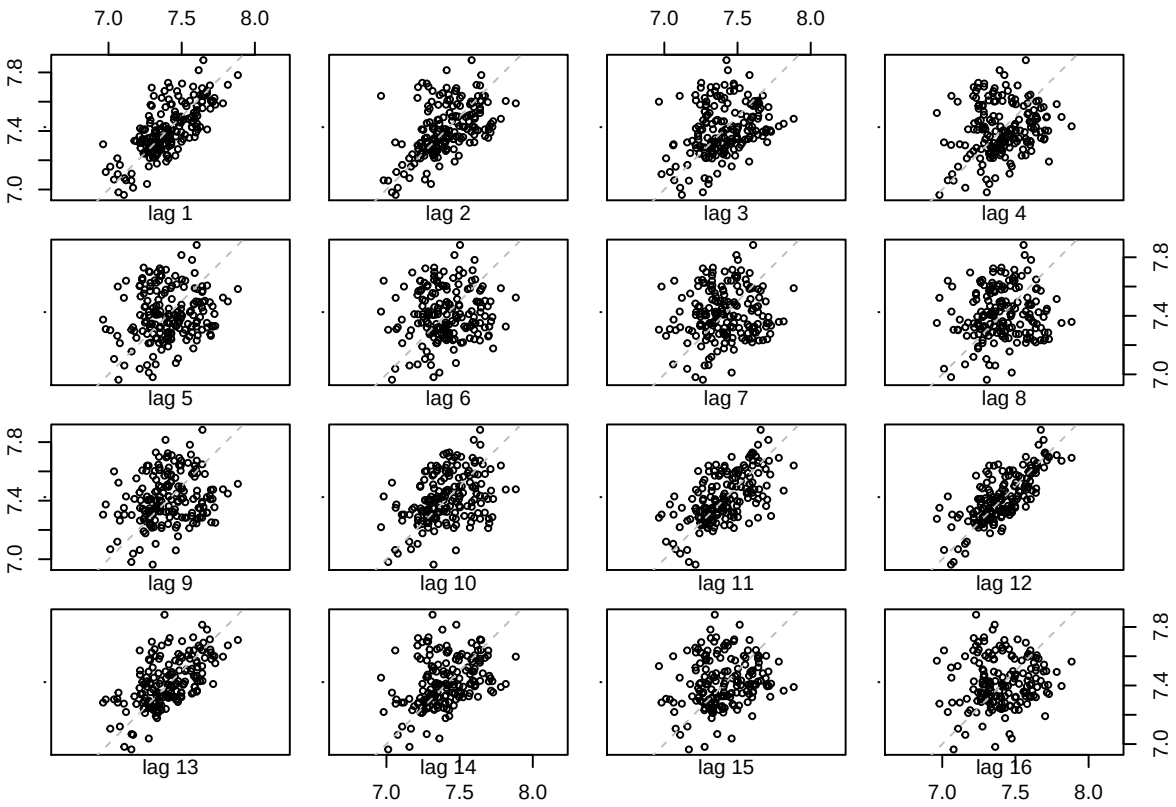
```
##      death       law        death_log          period
##  Min.   :1057    0:169   Min.   :6.963   Min.   :1969-01-01
##  1st Qu.:1462    1: 23   1st Qu.:7.287   1st Qu.:1972-12-24
##  Median :1631            Median :7.397   Median :1976-12-16
##  Mean   :1670            Mean   :7.406   Mean   :1976-12-15
##  3rd Qu.:1851            3rd Qu.:7.523   3rd Qu.:1980-12-08
##  Max.   :2654            Max.   :7.884   Max.   :1984-12-01
```



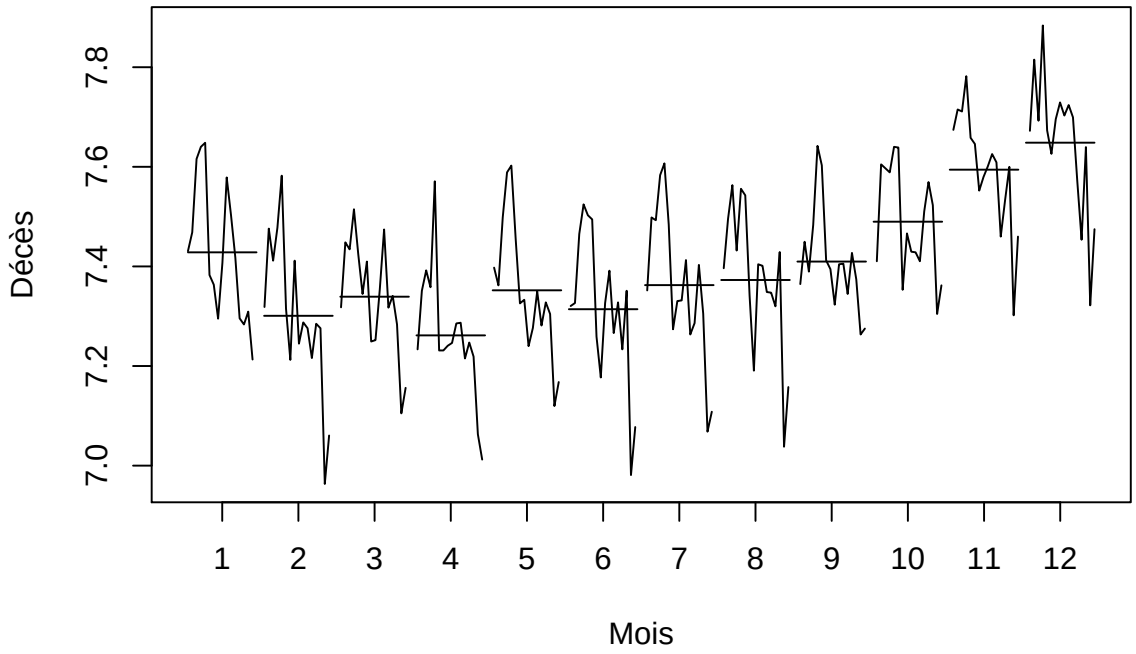


2 Lagplots

```
ukdeath$death_log %>% lag.plot(., lags = 16)
```



```
ukdeath$death_log %>% monthplot(., ylab = "Décès", xlab = "Mois")
```



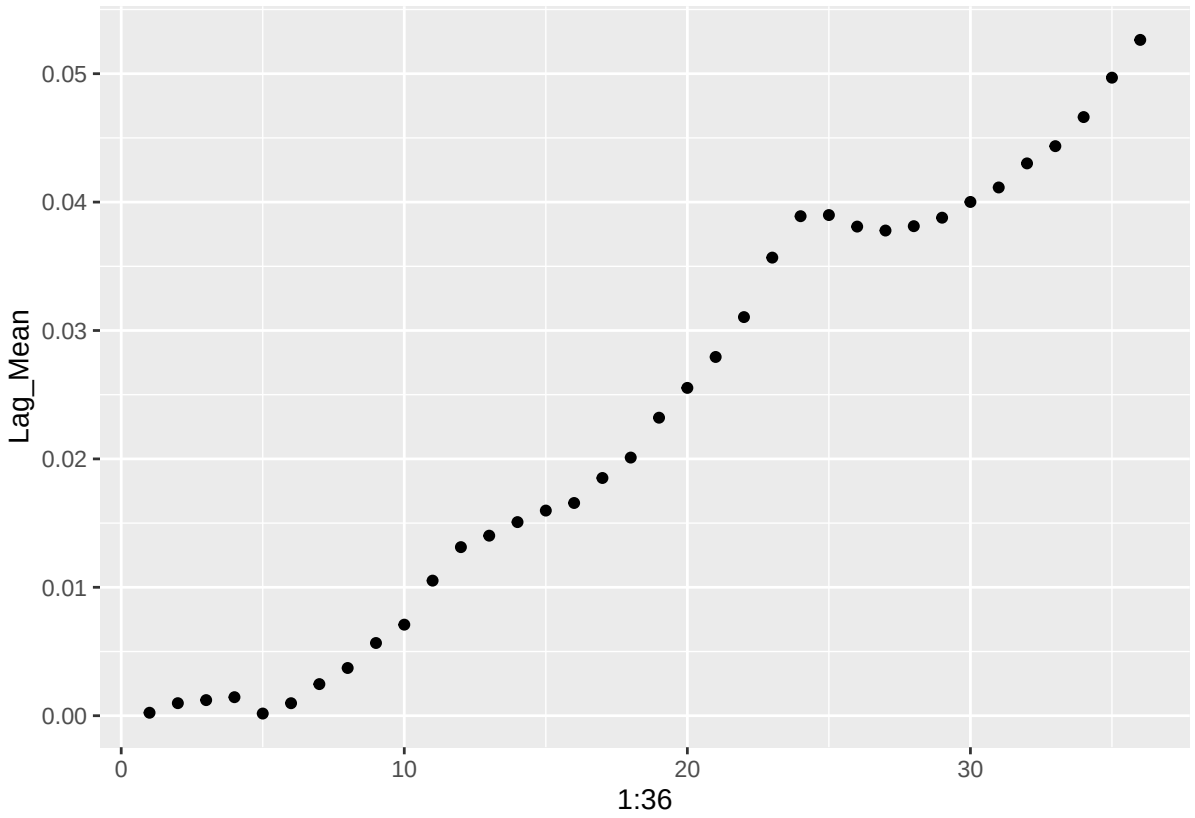
3 Trouver la période et le degré

3.1 La période (lag)

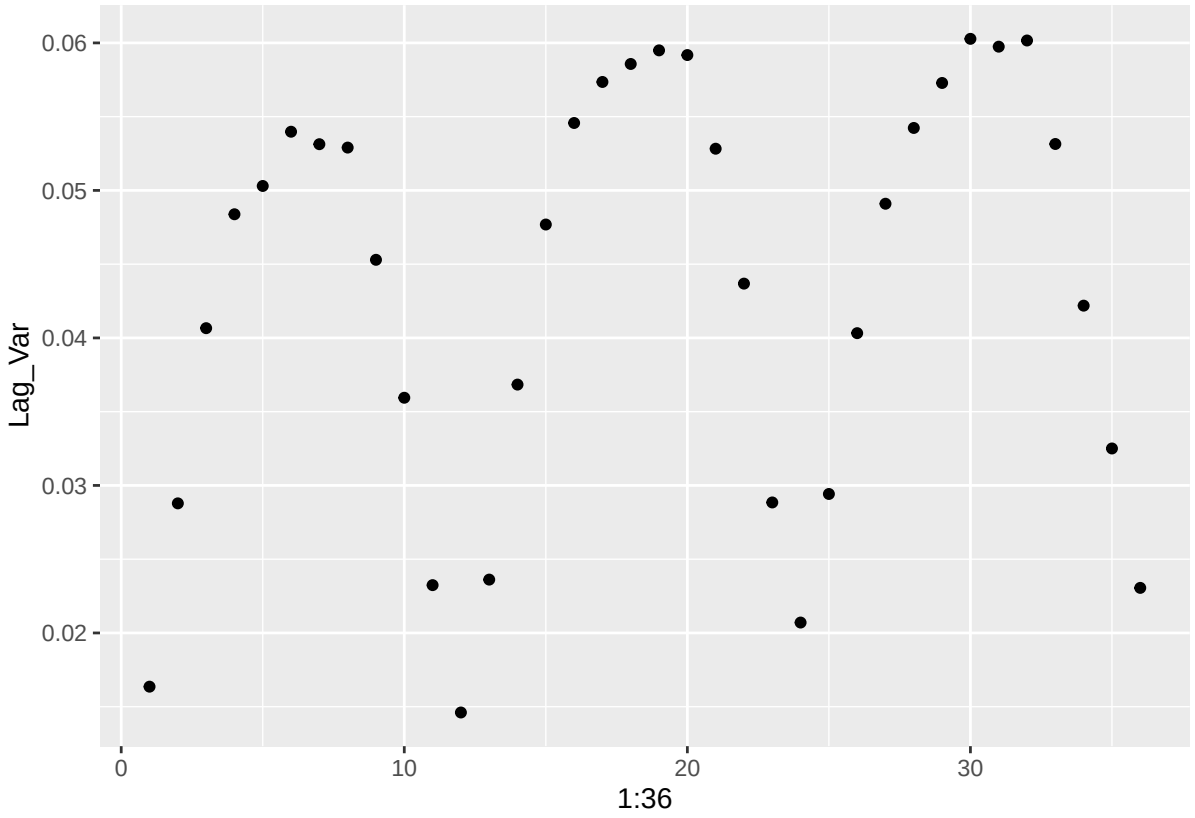
```
Lag_Mean <- NULL
Lag_Var <- NULL

for (ind in 1:36) {
  diff <- diff(ukdeath$death_log, ind, 1)
  Lag_Mean[ind] <- abs(mean(diff))
  Lag_Var[ind] <- var(diff)
}

ggplot() + aes(y = Lag_Mean, x = 1:36) + geom_point()
```




```
ggplot() + aes(y = Lag_Var, x = 1:36) + geom_point()
```

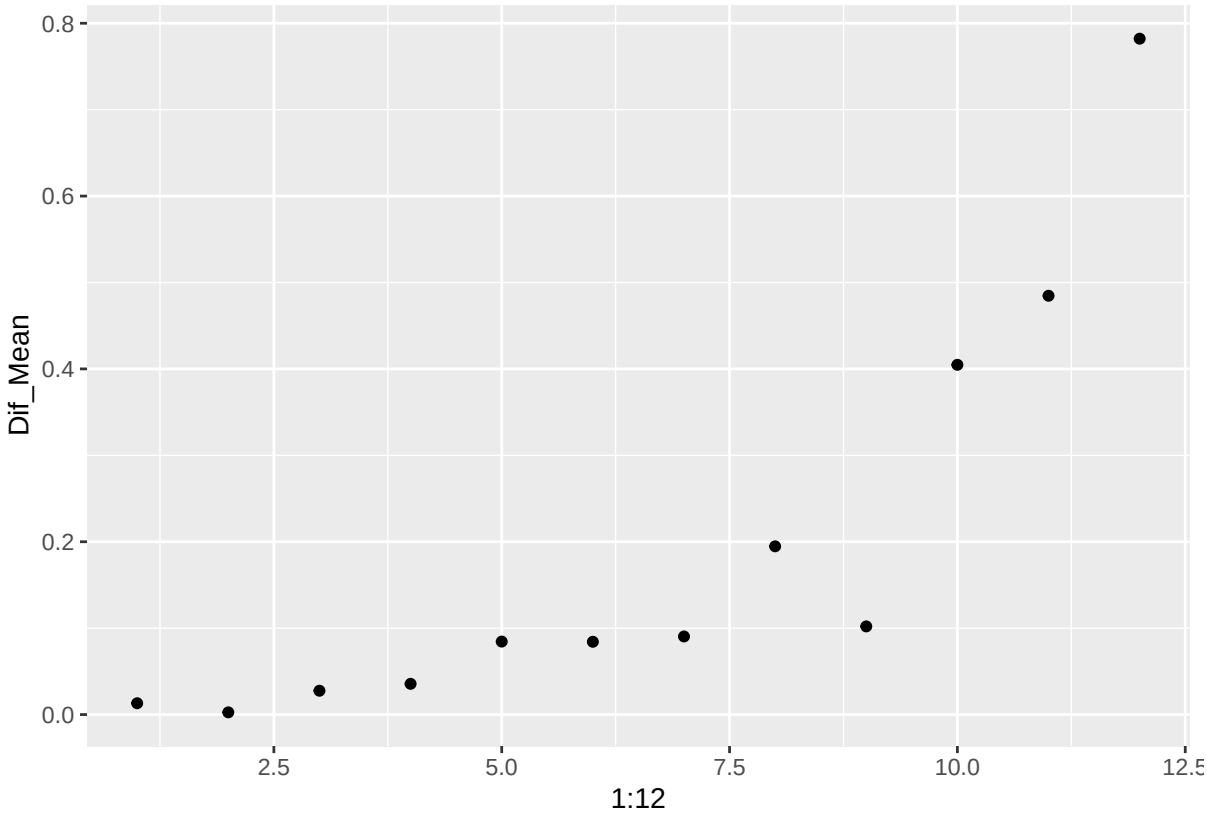


3.2 Le degré (differencies)

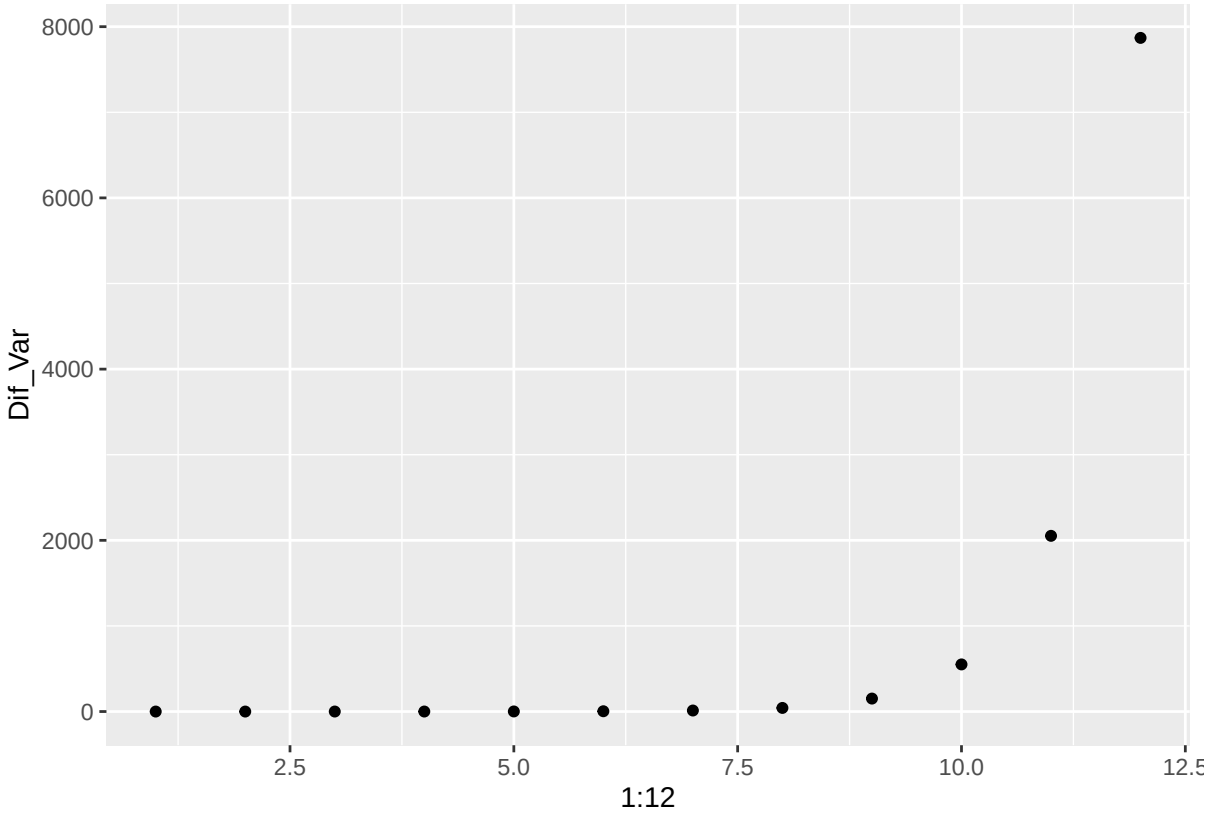
```
Dif_Mean <- NULL
Dif_Var <- NULL

for (ind in 1:12) {
  diff <- diff(ukdeath$death_log, 12, ind)
  Dif_Mean[ind] <- abs(mean(diff))
  Dif_Var[ind] <- var(diff)
}

ggplot() + aes(y = Dif_Mean, x = 1:12) + geom_point()
```



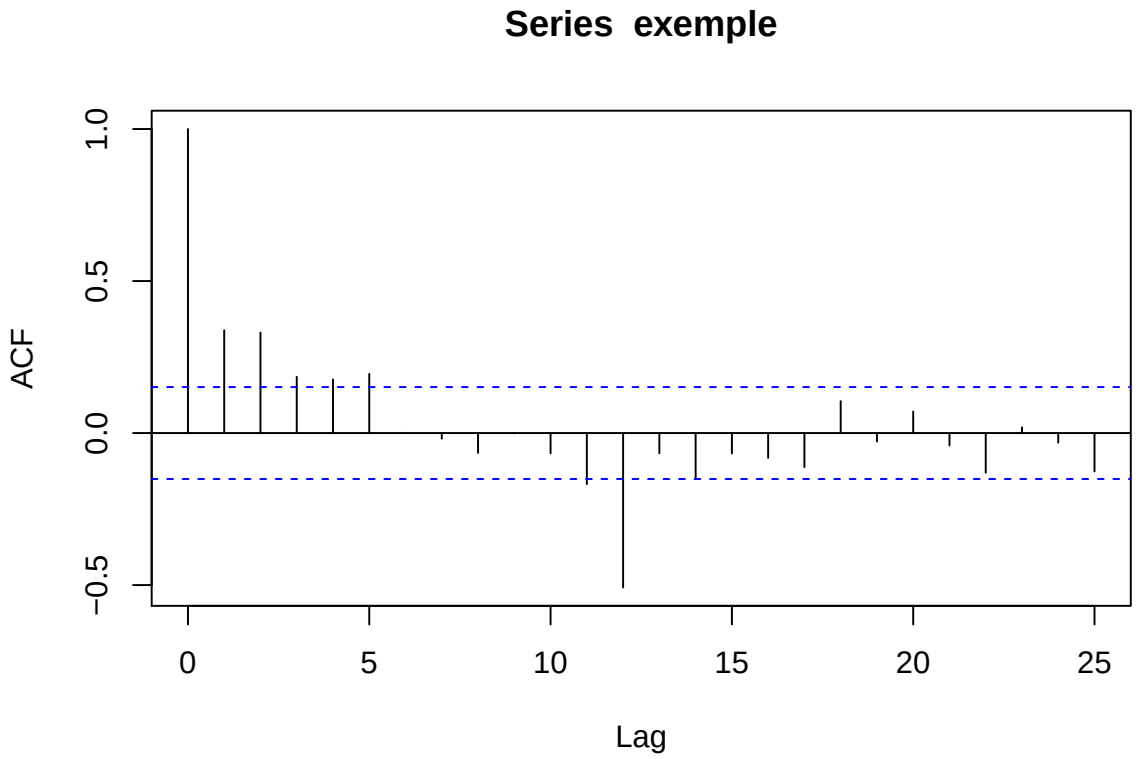

```
ggplot() + aes(y = Dif_Var, x = 1:12) + geom_point()
```



4 A choisir

lag : 12 diff : 1 ou 2

```
exemple <- diff(ukdeath$death_log, 12 , 2)
acf(exemple, 25)
```



5 Reprod

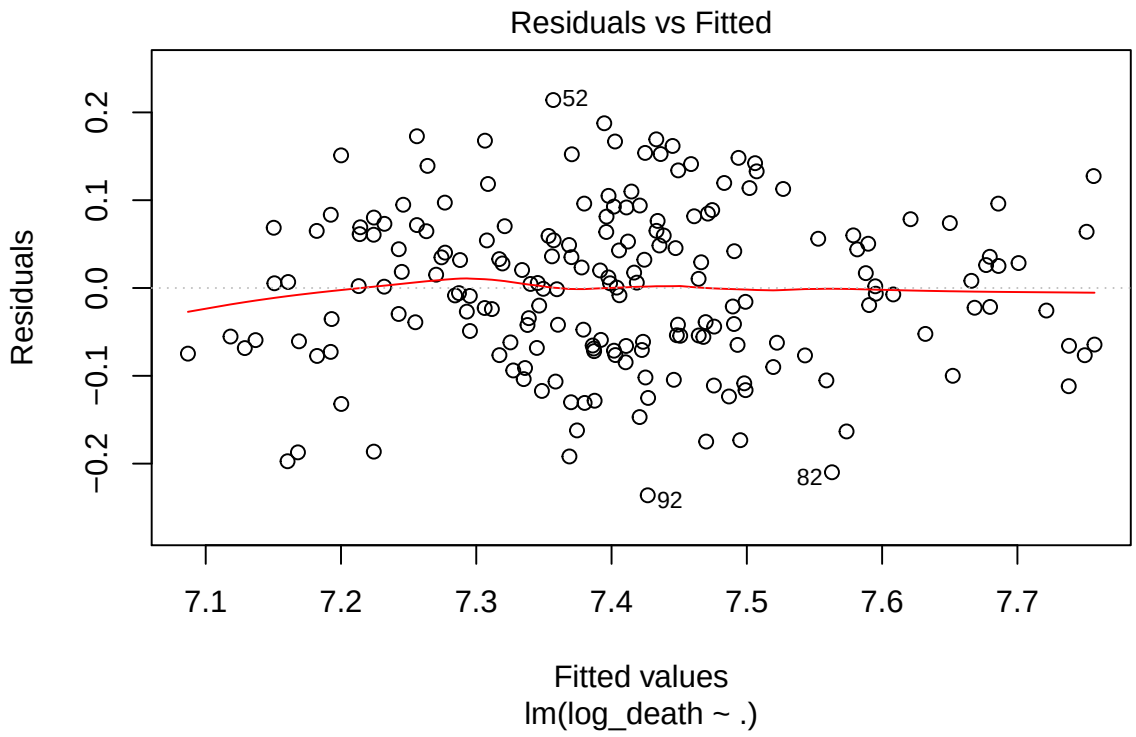
```
t <- 1:192
sinusoides <- t %o% c(rep(1:5, 2)) * pi / 6
sinusoides[, 1:5] <- sin(sinusoides[, 1:5])
sinusoides[, 6:10] <- cos(sinusoides[, 6:10])
sinusoides <- as.data.frame(sinusoides)
names(sinusoides) <-
  c(paste("sin_", 1:5, sep = ""), paste("cos_", 1:5, sep = ""))
```

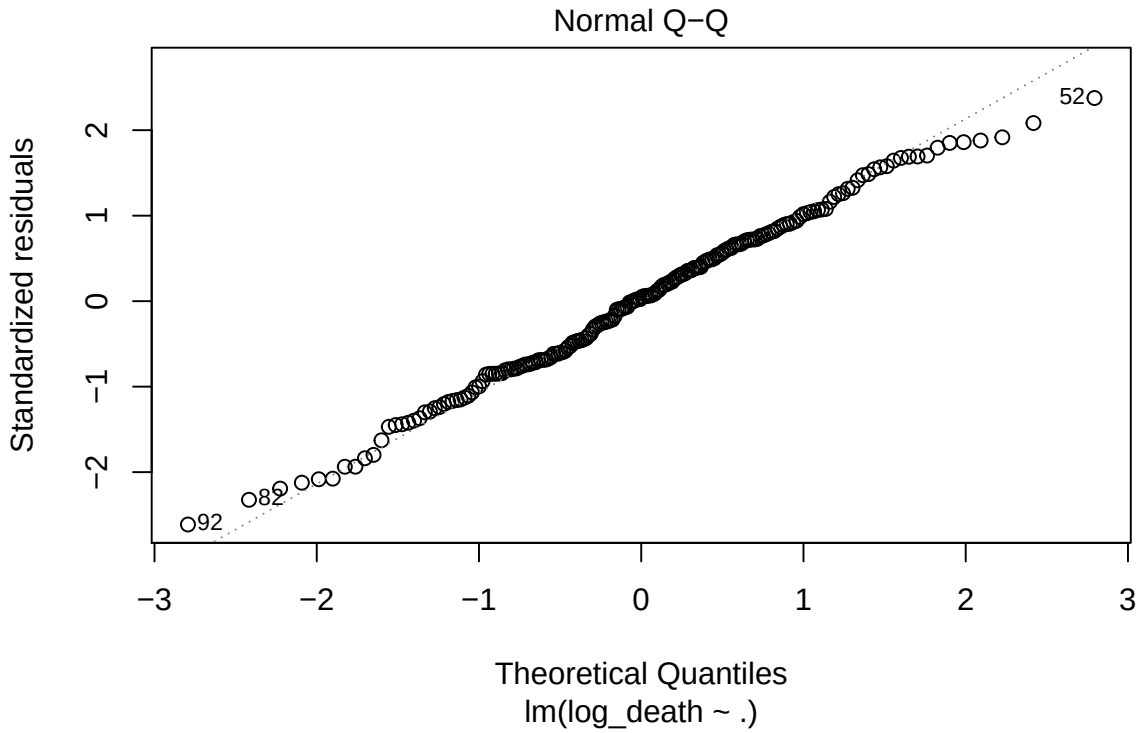
```
log_death <- ukdeath$death_log
df <- data.frame(log_death, t, t ^ 2, t ^ 3)
df <- cbind(df, sinusoides)
ModAddifitf <- lm(data = df, log_death ~ .)
summary(ModAddifitf)
```

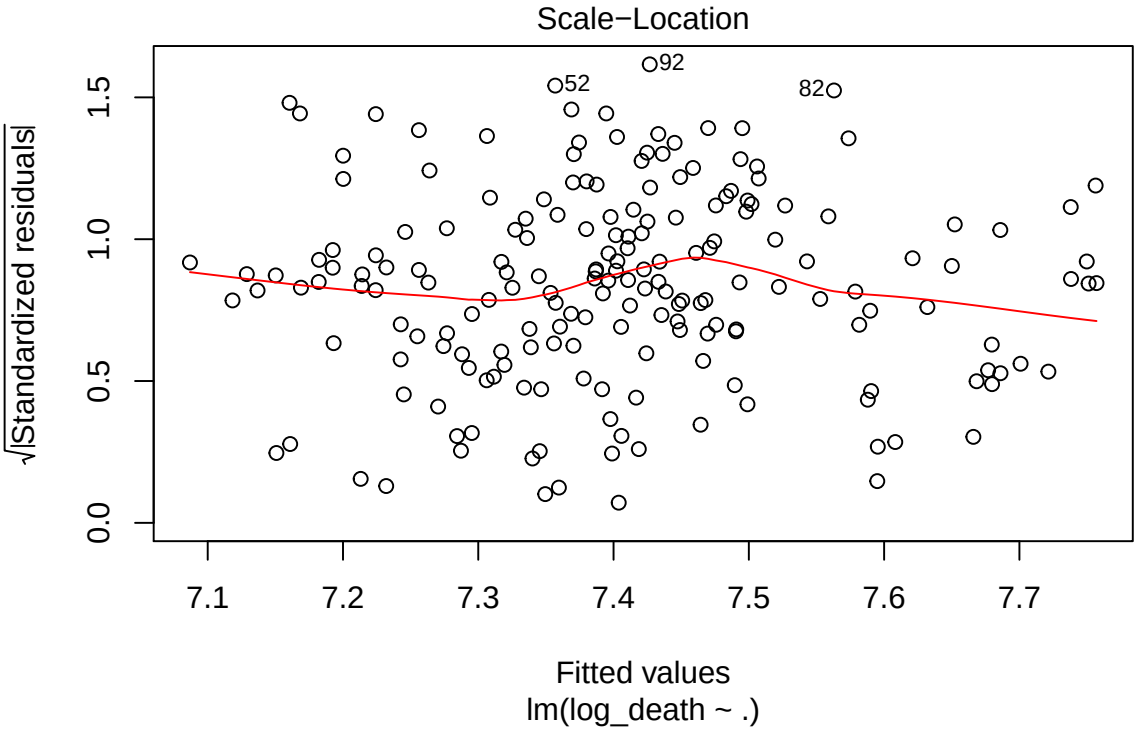
```
##
## Call:
## lm(formula = log_death ~ ., data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.236069 -0.064984  0.003402  0.064214  0.214065
##
```

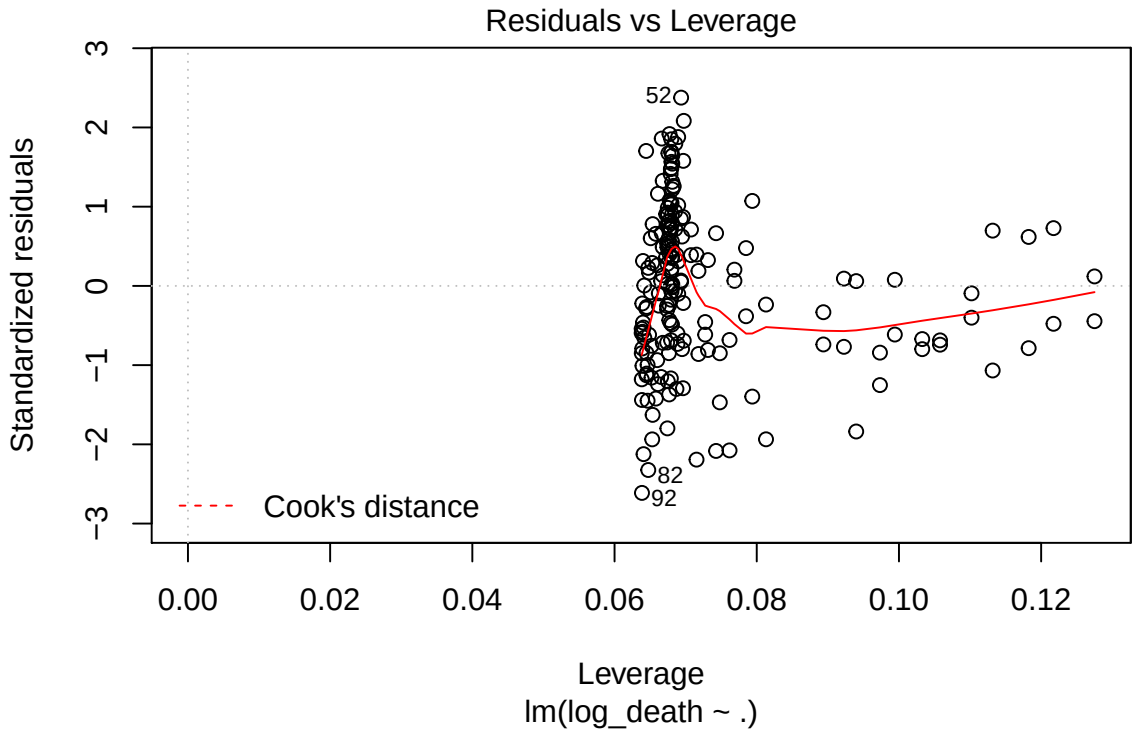
```
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.460e+00  2.760e-02 270.258  < 2e-16 ***
## t            2.068e-03  1.236e-03   1.673 0.096058 .
## t.2         -2.903e-05  1.486e-05  -1.954 0.052268 .
## t.3          5.937e-08  5.062e-08   1.173 0.242437
## sin_1       -7.471e-02  9.567e-03  -7.809 4.77e-13 ***
## sin_2       -3.595e-02  9.539e-03  -3.768 0.000223 ***
## sin_3       -1.897e-02  9.534e-03  -1.990 0.048128 *
## sin_4       -1.199e-02  9.532e-03  -1.258 0.210199
## sin_5        1.651e-02  9.531e-03   1.732 0.085045 .
## cos_1        1.147e-01  9.534e-03  12.034  < 2e-16 ***
## cos_2        6.297e-02  9.534e-03   6.606 4.42e-10 ***
## cos_3        3.101e-02  9.534e-03   3.253 0.001367 **
## cos_4        2.309e-02  9.534e-03   2.422 0.016458 *
## cos_5        2.564e-02  9.534e-03   2.689 0.007841 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09338 on 178 degrees of freedom
## Multiple R-squared:  0.7231, Adjusted R-squared:  0.7029
## F-statistic: 35.76 on 13 and 178 DF,  p-value: < 2.2e-16
```

```
plot(ModAddifitf)
```

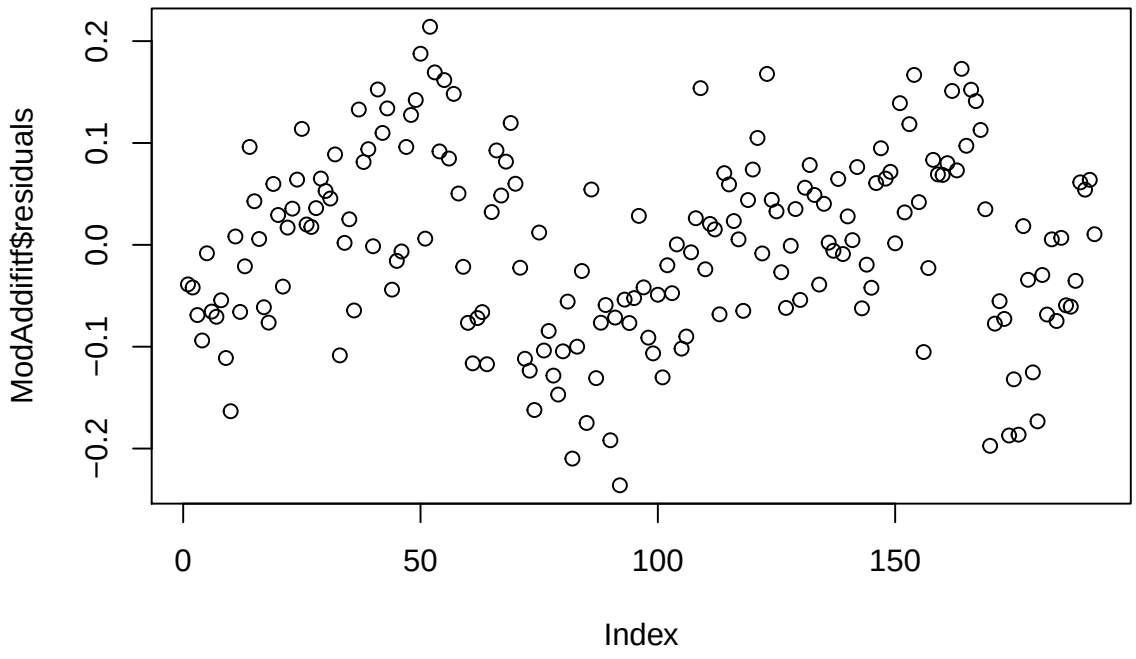



Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(log_death ~ .)
52
82
92
−3
−2
−1
0
1
2
3


Scale−Location
√|Standardized residuals|
Fitted values
lm(log_death ~ .)
52
82
92
0.0
0.5
1.0
1.5
7.1
7.2
7.3
7.4
7.5
7.6
7.7



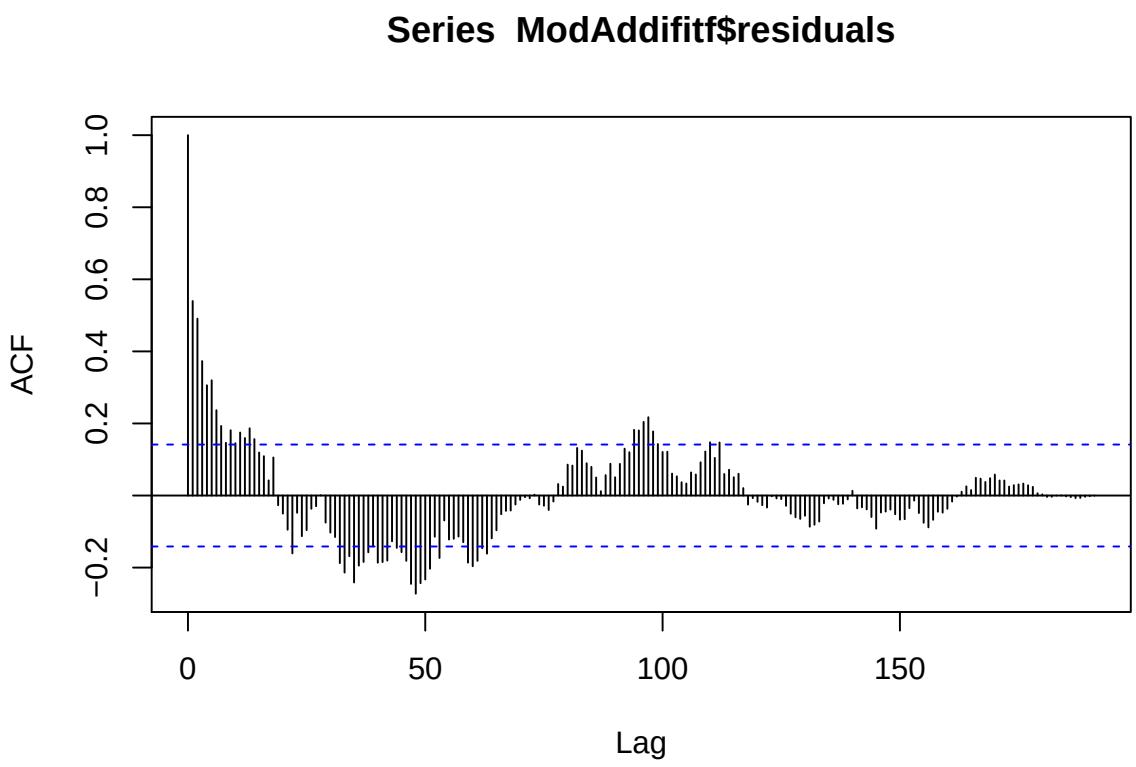

```
plot(ModAddifitf$residuals)
```



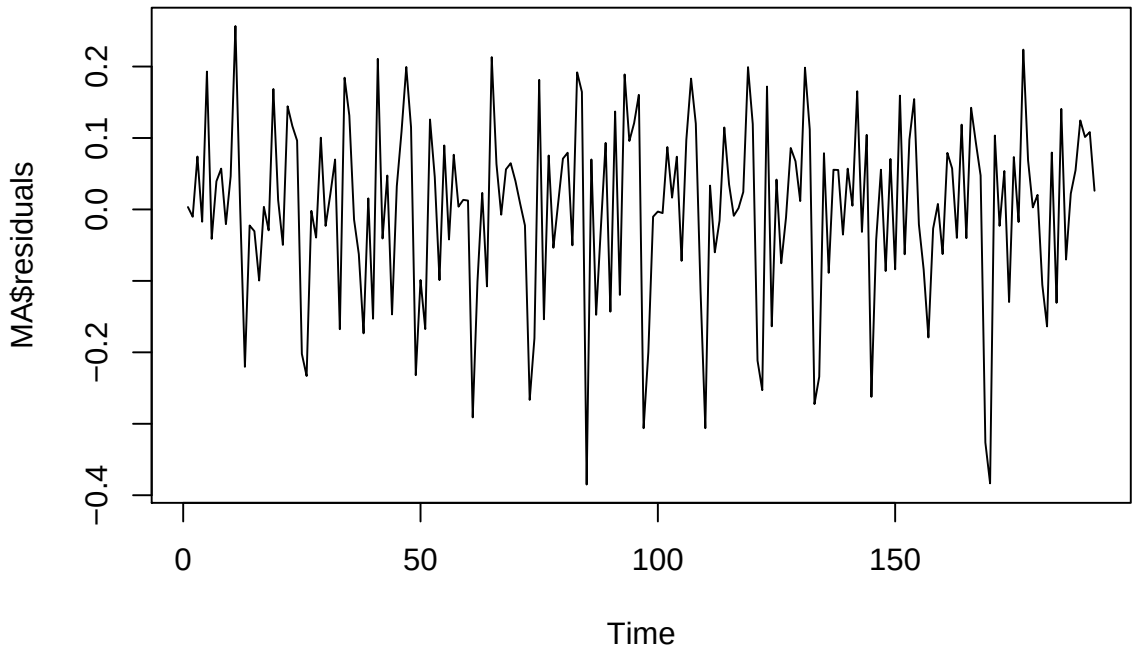
```r
t.test(ModAddifitf$residuals)
```

```
##
##  One Sample t-test
##
## data:  ModAddifitf$residuals
## t = 5.3199e-16, df = 191, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.01283293  0.01283293
## sample estimates:
##    mean of x
## 3.461175e-18
```

```r
acf(ModAddifitf$residuals, 192)
```




```r
MA <- arima(log_death, c(1, 2, 1))
plot(MA$residuals)
```


MA$residuals
−0.4
−0.2
0.0
0.1
0.2
0
50
100
150
Time