

T1-E03-regression_line_COL.R

jordi

Tue Oct 23 12:24:00 2018

```
setwd("~/Documents/CURS 2018-2019/PIE2")
COL <- read.csv2("./Dades/COL.csv")
#setwd("F:/windows")
COL <- read.csv2("./Dades/COL.csv")
p<-2
n<-dim(COL)[1]
library(car)
```

```
## Loading required package: carData
```

```
library(HH)
```

```
## Loading required package: lattice
```

```
## Loading required package: grid
```

```
## Loading required package: latticeExtra
```

```
## Loading required package: RColorBrewer
```

```
## Loading required package: multcomp
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: survival
```

```
## Loading required package: TH.data
```

```
## Loading required package: MASS
```

```
##
```

```
## Attaching package: 'TH.data'
```

```
## The following object is masked from 'package:MASS':
```

```
##
```

```
##      geyser
```

```
## Loading required package: gridExtra
```

```
##
```

```
## Attaching package: 'HH'
```

```
## The following objects are masked from 'package:car':
```

```
##
```

```
##      logit, vif
```

```
write("-----", "")
```

```
## -----
```

```
write("a)", "")
```

```
## a)
```

```
# MODEL
```

```
mod<-lm(C~W,COL)
```

```

# RESUM DEL MODEL
summary(mod)

##
## Call:
## lm(formula = C ~ W, data = COL)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -169.24  -39.81   -4.49   47.19  200.37
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 346.2251    33.1983   10.43 < 2e-16 ***
## W           -1.9835     0.5046   -3.93 0.000158 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 63.55 on 98 degrees of freedom
## Multiple R-squared:  0.1362, Adjusted R-squared:  0.1274
## F-statistic: 15.45 on 1 and 98 DF,  p-value: 0.0001581

#Càlculs opcionals: Intervals de confiança dels paràmetres
#confint(mod, level=0.99)

#Càlculs opcionals: SS1 Test dels paràmetres amb ordenació predeterminada
#anova(mod)

#Nota: SS3, els tests (F) sempre coincideixen amb els del resum (t),  $F=t^2$ 
#Anova(mod, ty=2)

write("-----", "")

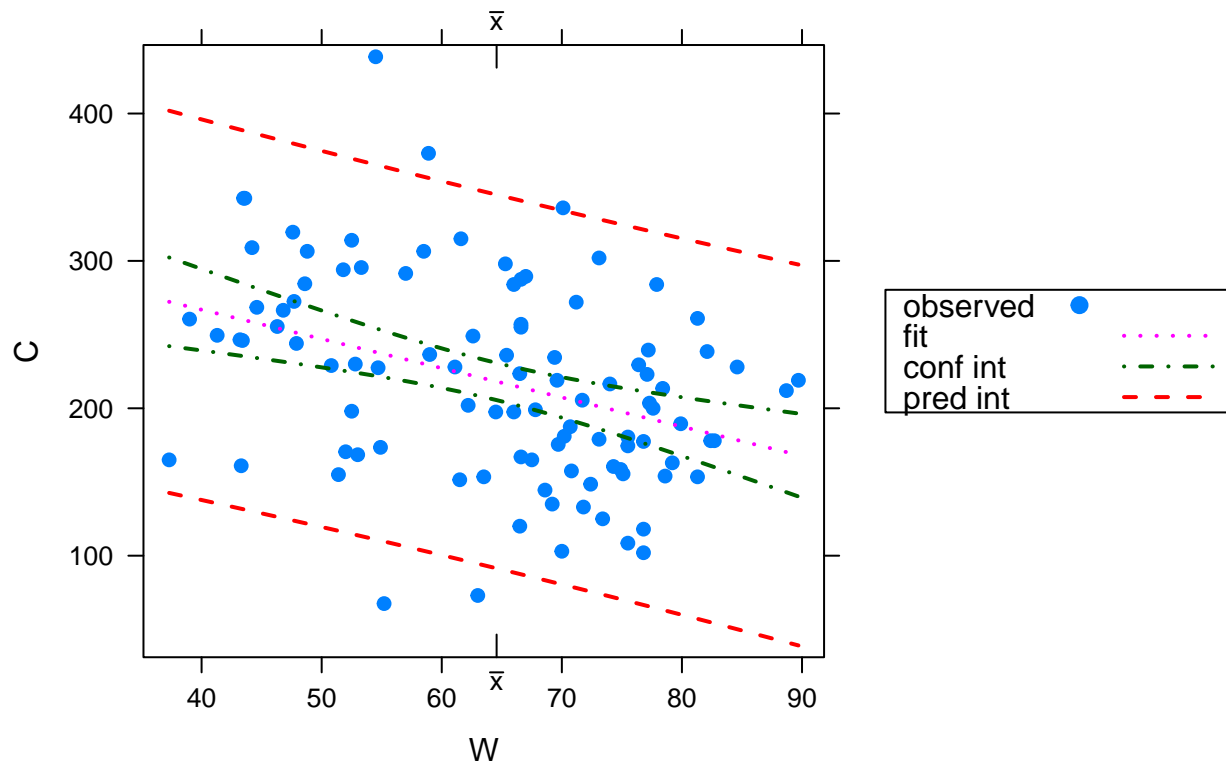
## -----
write("b)", "")

## b)

#Gràfica amb bandes de confiança i de predicció
ci.plot(mod)

```

95% confidence and prediction intervals for mod



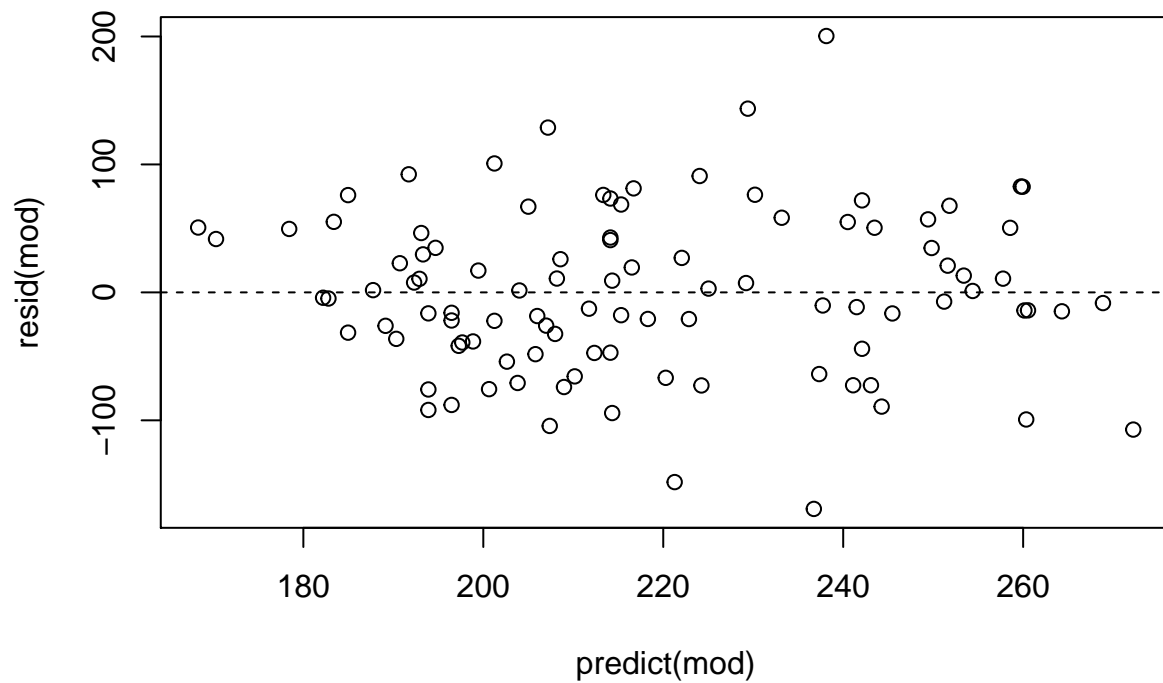
```
write("-----", "")
```

```
## -----
```

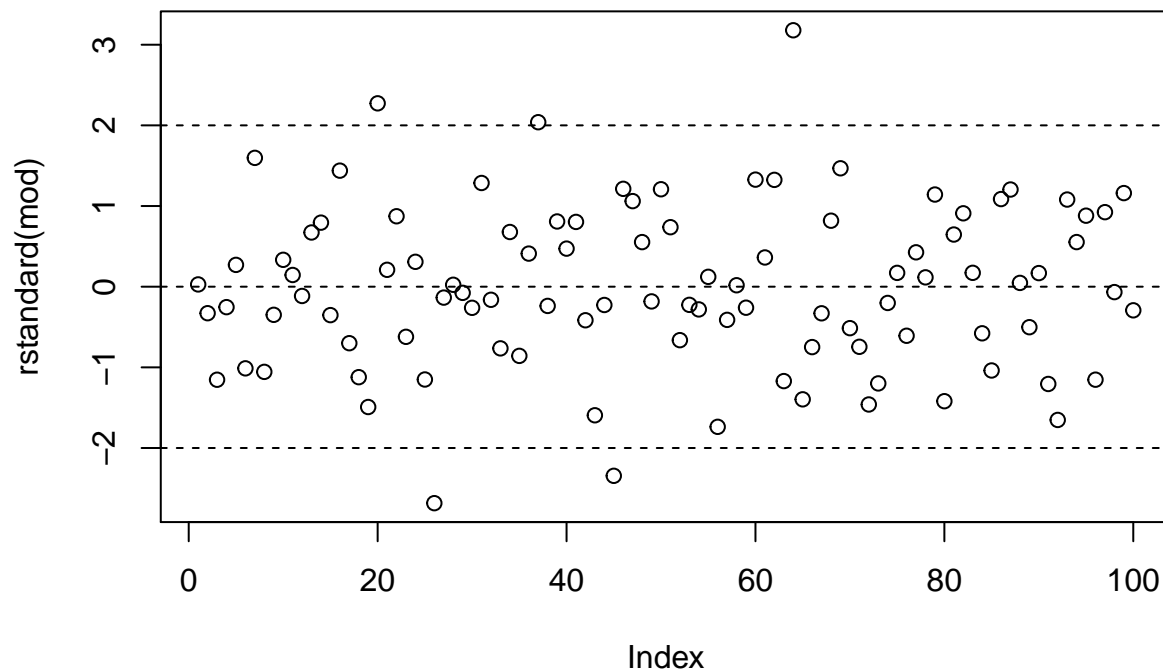
```
write("c)", "")
```

```
## c)
```

```
# Diagnòstic: TENDÈNCIES
plot(predict(mod), resid(mod))
abline(h=0, lty=2)
```



```
# Diagnostic: OUTLIERS (rstudent, rstandard)
plot(rstandard(mod))
abline(h=c(-2,0,2),lty=2)
```



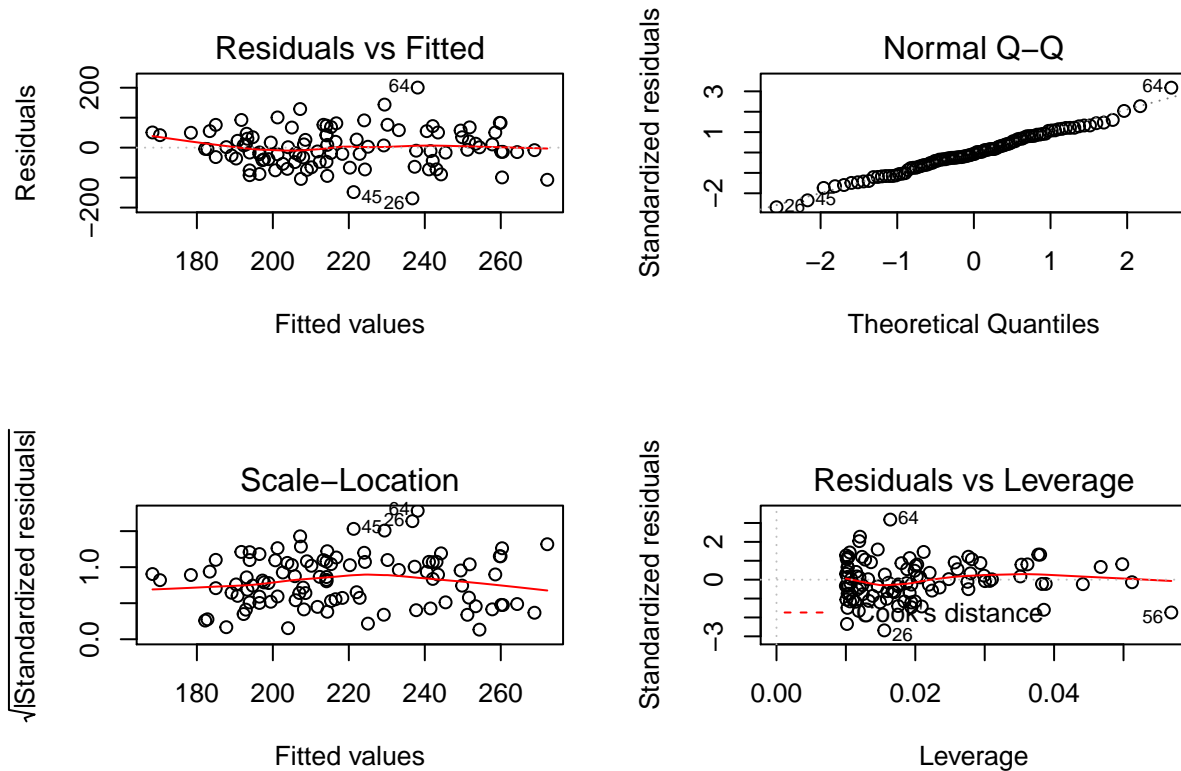
```
#plot(rstudent(mod),main="rstudent")
#abline(h=c(-2,0,2),lty=2)

# Diagnostic: LEVERAGE
#plot(hatvalues(mod))
#abline(h=c(0,2*mean(hatvalues(mod))),lty=2)
```

```
# Diagnòstic: INFLUÈNCIA (dffits, cooks.distance)
#plot(cooks.distance(mod))
#abline(h=c(0,4/n),lty=2)

#plot(dffits(mod),main="dffits")
#abline(h=c(-2*sqrt(p/n),0,2*sqrt(p/n)),lty=2)

#Diagnòstics de R
oldpar <- par( mfrow=c(2,2))
plot(mod,ask=F)
```



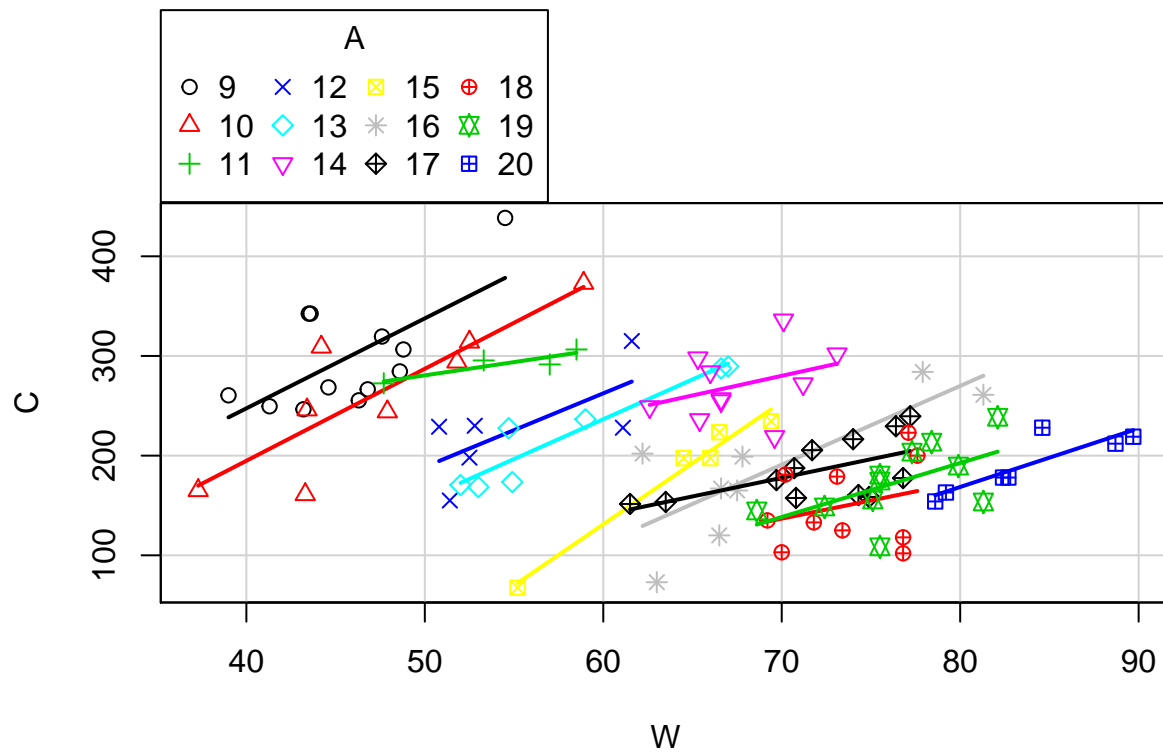
```
par(oldpar)

write("-----", "")

## -----

write("d)", "")

## d)
COL$A<-factor(COL$A)
sp(C~W|A,smooth=F,col=1:20, data=COL)
```



```
#COL$GA<-factor(ceiling((COL$A-8)/2))
#sp(C~W/GA,smooth=F, data=COL)
```