Régression linéaire sur le marché du Crédit

Theddy Morin

2022-11-09

# Modèle de régression linéaire sur le marché du Crédit

Dans ce document, on va procéder à la régression du nombre de jours de retard (qui est la variable expliquée) sur les variables Crédit et Type qui sont les variables explicatives

## Importation des données

Les données suivantes vont servir à procéder à la régression linéaire

dataR = read.csv("CreditBancaire.csv")  
attach(dataR)

## Régressons le nombre de jours de retard sur les variables Credit et Type

Le modèle de régression linéaire se présente ainsi:

dataR$Type = as.factor(dataR$Type)  
rlm = lm(Jours ~ Credit + Type, data = dataR)  
rlm

##   
## Call:  
## lm(formula = Jours ~ Credit + Type, data = dataR)  
##   
## Coefficients:  
## (Intercept) Credit TypeConsommation TypeProduction   
## 8.519e+01 -5.484e-06 -1.486e+01 1.498e+01

## Résultats détaillés de la régression

Voici en détails les résultats détaillés de la régression du nombre de jours de retard sur le montant et le type de crédit octroyé

summary(rlm)

##   
## Call:  
## lm(formula = Jours ~ Credit + Type, data = dataR)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -96.26 -69.22 -49.06 17.29 270.22   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.519e+01 1.275e+01 6.681 1.63e-10 \*\*\*  
## Credit -5.484e-06 1.182e-05 -0.464 0.643   
## TypeConsommation -1.486e+01 1.472e+01 -1.009 0.314   
## TypeProduction 1.498e+01 2.532e+01 0.591 0.555   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 104 on 241 degrees of freedom  
## Multiple R-squared: 0.007512, Adjusted R-squared: -0.004842   
## F-statistic: 0.608 on 3 and 241 DF, p-value: 0.6104

## Tableau partiel des coéfficients

Le tableau suivant affiche les statistiques de student et les probabilités par variable liées à la régression linéaire

summary(rlm)$coefficients[,3]

## (Intercept) Credit TypeConsommation TypeProduction   
## 6.6809256 -0.4637417 -1.0090929 0.5913659

summary(rlm)$coefficients[,4]

## (Intercept) Credit TypeConsommation TypeProduction   
## 1.630050e-10 6.432514e-01 3.139421e-01 5.548295e-01

## Tableau partiel des coéfficients

Le tableau suivant affiche les coéfficients estimés et les statistiques de student par variable liées à la régression linéaire

summary(rlm)$coefficients[,1]

## (Intercept) Credit TypeConsommation TypeProduction   
## 8.519082e+01 -5.483559e-06 -1.485726e+01 1.497503e+01

summary(rlm)$coefficients[,3]

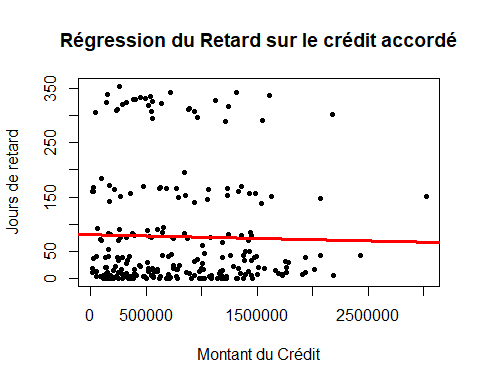
## (Intercept) Credit TypeConsommation TypeProduction   
## 6.6809256 -0.4637417 -1.0090929 0.5913659

## Visualisation graphique de la régression

Le graphique suivant permet d’apprécier les écarts du nuage de données par rapport à la droite de régression linéaire et ainsi avoir une idée de l’erreur standard

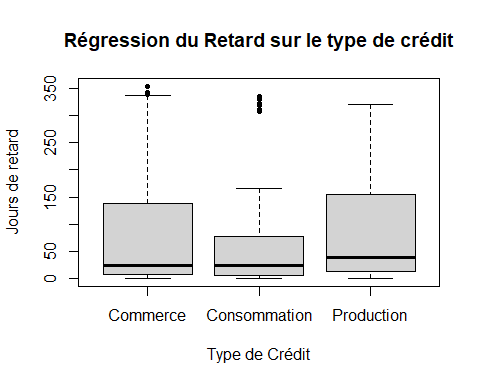
### Visualisons d’abord la corrélation du nombre de jours de retard et le montant de crédit accordé

plot(dataR$Credit, dataR$Jours, pch = 20, main = 'Régression du Retard sur le crédit accordé', xlab = 'Montant du Crédit', ylab = 'Jours de retard')  
abline(lm(Jours ~ Credit, data = dataR), col = 'red', lwd = 3)



### Visualisons maintenant la corrélation entre le nombre de jours de retard et le type de crédit accoré

plot(dataR$Type, dataR$Jours, pch = 20, main = 'Régression du Retard sur le type de crédit', xlab = 'Type de Crédit', ylab = 'Jours de retard')



## Création d’une fonction de régression

La fonction suivante permet de régresser une variable y sur une autre notée x, sachant que ces deux variables sont deux colonnes d’un data frame noté data

fregress <- function(data){   
 Estimates = summary(lm(y ~ ., data))$coefficients[, 1]  
 Stat\_Student = summary(lm(y ~ ., data))$coefficients[, 3]  
   
 return(data.frame(Estimates,Stat\_Student))  
}

### Test de la fonction

Testons la fonction qui vient d’être créé avec des nombres aléatoires

y<-c(23,67,34,90,109,345,4)  
x<-c(53,91,23,45,567,987,3)  
data <- data.frame(y,x)  
fregress(data)

## Estimates Stat\_Student  
## (Intercept) 24.295872 1.098911  
## x 0.283736 5.550305