

UE Probabilité et Statistique - Mathématiques

Devoir Maison n°8

Vincent Escoffier, Adrien Jallais, Théo Martel, Louis Muzellec.

27 janvier 2022

Contents

Préparation de l'environnement	1
Exercice 1	2
Description des données	2
Calculer un intervalle de confiance	2
Exercice 2	4
Description du modèle de données	4
Exercice 3	5
Description des données	5

Préparation de l'environnement

```
library(readr)
library(dplyr)
library(DescTools)
```

Exercice 1

Description des données

On importe les données avec la commande suivante :

```
d1 <- read_delim("../data/donnees.exo1.csv", ";", col_names = TRUE, escape_double = FALSE,
  show_col_types = FALSE)
```

L'effectif des données est de 35.

La table suivante résume la dispersion de nos données :

```
##          id          sinistre
## Min.      : 1.0    Min.      :0.0000
## 1st Qu.: 9.5    1st Qu.:0.0000
## Median :18.0    Median :0.0000
## Mean    :18.0    Mean     :0.5429
## 3rd Qu.:26.5    3rd Qu.:1.0000
## Max.    :35.0    Max.      :3.0000
```

Il est admis que l'occurrence des sinistres suit une loi de Poisson. Ainsi on note :

- X^2 : La valeur critique du khi-carré
- N : Le nombre d'événements observés
- α : Le niveau de signification

Calculer un intervalle de confiance

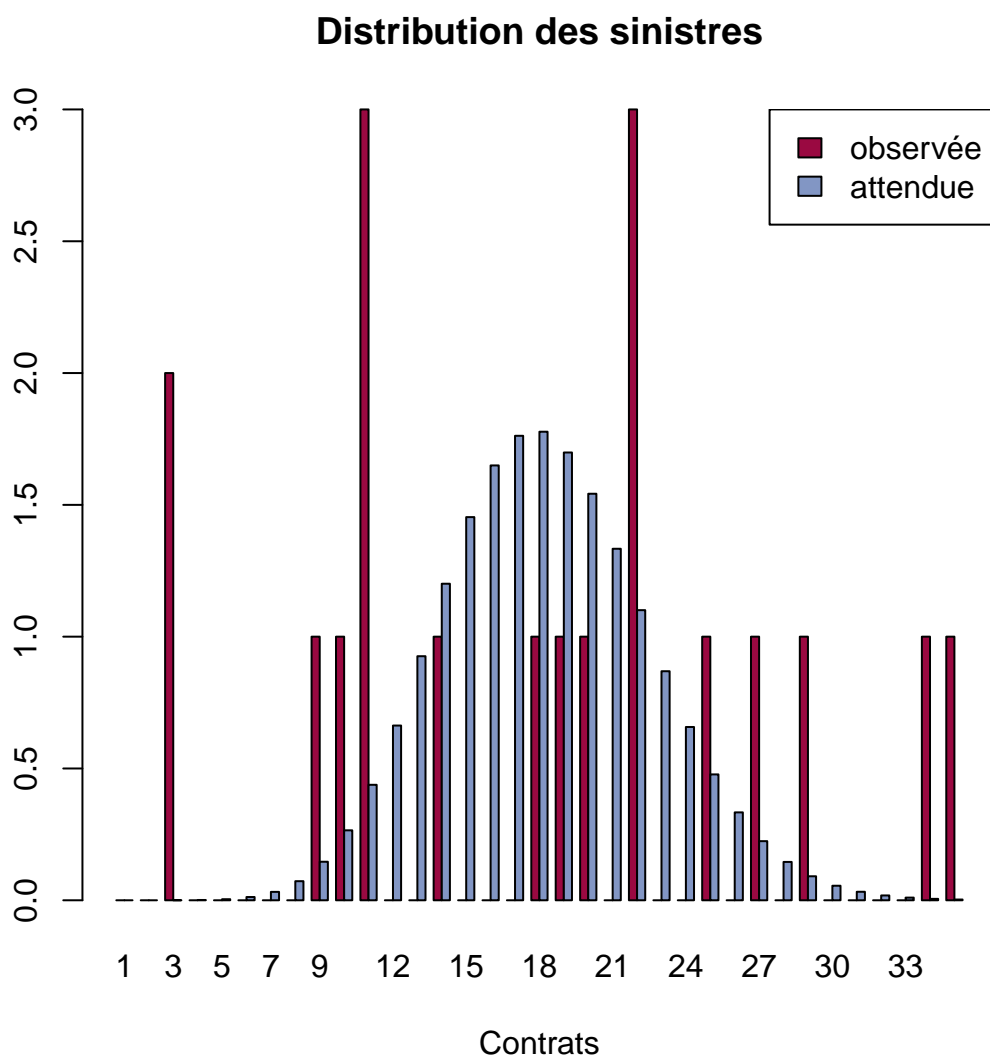
La formule de l'intervalle de confiance de Poisson est $IC = [0, 5 * X^2 N, \alpha/2, 0, 5 * X^2 (N + 1), 1 - \alpha/2]$.

On utilise les notations suivantes :

```
n <- sum(d1$sinistre)
x <- sum(d1$id * d1$sinistre)
lambda <- x/n
```

Afin de pouvoir afficher le graphique suivant :

```
exp <- dpois(d1$id, lambda) * n
barplot(rbind(d1$sinistre, exp * n/sum(exp)), names = d1$id, beside = TRUE, col = c(hred,
  hblue), main = "Distribution des sinistres", xlab = "Contrats")
legend("topright", legend = c("observée", "attendue"), fill = c(hred, hblue), bg = "white")
```



Ce graphique nous permet d'illustrer le calcul des intervalles de confiances avec les commandes suivantes :

```
ICL <- PoissonCI(x = x, n = n, conf.level = 0.95, sides = "left")
ICR <- PoissonCI(x = x, n = n, conf.level = 0.95, sides = "right")
```

L'intervalle de confiance à gauche est estimé à : 18.158.

L'intervalle de confiance à droite est estimé à : 18.158.

Exercice 2

Description du modèle de données

Exercice 3

Description des données

On importe les données avec les commandes suivantes :

```
d3 <- read_delim("../data/donnees.exo3.csv", ";", col_names = TRUE, escape_double = FALSE,
  show_col_types = FALSE, locale = locale(decimal_mark = ","))
d3$x <- as.numeric(d3$x)
d3$y <- as.numeric(d3$y)
d3 <- select(d3, x, y)
```

L'effectif des données est de 10.

La table suivante résume la dispersion de nos données :

##	x	y
##	Min. :0.500	Min. : 0.400
##	1st Qu.:1.625	1st Qu.: 2.345
##	Median :2.750	Median : 6.260
##	Mean :2.750	Mean : 7.620
##	3rd Qu.:3.875	3rd Qu.:11.852
##	Max. :5.000	Max. :19.900