Ajustement d'une V.A discrète par adjustr

Abdessabour MOUTIK

2020-07-30

Table des matières

Table des figures

La variable aléatoire **X** a étudié est le nombre de sinistres. D'abord, nous importons le fichier `nombresinistre30.csv` a notre session R on utilisons la fonction `read.csv` :

```
db.nbrsin <- read.csv2(path)
```

On affecte la deuxième colone de `db.nbrsin` a notre vecteur `x`

```
x <- db.nbrsin[[2]]
```

On a l'espérance et la variance de X sont :

$$\bar{X} = 1.78826086956522 \quad \text{ET} \quad \hat{\sigma}_X^2 = 0.969066134614294$$

On remarque que

$$\bar{X} > \hat{\sigma}_X^2$$

donc la distribution la mieux ajustée peut être celle d'une loi **binomiale**.

J'appelle alors ma fonction `getgoodfit` qui retourne une liste contenant l'*estimation*, les *effectifs théorique*, le *vecteur de probabilté*, le *degré de liberté*, la *p-value*, et l'$\tilde{\chi}^2$ pour les distributions données dans la colonne `distributions` du résultat. Simultanément, cette fonction crée le diagramme à barres des *effectifs observés* et les rootogrammes suspendus pour chaque ajustement.

```
result <- getgoodfit(x, short=FALSE, plots.as.vars = TRUE)
```

```
## [1] "Error occured whilst estimating the size n = NA of the negative binomial distribution."
```

Les distributions utilisées sont :

```
result$distributions
```

```
## [1] "pois"   "binom"  "nbinom"
```

On trace l'histogramme des nombres de sinistres **X**.

```
result$Xplot
```

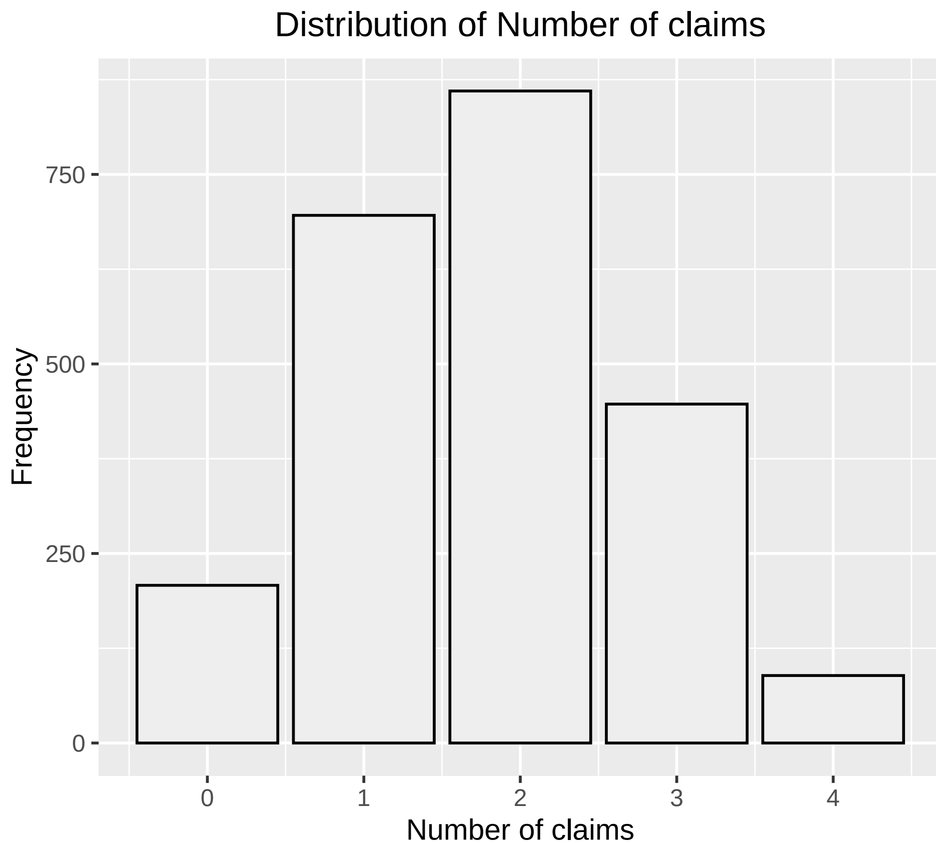

FIGURE 1 – Distribution des Nombre de sinistres

Ajustement par une distribution de loi de poisson

Après avoir utilisé `vcd::goodfit(x, "pois")` pour obtenir la distribution de poisson ajustée dans le corps de notre fonction `getgoodfit`. On retrouvera l'*estimation* suivantes:

```
result$pois$estimate
```

```
## $lambda
## [1] 1.788261
```

On affiche le rootograme suspendu :

```
result$pois$plot
```

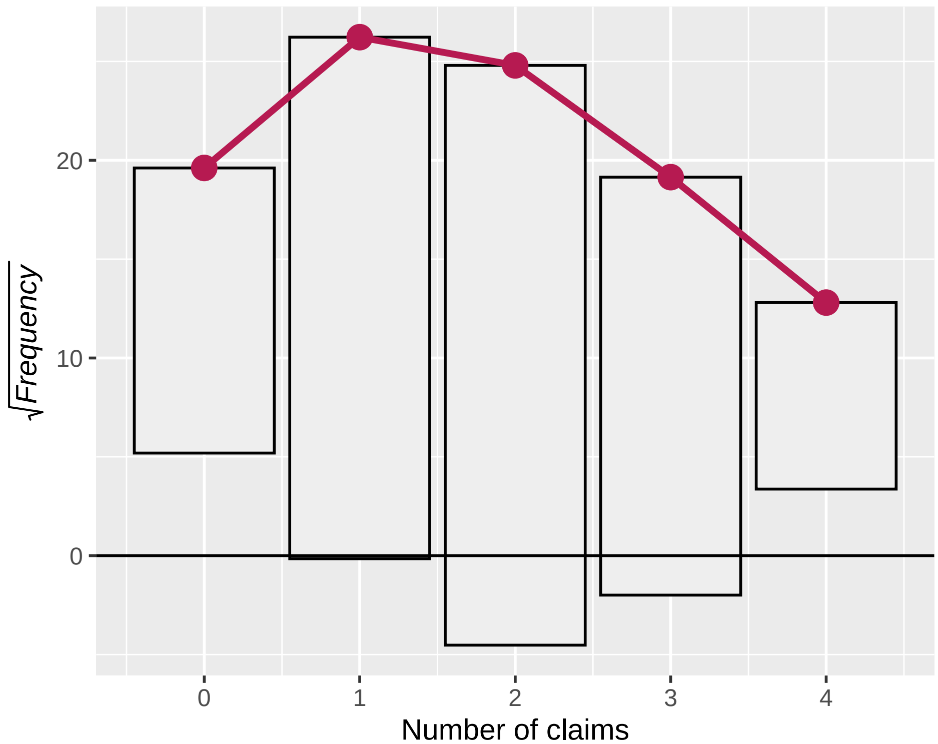

Figure 2 – Ajustement par une distrubition $\sim \mathcal{P}\,(1.79)$

On affiche les valeurs de **X**, les *effectifs théoriques*, les *effectifs observés* et le *vecteur de probabilté* :

```
result$pois
```

	Effectif		
X	**Obserevé**	**Ajusté**	**Probabilité**
0	208	384.677	0.167
1	696	687.902	0.299
2	860	615.075	0.267
3	447	366.638	0.159
4	89	163.911	0.107

En utilisant le test de $\tilde{\chi}^2{}_3$ on retrouvera la

$$pvalue = 0$$

et

$$\tilde{\chi}^2 = 296.33091310395$$

Le 95^{e} centile d'une loi χ_3^2

$$\chi_3^2(95\%) = 7.81472790325118$$

on a donc

$$\tilde{\chi}^2 > \chi_3^2(95\%)$$

et

$$pvalue < 5\%$$

On rejete l'hypthèse Nulle $\mathcal{H}_0$. Donc la distribution de loi **de poisson** ne permet pas de modèlisé celle de **X**.

Ajustement par une distribution de loi binomiale

Après avoir utilisé `vcd::goodfit(x, "binom", size=4)` pour obtenir la distribution binomiale ajustée dans le corps de notre fonction `getgoodfit`. On retrouvera l'*estimation* suivantes:

```
result$binom$estimate
```

```
## $prob
## [1] 0.4470652
##
## $size
## [1] 4
```

On affiche le rootograme suspendu :

```
result$binom$plot
```

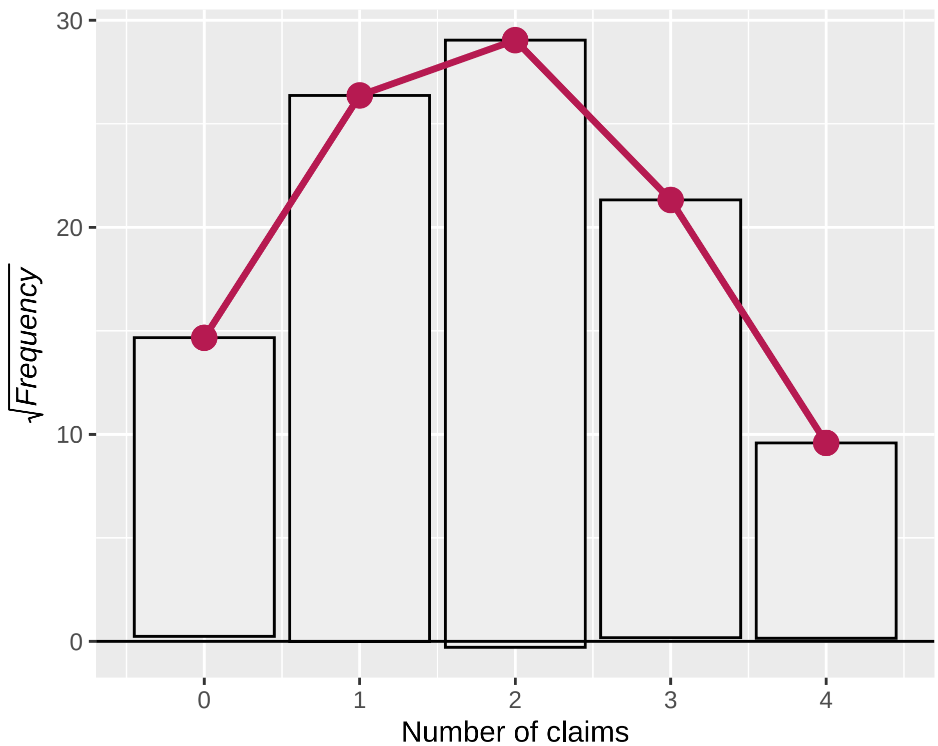

FIGURE 3 – Ajustement par une distrubition $\sim \mathcal{B}\,(0.45, 4)$

On remarque que la distribution binomiale est la mieux ajustée à la distribution de **X**.

On affiche les valeurs de **X**, les *effectifs théoriques*, les *effectifs observés* et le *vecteur de probabilté* :

```
result$binom
```

	Effectif		
X	**Obserevé**	**Ajusté**	**Probabilité**
0	208	214.993	0.093
1	696	695.313	0.302
2	860	843.274	0.367
3	447	454.542	0.198
4	89	91.878	0.040

En utilisant le test de $\tilde{\chi}^2{}_2$ on retrouvera la

$$pvalue = 0.678694205378728$$

et

$$\tilde{\chi}^2 = 0.775169226266107$$

Le 95^{e} centile d'une loi χ^2_2

$$\chi^2_2(95\%) = 5.99146454710798$$

on a donc

$$\tilde{\chi}^2 < \chi^2_2(95\%)$$

et

$$pvalue > 5\%$$

On accepte l'hypothèse Nulle $\mathcal{H}_0$. Donc la distribution de loi **binomiale** permet de bien modéliser la distribution la valeur aléatoire **X**.

Ajustement par une distribution de loi binomiale négative

Après avoir utilisé `vcd::goodfit(x, "nbinom", size=4)` pour obtenir la distribution binomiale négative ajustée dans le corps de notre fonction `getgoodfit`. On retrouvera l'*estimation* suivantes:

```
result$nbinom$estimate
```

```
## $size
## [1] 4
##
## $prob
## [1] 0.6910539
```

On affiche le rootograme suspendu :

```
result$nbinom$plot
```

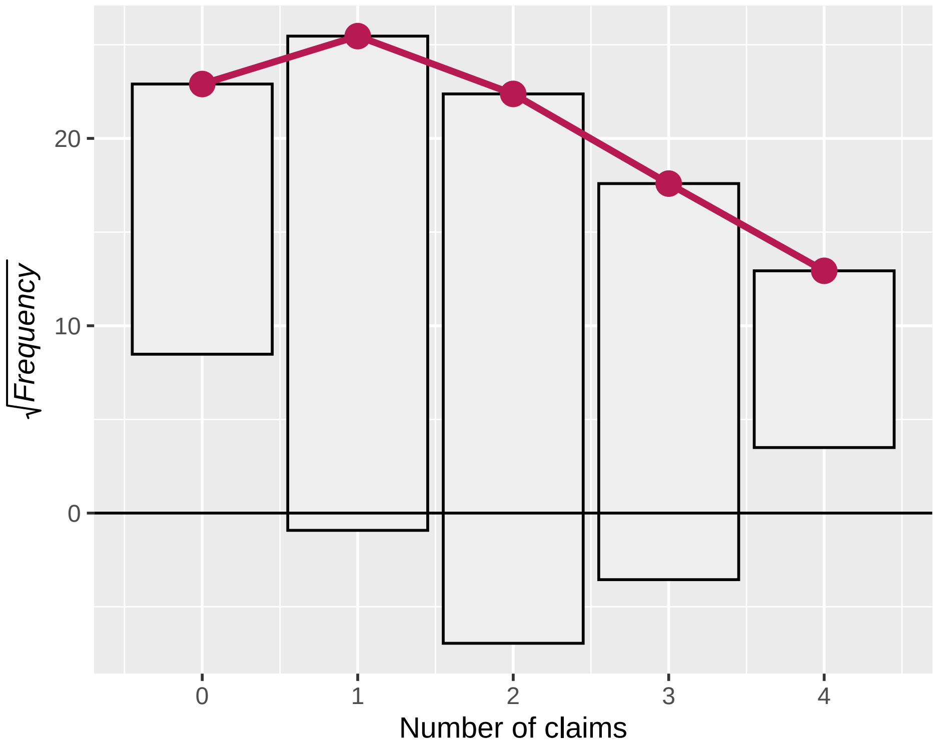

FIGURE 4 – Ajustement par une distrubition $\sim \mathcal{NB}\,(4, 0.69)$

On affiche les valeurs de **X**, les *effectifs théoriques*, les *effectifs observés* et le *vecteur de probabilté* :

```
result$nbinom
```

	Effectif		
X	**Obserevé**	**Ajusté**	**Probabilité**
0	208	524.536	0.228
1	696	648.214	0.282
2	860	500.658	0.218
3	447	309.353	0.135
4	89	167.253	0.138

En utilisant le test de $\tilde{\chi}^2{}_2$ on retrouvera la

$$pvalue = 0$$

et

$$\tilde{\chi}^2 = 677.908769377782$$

Le 95e centile d'une loi χ^2_2

$$\chi^2_2(95\%) = 5.99146454710798$$

on a donc

$$\tilde{\chi}^2 > \chi^2_2(95\%)$$

et

$$pvalue < 5\%$$

On rejete l'hypthèse Nulle $\mathcal{H}_0$. Donc la distribution de loi **binomiale négative** ne permet pas de modèlisé celle de $\mathbf{X}$.

Conclusion : La distribution la mieux ajustée est celle de la loi binomiale

On écrit :

$$\mathbf{X} \sim \mathcal{B}(0.45, 4)$$