TD Stat Inf

Julien JACQUES

15/11/2018

Exercice 4 TD2

Saisissons les données (sans la valeur 25) puis réalisons un test de Wilcoxon

```
x=c(25.6,24.5,24.3,29.5,24.1,24.8,24.7,25.2,24.9)
wilcox.test(x,mu = 25,alternative = "less")
```

```
## Warning in wilcox.test.default(x, mu = 25, alternative = "less"): cannot
## compute exact p-value with ties
```

```
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  x
## V = 17.5, p-value = 0.2968
## alternative hypothesis: true location is less than 25
```

La p-value de 0.2968 ne nous permet pas de rejetter H0. On ne peut pas affirmer que ce départment ne respecte pas la loi SRU.

Exercice 5 TD2

```
appels=0:9
nb=c(1,4,7,11,10,9,5,3,2,1)
```

On estime à la main les moyennes et variances du nombre d'appels

```
m=1/sum(nb)*sum(appels*nb)
print(m)
```

```
## [1] 4
```

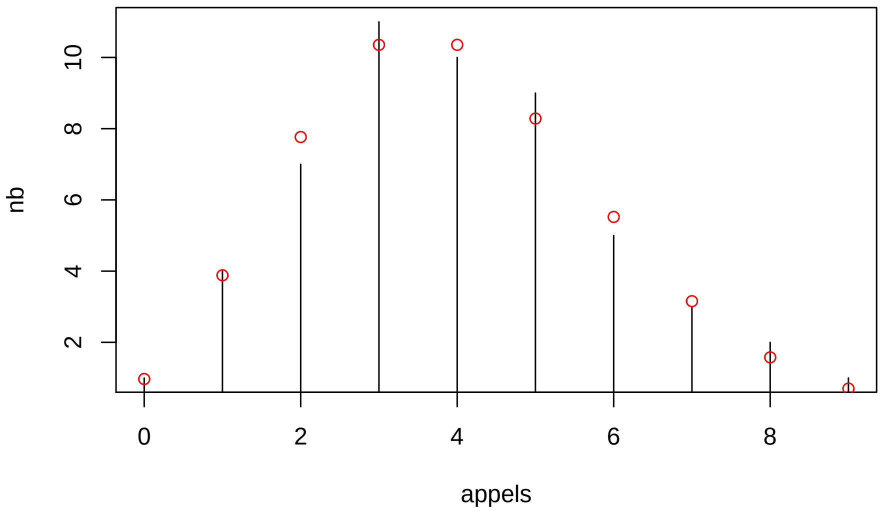
```
s2= 1/(sum(nb)-1)*sum(nb*(appels-m)^2)
print(s2)
```

```
## [1] 3.923077
```

La loi probable pour ces données est une loi de Poisson de paramètre 4.

Représentons l'histogramme ainsi que la distributoin selon une P(4)

```
plot(appels,nb,type='h')
points(appels,sum(nb)*dpois(appels,lambda = m),col=2)
```



Nous allons tester l'adéquation à la P(4) par un test du Chi2.

Calculons les probabilités théoriques et effectifs théoriques. Pour les probabilités théoriques de la classe 9 et +, on va la calculer comme 1 moins la somme de toutes les autres probabilités déjà calculées

```
pri=dpois(appels,lambda = m)
pri[10]=1-sum(pri[1:9])
```

Affichons le tableau contenant les nombres d'appels x_i observés par heure, les effectifs observés N_i, les probabilités théoriques p_i ainsi que les effectifs théoriques np_i :

```
n=sum(nb)
M=matrix(0,4,length(appels))
row.names(M)=c("xi","Ni","pi","npi")
M[1,]=appels
M[2,]=nb
M[3,]=pri
M[4,]=n*pri
print(M,digits=2)
```

```
##      [,1]  [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## xi  0.000 1.000 2.00  3.0  4.0 5.00  6.0 7.00 8.00 9.000
## Ni  1.000 4.000 7.00 11.0 10.0 9.00  5.0 3.00 2.00 1.000
## pi  0.018 0.073 0.15  0.2  0.2 0.16  0.1 0.06 0.03 0.021
## npi 0.971 3.883 7.77 10.4 10.4 8.28  5.5 3.16 1.58 1.132
```

Les effectifs théoriques dans les cases *sur les bords* sont inférieurs à 5. Il faut les regrouper :

```
regroup1=c(NA,sum(M[2,1:3]),sum(M[3,1:3]),sum(M[4,1:3]))
regroup2=c(NA,sum(M[2,8:10]),sum(M[3,8:10]),sum(M[4,8:10]))
Mregroupe=cbind(regroup1,M[,4:7],regroup2)
print(Mregroupe,digits=2)
```

```
##     regroup1                     regroup2
## xi        NA  3.0  4.0 5.00 6.0        NA
## Ni     12.00 11.0 10.0 9.00 5.0      6.00
## pi      0.24  0.2  0.2 0.16 0.1      0.11
## npi    12.62 10.4 10.4 8.28 5.5      5.87
```

Calculons les écarts $\frac{(N_i-np_i)^2}{np_i}$

```r
Mregroupe=rbind(Mregroupe,((Mregroupe[2,]-Mregroupe[4,])^2)/Mregroupe[4,])
row.names(Mregroupe)[5]='ecarts'
print(Mregroupe,digits=2)
```

```
##        regroup1                          regroup2
## xi           NA  3.00  4.000 5.000 6.000       NA
## Ni        12.00 11.00 10.000 9.000 5.000   6.0000
## pi         0.24  0.20  0.195 0.156 0.104   0.1107
## npi       12.62 10.35 10.354 8.284 5.522   5.8657
## ecarts     0.03  0.04  0.012 0.062 0.049   0.0031
```

```r
D2=sum(Mregroupe[5,])
print(D2)
```

```
## [1] 0.1972414
```

La somme de la dernière colonne est la statistique de test D^2, qu'il faut comparer au quantile de la loi du Chi2 (d'ordre $1-\alpha$) à 4 degrés de liberté : 6 (nombres de cases après regroupement) -1 (car on a estimé la moyenne pour connaitre le lambda à tester) - 1

```r
print(qchisq(0.95,4))
```

```
## [1] 9.487729
```

La valeur de D^2 est très en dessous du quantile $\chi^2_{4,0.95}$, on ne rejette pas H0, la loi de Poisson de paramètre 4 ne peut pas être rejettée. On l'accepte donc (avec un risque de seconde espèce β qu'il faudrait calculer…)

Finalement, on peut faire le test directement sous R avec la fonction *chisq.test*. Attention, cette fonction permet de faire une test d'ajustement mais aussi un test d'adéquation. Pour faire un test d'adéquation, il faut lui donner le vecteur d'effectifs observés ainsi que les probabilités théoriques :

```r
chisq.test(nb,pri)
```

```
## Warning in chisq.test(nb, pri): Chi-squared approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  nb and pri
## X-squared = 70, df = 64, p-value = 0.2833
```

On remarque le warning indiquant que *l'approximation du chi2 peut-être incorrecte* (problème des effectifs théoriques inférieurs à 5). Sous R, plutôt que de regrouper à la main (et perdre en puissance), on peut simuler la loi (au lieu d'utiliser la loi du chi2). Cela se fait comme ceci :

```r
chisq.test(nb,pri,simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  nb and pri
## X-squared = 70, df = NA, p-value = 1
```

Attention, on n'utilise cette option seulement quand on obtient un warning nous indiquant que l'approximation du chi2 peut être fausse.

La pvalue est bien au dessus de 5%, on ne rejette pas H0.