

TD 3 - χ^2 et Tests non-Paramétriques

1 Exercice 1

Tres largement inspire du document r5_prof.pdf, joint.

1. Charger les données du fichier : `tabac.txt`
2. Créer une variable de type TRUE/FALSE si tabac $\neq 0$
3. Créer le tableau de contingence entre fumeur et probleme pulmonaire et faire le test du χ^2
4. Vérifier que le problème pulmonaire n'est pas lié au sexe.
5. On peut refaire le test avec d'autres variables pour vérifier qu'il n'y a pas de biais.

Code R

```
> exo1<-read.table("tabac.txt", header=TRUE)
> exo1["fumeur"]<-exo1$tabac >0
> chisq.test(exo1$probleme_pulmonaire,exo1$fumeur)
```

Pearson's Chi-squared test with Yates' continuity correction

```
data: exo1$probleme_pulmonaire and exo1$fumeur
X-squared = 26.8279, df = 1, p-value = 2.224e-07
```

//////// Si l'on met le resultat dans une variable, on peut alors
interroger differents champs :

```
> res<-chisq.test(exo1$probleme_pulmonaire,exo1$fumeur)
```

```
> res$observed
```

```
                exo1$fumeur
exo1$probleme_pulmonaire FALSE TRUE
                FALSE    46   19
                TRUE     5   30
```

```
> res$expected
```

```
                exo1$fumeur
exo1$probleme_pulmonaire FALSE  TRUE
                FALSE 33.15 31.85
                TRUE  17.85 17.15
```

```
> res$p.value
```

```
[1] 2.224062e-07
> res$parameter
df
  1
$
```

2 Exercice 2

1. Mettre en pratique dans R l'exemple sur les prélèvements sanguins réalisés sur les NativeAméricains et les Caucasiens.

3 Exercice 3

On doit comparer deux séries de longévités d'animaux rares nés en captivité dans deux types de zoos qui utilisent des méthodes d'élevage différentes. La question est de savoir si un type de méthode d'élevage est meilleur que l'autre, en comparant 10 zoos utilisant la première méthode et 8 zoos utilisant la deuxième méthode. Ce type de naissance étant très rare, on a pu observer un seul cas par zoo. La distribution des longévités est une variable notoirement distincte de la loi normale (en particulier à cause de la mortalité en bas âge, qui peut être particulièrement élevée). Les longévités (en années) obtenues dans les deux types de zoos sont les suivantes :

Zoo de type A ($n_A = 10$ naissances, dans 10 zoos) : 1, 1, 1, 1, 5, 6, 6, 8, 9, 12

Zoo de type B ($n_B = 8$ naissances, dans 8 zoos) : 1, 7, 7, 8, 10, 13, 15

1. Charger les données comme ci-dessous
2. Appliquer le test de Wilcoxon .
3. Interpretation : Dans une table d'interprétation du test pour un effectif de , pour un risque de 0.05 on a un seuil de 29, comme 17.5 est inférieur à cette valeur, on rejette H_0 - en fait on peut même la rejeter avec un risque de 0.01 car le seuil est à 21.

Code R

```
> zoo1<-c(1, 1, 1, 1, 5, 6, 6, 8, 9, 12)
> zoo2<-c(1, 7, 7, 8, 10, 13, 15)
> wilcox.test(zoo1,zoo2)
```

Wilcoxon rank sum test with continuity correction

data: zoo1 and zoo2

W = 17.5, p-value = 0.0924

alternative hypothesis: true location shift is not equal to 0

Warning message:

In wilcox.test.default(zoo1, zoo2) : cannot compute exact p-value with ties

4 Exercice 4

La variable mesurée est la longueur de la mâchoire inférieure (en mm) de 10 chacals mâles et 10 chacals femelles (*Canis Aureus*) conservées au British Museum . La variable mesurée diffère-t-elle entre les sexes dans cette espèce ?

mâles 120 107 110 116 114 111 113 117 114 112

femelles 110 111 107 108 110 105 107 106 111 111

1. Après avoir rentrer les données, afficher l'histogramme pour constater que la distribution n'est pas normale.
2. Si on crée une matrice, on peut afficher le classement par rang
3. Le test de Wilcoxon rejete l'hypothese nulle d'égalité entre les males et les femelles pour la taille de leur machoire avec un risque de 0.0053

Code R

```
>y1 <- c(120,107,110,116,114,111,113,117,114,112)
>y2 <- c(110,111,107,108,110,105,107,106,111,111)
>hist(c(y1,y2), nclass = 8, col=grey(0.7), main="longueur de la mâchoire")
> mat <-c(y1,y2)
> rank(mat)
>> wilcox.test(y1, y2, correct = F)
```

Wilcoxon rank sum test

data: y1 and y2

W = 87.5, p-value = 0.004301

alternative hypothesis: true location shift is not equal to 0

Warning message:

```
In wilcox.test.default(y1, y2, correct = F) :
cannot compute exact p-value with ties
```