Fiche TD avec le logiciel R : tdr311

Choisir un test

D. Chessel, S. Dray, A.B. Dufour & J.R. Lobry

Corrigés d'exercices de la fiche tdr31

1 Silex

On a mesuré la dureté de 12 silex provenant de deux régions A et B, et on a classé les silex du moins dur au plus dur :

Rang	1	2	3	4	5	6	7	8	9	10	11	12
Origine	**A**	**A**	**A**	**B**	**A**	**A**	**B**	**A**	**B**	**B**	**B**	**B**

La dureté est-elle la même dans les deux régions ?

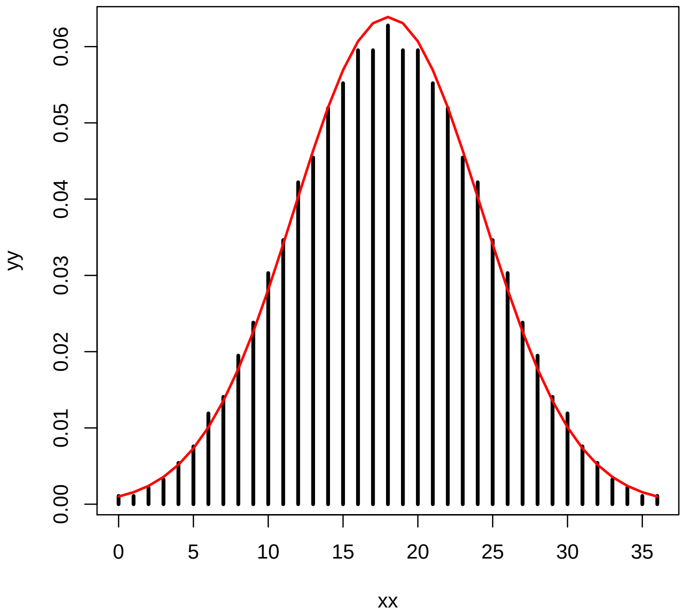
L'hypothèse à tester H_0 est "la dureté des silex est la même dans les deux régions". Le nombre de silex dans chaque échantillon est faible. Disposant des rangs, on ne peut que réaliser un test de Wilcoxon sur échantillons indépendants.

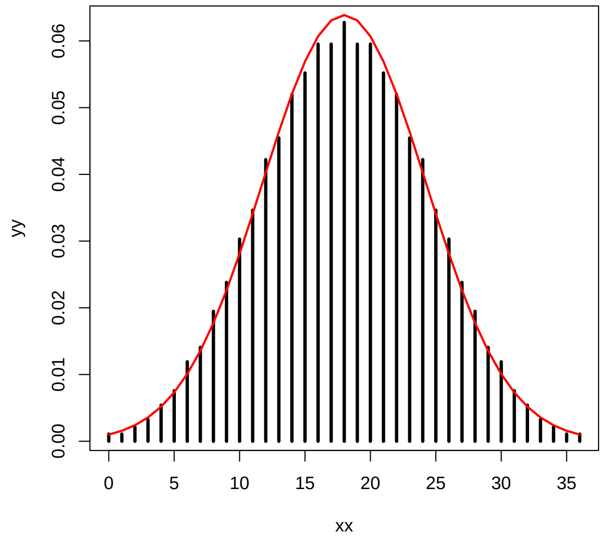
```
 originA <- c(1,2,3,5,6,8)
 originB <- c(4,7,9,10,11,12)
 sum(originA)
[1] 25
 sum(originB)
[1] 53
 wilcox.test(originA,originB,exact=T)
        Wilcoxon rank sum test
data:  originA and originB
W = 4, p-value = 0.02597
alternative hypothesis: true location shift is not equal to 0
```

La valeur de W=4 correspond à $25 - \frac{m(m+1)}{2}$ où m est le nombre de silex dans la région A. La plus petite somme des rangs de 6 valeurs prises parmi 12 est $1 + 2 + ... + 6 = 21$ La plus grande somme des rangs de 6 valeurs prises parmi 12 est $7 + 8 + ... + 12 = 57$ La plus petite valeur de W possible est donc 21-21 soit 0. La plus grande valeur de W possible est 57-21 soit 36. La distribution de Wilcoxon est :

```
 xx <- 0:36
 yy <- dwilcox(xx,6,6)
 2*pwilcox(4,6,6)
```

```
[1] 0.02597403
plot(xx,yy,type="h",lwd=3)
m1 <- sum(xx*yy)
m2 <- sum(yy*(xx-m1)*(xx-m1))
lines (xx,dnorm(xx,m1,sqrt(m2)),lwd=2,col="red")
```





Conclusion : On rejette l'hypothèse d'égalité de la dureté des silex dans les deux régions. A noter : la qualité de l'approximation normale pour la loi, même pour de petits échantillons équilibrés.

2 Fécondité

On a étudié[2] la fécondité d'une guêpe parasite (*Diadromus pulchellus*) en fonction de la présence d'un hôte au stade cocon (*Acrolepia assectella*). Deux lots A et B de parasites sont formés. Aux parasites du lot A, on présente un hôte les jours (du 1^{er} au $35^{ème}$ jour), à ceux du lot B à partir du $6^{ème}$ jour seulement (du $6^{ème}$ au $35^{ème}$ jour). Les nombre d'œufs pondus par chaque insecte sont les suivantes :

Lot A (13 insectes) : `98, 84, 63, 75, 84, 66, 56, 48, 48, 109, 85, 95, 106`
Lot B (18 insectes) : `124, 83, 75, 123, 105, 108, 155, 128, 56, 72, 96, 45, 71, 45, 73, 60, 89, 83`

Commenter les résultats. L'hypothèse à tester H_0 est "En moyenne, le nombre d'œufs pondus par chaque insecte est le même que l'hôte soit présenté tous les jours ou seulement à partir du sixième jour". Il n'y a rien à dire sur les effectifs dans les deux groupes.

```
lotA <- c(98,84,63,75,84,66,56,48,48,109,85,95,106)
lotB <- c(124,83,75,123,105,108,155,128,56,72,96,45,71,45,73,60,89,83)
summary(lotA)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 48.00   63.00   84.00   78.23   95.00  109.00 
summary(lotB)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
 45.00   71.25   83.00   88.39  107.25  155.00 
```

2.1 Ce qu'il ne faut pas faire

Testons la normalité des distributions.

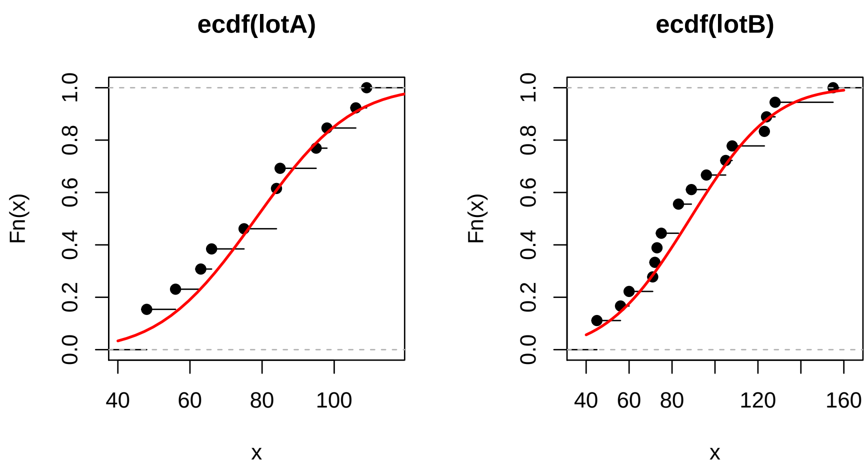
```
ks.test(lotA,"pnorm",mean(lotA),sd(lotA))
        One-sample Kolmogorov-Smirnov test
data:  lotA
D = 0.14746, p-value = 0.94
alternative hypothesis: two-sided
```

Le test de Kolmogorov-Smirnov, dans la fonction `ks.test` est réservé soit à la comparaison de deux échantillons, soit à l'ajustement d'une loi dont les paramètres sont connus *a priori* :

```
If a single-sample test is used, the parameters specified in '...'
     must be pre-specified and not estimated from the data. There is
     some more refined distribution theory for the KS test with
     estimated parameters (see Durbin, 1973), but that is not
     implemented in 'ks.test'.
```

2.2 Ce qu'il faut faire

```
shapiro.test(lotA)
        Shapiro-Wilk normality test
data:  lotA
W = 0.94314, p-value = 0.4988
shapiro.test(lotB)
        Shapiro-Wilk normality test
data:  lotB
W = 0.9589, p-value = 0.5806
par(mfrow=c(1,2))
w0 <- seq(40,160,le=50)
plot(ecdf(lotA))
lines(w0,pnorm(w0,mean(lotA),sd(lotA)),lwd=2,col="red")
plot(ecdf(lotB))
lines(w0,pnorm(w0,mean(lotB),sd(lotB)),lwd=2,col="red")
```



Conclusion : Rien ne s'oppose à la normalité des variables.

2.3 Ce qu'il ne faut pas faire

Testons l'égalité des deux variances par un test de Fisher.

```
var(lotA)
[1] 434.6923
var(lotB)
[1] 932.7222
var(lotB)/var(lotA)
[1] 2.145707
1-pf( var(lotB)/var(lotA),17,12)
[1] 0.09149241
```

En faisant cela, on fait un test unilatéral contre l'hypothèse alternative que la variance de A est plus grande que la variance de B. Il n'y a rien qui indique que ceci ait un sens biologique. Le rapport inverse porte sur l'hypothèse alternative que la variance de B est plus grande que la variance de A. Il n'y a pas plus d'indications que ceci ait un sens biologique. On se concentre sur un test bilatéral.

2.4 Ce qu'il faut faire

```
var.test(lotA,lotB)
        F test to compare two variances
data:  lotA and lotB
F = 0.46605, num df = 12, denom df = 17, p-value = 0.183
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.164979 1.458093
sample estimates:
ratio of variances
         0.4660469
bartlett.test(c(lotA,lotB),as.factor(rep(c("A","B"),c(13,18))))
        Bartlett test of homogeneity of variances
data:  c(lotA, lotB) and as.factor(rep(c("A", "B"), c(13, 18)))
Bartlett's K-squared = 1.8545, df = 1, p-value = 0.1733
```

Conclusion : On ne peut pas rejeter l'égalité des variances. Afin de prendre une décision, compte tenu de ce qui précède, nous allons appliquer le test de Student sur échantillons indépendants.

```
t.test(lotA,lotB,var.equal=T)
        Two Sample t-test
data:  lotA and lotB
t = -1.0353, df = 29, p-value = 0.3091
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -30.224771   9.908532
sample estimates:
mean of x mean of y
 78.23077  88.38889
wilcox.test(lotA,lotB)
        Wilcoxon rank sum test with continuity correction
data:  lotA and lotB
W = 102, p-value = 0.5614
alternative hypothesis: true location shift is not equal to 0
y <- c(lotA,lotB)
fac <- as.factor(rep(c("A","B"),c(13,18)))
anova(lm(y~fac))
Analysis of Variance Table
Response: y
          Df  Sum Sq Mean Sq F value Pr(>F)
fac        1   778.9  778.90  1.0719 0.3091
Residuals 29 21072.6  726.64
```

Conclusion : Rien dans ces petits échantillons ne permet de mettre en évidence une différence de moyenne. Ou encore : les moyennes ne sont pas significativement différentes (ce qui ne veut pas dire qu'elles sont égales, évidemment).

3 Alcool

On a étudié[1] le temps de réaction nécessaire pour arrêter une automobile chez des sujets sous l'influence de trois onces d'alcool (environ 0,09 l). On mesure le temps de réaction (en 100ème de seconde) avant et après l'ingestion d'alcool.

Cobaye	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Avant	33	29	26	23	21	36	27	38	22	33	42	35	22	39	37
Après	46	41	37	37	30	43	38	47	33	42	54	48	33	54	50

Conclusion ?

```
après <- c(46,41,37,37,30,43,38,47,33,42,54,48,33,54,50)
avant <- c(33,29,26,23,21,36,27,38,22,33,42,35,22,39,37)
après - avant
[1] 13 12 11 14  9  7 11  9 11  9 12 13 11 15 13
```

Dans la mesure où toutes les différences sont positives, il paraît absurde de faire un test. Les temps de réactions sont plus longs lorsqu'un individu a bu de l'alcool. Plus précisément, si l'alcool n'avait aucun effet (ça se saurait) on aurait 1 chance sur deux pour chacune des personnes testées, donc 2 chances sur 32768 de trouver tous les résultats dans le même sens. Le résultat est significatif au seuil de 1 pour 1000.

4 Filtres

Dans une étude sur le traitement des eaux usées, l'efficacité de deux filtres, l'un en fibre de verre et l'autre en papier filtre Whatman 40 a été testée[3]. Sur des prélèvements de 200 millilitres d'eau provenant d'une usine de pâtes à papier, la quantité de solides en suspension retenus par les deux filtres a été mesurée.

Prélèvement	1	2	3	4	5	6	7	8
Verre	68	86	91	77	100	138	190	110
Papier	64	75	72	63	96	120	184	96

L'efficacité du filtre de verre est-elle supérieure à celle du papier filtre ?

Dans la mesure où toutes les différences sont positives, il paraît inutile de faire un test. Un test bilatéral immédiat donne $p = \frac{2}{2^8} = 0.0078125$ Le cas est donc moins trivial que le précédent. On peut tester le caractère supérieur du verre, donc utiliser un test unilatéral. Pour vérifier :

```
verre <- c(68,86,91,77,100,138,190,110)
papier <- c(64,75,72,63,96,120,184,96)
wilcox.test(verre,papier, alternative="greater",paired=T)
        Wilcoxon signed rank test with continuity correction
data:  verre and papier
V = 36, p-value = 0.007015
alternative hypothesis: true location shift is greater than 0

t.test(verre,papier, alternative="greater",paired=T)
        Paired t-test
data:  verre and papier
t = 5.2876, df = 7, p-value = 0.0005693
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 7.219048      Inf
sample estimates:
mean of the differences
                  11.25
```

Conclusion : L'efficacité du filtre de verre est supérieure à celle du papier filtre.

5 Force

On mesure la force statique par dynamométrie manuelle (exprimée en kg) chez 10 enfants atteint de trisomie 21. Un premier test est réalisé en septembre. Puis, sur une période de six mois, ces enfants essaient de développer, sous forme de jeux, leur force statique. Un second test est réalisé en février.

33	42.5	54	60	61	68	68	69	72	86
38	45	52	63	61	75	66	70	81	90

Peut-on dire que le programme suivi par ces enfants a permis d'améliorer leur résultat ?

```
sept <- c(33,42.5,54,60,61,68,68,69,72,86)
fev <- c(38,45,52,63,61,75,66,70,81,90)
fev-sept
[1]  5.0  2.5 -2.0  3.0  0.0  7.0 -2.0  1.0  9.0  4.0
rank(abs(fev-sept))
[1]  8.0  5.0  3.5  6.0  1.0  9.0  3.5  2.0 10.0  7.0
rank(abs(fev-sept))-1
[1] 7.0 4.0 2.5 5.0 0.0 8.0 2.5 1.0 9.0 6.0
wilcox.test(sept,fev,paired=T)
        Wilcoxon signed rank test with continuity correction
data:  sept and fev
V = 5, p-value = 0.04383
alternative hypothesis: true location shift is not equal to 0
t.test(sept,fev,paired=T)
        Paired t-test
data:  sept and fev
t = -2.3853, df = 9, p-value = 0.04087
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -5.3580295 -0.1419705
sample estimates:
mean of the differences
                  -2.75
```

Apparemment, la somme des rangs des différences négatives est 7 mais lorsqu'il y a des sujets qui n'ont ni progressé, ni régressé, aucun signe ne peut leur être attribué. Donc la série est modifiée et ne comporte plus que 9 individus. La somme des rangs des différences négatives est donc finalement 5. L'identité des résultat entre les deux tests est remarquable. On a une indication sérieuse de l'augmentation de la force des enfants, mais il peut y avoir confusion de facteur avec la croissance. Remarque : dans les données appariées, l'étude de la distribution des échantillons n'a pas de sens.

6 Test de Wilcoxon : informations complémentaires

- Dans le cas des petits effectifs–($n \leq 15$) où n est le nombre de différences non nulles)–on peut lire par exemple la probabilité (cas unilatéral) dans la table[4] correspondante. La valeur de p est alors 0.039.
- Dans le cas des grands effectifs ($n > 15$), on calcule la probabilité à partir des informations suivantes :
- si n <50 et si la distribution ne comporte ni ex-aequo ni différence nulle, `wilcox.test` calcule la probabilité exacte.
- si n est quelconque et si la distribution comporte des ex-aequo ou des différences nulles, la valeur calculée de la statistique du test est :

$$T = \frac{\left| W - \frac{n(n+1)}{4} \right| - \frac{1}{2}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{1}{2}\left(\frac{\sum (t_i^3 - t_i)}{24} \right)}}$$

où t_i est le nombre d'ex-aequo associé à une même valeur. La statistique du test converge alors vers une loi normale centrée réduite.

— si $n > 50$ et si la distribution ne comporte pas d'ex-aequo, la valeur calculée de la statistique du test est :

$$T = \frac{\left| W - \frac{n(n+1)}{4} \right| - \frac{1}{2}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}.$$

La statistique du test converge alors vers une loi normale centrée réduite. Dans les deux cas précédents, si on pose comme argument `correct=F`, la correction (-1/2) du numérateur est supprimée.

Références

[1] D.E. Honkle, W. Wiersma, and S.G. Jurs. *Applied Statistics for the Behavorial Sciences.* Houghton Mifflin, 1988.

[2] V. Labeyrie. Contribution à l'étude de la dynamique des populations d'insectes. influence de l'hôte *Acrolepia assectella* z. sur la multiplication d'un hymenoptère ichneumonidae (*Diadromus* sp.). *Entomophaga*, Mémoire hors-série 1, 1960.

[3] B. Scherrer. *Biostatistique.* Gaëtan Morin éditeur, 1984.

[4] S. Siegel and N.J. Castellan Jr. *Nonparametric Statistics for the Behavorial Sciences.* McGraw-Hill, 1989.