L3 Pro "Biotechnologies végétales et création variétale"
TP 3 : tests classiques avec R

D. Chessel, A-B. Dufour, T. Jombart, J.R. Lobry & M. Bailly-Bechet

Automne 2012

1 Anatomie d'un test

On dispose de données sur le poids d'un échantillon de 15 chats sauvages attrapés dans les Ardennes :

```
poids<-c(5.72,4.06,4.30,3.45,3.08,2.68,4.86,5.76,4.90,4.28,5.13,3.93,4.95,4.72,5.85)
```

On veut vérifier si le poids de ces chats est similaire au poids de référence de l'espèce en France, 5 kilos. On peut calculer la moyenne :

```
mean(poids)
```

```
[1] 4.511333
```

Ces chats sont plus légers ; mais on ne peut pas conclure, *uniquement sur la base de la différence observée*, que les chats sauvages des Ardennes sont plus légers que la moyenne des chats sauvages français. En effet, il est possible que les chats attrapés soient plus léger *par hasard*, mais que l'ensemble des chats sauvages des Ardennes ait en moyenne le même poids que leurs homologues français. Pour savoir si l'écart observé est du au hasard on va pratiquer un test statistique ; dans cette situation ce sera un test de *comparaison à une moyenne de référence*.

Comme tout test, celui-ci va chercher à distinguer entre deux hypothèses, l'une simple et mathématiquement précise (H_0) et l'autre plus floue, qui regroupe tout ce qui n'est pas inclus dans la précédente (H_1). Ici donc on a :

H_0 : la moyenne des poids des chats sauvages des Ardennes est *la même* que celle des chats sauvages français.

H_1 : la moyenne des poids des chats sauvages des Ardennes est *différente de* celle des chats sauvages français.

La commande permettant de réaliser ce test sous R est :

```
t.test(poids,mu=5)
```

```
        One Sample t-test

data:  poids
t = -1.9738, df = 14, p-value = 0.06848
alternative hypothesis: true mean is not equal to 5
95 percent confidence interval:
 3.980326 5.042341
sample estimates:
mean of x
 4.511333
```

Que veulent dire les résultats affichés ?

One Sample t-test : le test utilisé par .

data : le jeu de données testé.

t= : la valeur de la statistique observée, que l'on retrouverait si on avait fait le test à la main.

df= : le nombre de degrés de liberté employé dans l'interprétation du test, qui nous serait utile si on faisait le travail à la main.

p-valeur= : la p-valeur du test, soit l'un des éléments les plus importants pour nous. C'est à partir de cette valeur que l'on va raisonner par la suite.

alternative hypothesis : l'hypothèse alternative du test ; rappelle toujours cette hypothèse, sans rapport avec la p-valeur et le fait que cette hypothèse soit plus ou moins proche de la vérité.

95 percent confidence interval : l'estimation par intervalle de confiance à 95% de confiance de la moyenne du poids des chats sauvages des Ardennes.

sample estimates : les valeurs intermédiaires calculées par , ici la moyenne de notre vecteur.

Pour conclure sur ce test, il faut interpréter la p-valeur renvoyée par , qui représente les chances que l'échantillon analysé ait été obtenu si H_0 était vraie. Si ces chances sont trop faibles, on concluera qu'on ne peut pas accepter H_0, et donc que H_1 doit être vraie ; si elle est trop grande, on en concluera que l'on ne peut pas rejeter H_0, et on conservera cette hypothèse *parce qu'elle est la plus simple*. Dans le premier cas (accepter H_1), on connaît bien le risque que l'on prend de se tromper : c'est la p-valeur. Dans le second (conserver H_0) par défaut, on ne le connaît pas bien : c'est le risque de deuxième espèce noté β.

Classiquement, on fixe le seuil sur la p-valeur à 5% ; mais cela pourra évoluer en fonction du nombre de tests effectués et de leur enjeu.

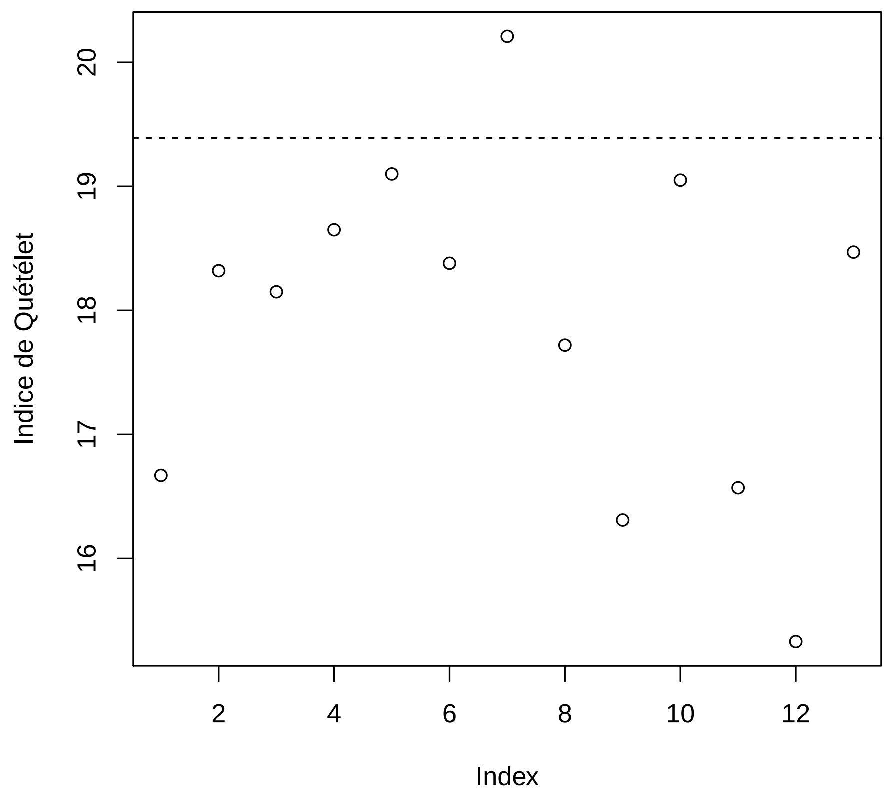
2 Exemples et exercices de tests de comparaison de moyennes

2.1 Indice de Quételet (ou Indice de Masse Corporelle) et Syndrôme de Turner

On a mesuré l'indice de Quételet chez 13 jeunes filles âgées de 14 ans, atteintes du syndrôme de Turner, qui est une monosomie du chromosome X. On a :

```
quetelet <- c(16.67,18.32,18.15,18.65,19.10,18.38,20.21,17.72,
16.31,19.05,16.57,15.33,18.47)
```

1. Calculer la moyenne, la variance et l'écart-type de ces données.
2. Sachant que l'indice moyen de référence pour des jeunes filles du même âge est de 19.39, peut-on conclure que l'indice de Quételet moyen de l'échantillon est identique à l'indice moyen de référence ?
3. Tracez un graphe simple de ces données (`plot`). Ajoutez-y une ligne horizontale à la valeur de référence (`abline`). Ce graphe met-il en évidence un signal en accord avec le résultat du test ?



2.2 Retard de croissance et Indice de Quételet

Précédemment, nous avons comparé la moyenne de l'indice de Quételet chez 13 jeunes filles atteintes du syndrôme de Turner.

```
turner <- quetelet
```

Nous possédons également l'indice de Quételet chez 8 jeunes filles atteintes d'un déficit en hormone de croissance.

```
deficit <- c(16.79,19.22,23,20.52,16.06,20,16.61,17.81)
```

Existe-t-il une différence, en moyenne, de l'indice de Quételet entre ces deux pathologies de croissance ? Cette question se modélise par un test, dont les hypothèses sont les suivantes :

H_0 : l'indice de Quételet, en moyenne, *est identique* chez les personnes souffrant d'une ou l'autre pathologie.

H_1 : l'indice de Quételet, en moyenne, *est différent* chez les personnes souffrant en fonction de la pathologie dont elles souffrent.

Avant de pouvoir le vérifier, il nous faut savoir si les variances des deux populations sont identiques ou non. Pour cela, on ne calcule pas simplement les variances (`var`), mais on effectue un test préliminaire. Ces hypothèses sont :

H'_0 : les variances des indices de Quételet, *sont identiques* chez les personnes souffrant d'une ou l'autre pathologie.

H'_1 : les variances des indices de Quételet, *sont différentes* chez les personnes souffrant d'une ou l'autre pathologie.

```
var.test(turner,deficit)


        F test to compare two variances

data:  turner and deficit
F = 0.3232, num df = 12, denom df = 7, p-value = 0.08294
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.06927673 1.16574614
sample estimates:
ratio of variances
         0.3232334
```

Ici, la p-valeur étant supérieure à 5%, on va conserver H_0 et considérer les variances égales ; on fait donc le test de comparaison de deux moyennes ainsi :

```
t.test(turner,deficit,var.eq=T)


        Two Sample t-test

data:  turner and deficit
t = -1.0317, df = 19, p-value = 0.3151
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.524538  0.857423
sample estimates:
mean of x mean of y
 17.91769  18.75125
```

On voit qu'on ne peut pas conclure à une différence des indices de Quételet en fonction de la pathologie, ni au niveau des variances, ni à celui des moyennes. Si on avait conclu que les variances étaient différentes, la commande pour le test de comparaison de moyennes aurait simplement été :

```
t.test(turner,deficit,var.eq=F)
```

Notez que, même si les formats des réponses sont très proches, R ne fait pas le même test dans les deux cas : quand les variances sont égales il effectue un test de Student, tandis que quand elles sont différentes il effectue un test de Welch – qui est une généralisation du test de Student, moins puissante mais moins contraignante.

2.3 Dépense énergétique et poids

Les données contiennent les dépenses énergétiques sur une période de 24 heures (en MJ) en fonction du statut pondéral obèse (`obese`) ou mince (`lean`) de 22 femmes. Peut-on mettre en évidence une différence de dépense énergétique selon le statut pondéral ?

```
depenergie <- c(9.21,7.53,7.48,8.08,8.09,10.15,8.40,10.88,6.13,7.90,11.51,
                12.79,7.05,11.85,9.97,7.48,8.79,9.69,9.68,7.58,9.19,8.11)
statpond <- c("obese","lean","lean","lean","lean","lean","lean","lean","lean",
              "lean","obese","obese","lean","obese","obese","lean","obese",
              "obese","obese","lean","obese","lean")
```

2.4 Tests univariés

Reprenons l'exemple de la partie précédente (partie 2.3). La procédure que vous avons employée teste pour *n'importe quelle différence* entre les dépenses énergétiques chez les femmes minces ou obèses. Supposons maintenant que nous sachions à l'avance, sans avoir vu les données, que les femmes obèses devraient avoir des dépenses énergétiques plus importantes, parce que leur surcharge pondérale oblige leur organisme à dépenser plus pour des fonctions de mobilité de base. Dans ce cas, on peut réaliser un test unilatéral, qui aura pour hypothèses :

H_0 : les dépenses énergétiques des gens obèses sont *inférieures ou égales* à celles des gens minces.

H_1 : les dépenses énergétiques des gens obèses sont *supérieures* à celles des gens minces.

La procédure statistique pour tester cela consiste à ajouter une option au test de Student (après avoir vérifié l'égalité des variances comme vous avez du le faire précédemment) :

```
t.test(depenergie[statpond=="obese"],depenergie[statpond=="lean"],
var.equal=T,alternative="greater")
```

```
        Two Sample t-test

data:  depenergie[statpond == "obese"] and depenergie[statpond == "lean"]
t = 3.9456, df = 20, p-value = 0.0003995
```

```
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 1.256118      Inf
sample estimates:
mean of x mean of y
10.297778  8.066154
```

Remarquez la manière dont l'hypothèse alternative etl'intervalle de confiance sont donnés. Que concluez-vous ?

Supposons qu'un autre chercheur fasse l'hypothèse que les femmes obèses ont des dépenses énergétiques moindres que les personnes minces, car la régulation thermique est moins coûteuse chez elles. Avec les mêmes données, faites le test correspondant et concluez.

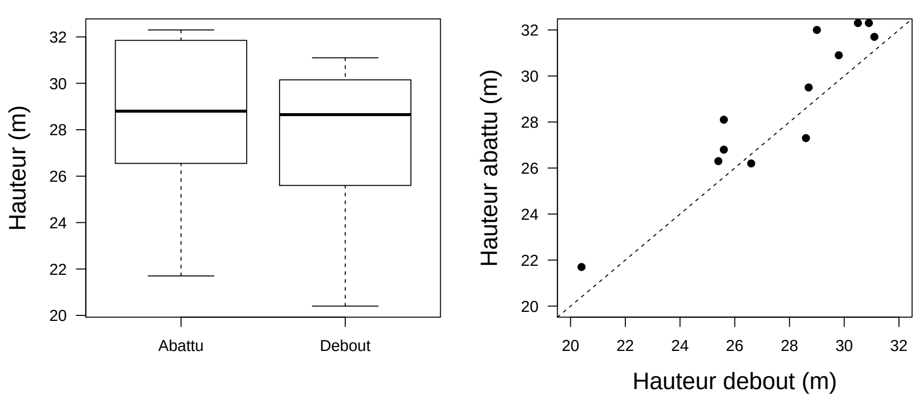
Notez que cette démarche peut s'appliquer à toutes les procédures de tests déjà vues et à venir.

2.5 Tests de données appariées

Les données appariées sont un cas un peu particulier, dans lequel le test compare des données qui ne sont pas indépendantes, mais qui sont mesurées sur les mêmes individus. On veut donc distinguer la variabilité globale inter-individus avec la variabilité au sein de chaque individu. Prenons l'exemplde la mesure de hauteurs d'arbres, avant et après abattage :

```
debout = c(20.4,25.4,25.6,25.6,26.6,28.6,28.7,29.0,29.8,30.5,30.9,31.1)
abattu = c(21.7,26.3,26.8,28.1,26.2,27.3,29.5,32.0,30.9,32.3,32.3,31.7)
```

Il s'agit bien des mêmes arbres. Comme d'habitude, traçons un graphe pour bien comprendre ces données (pourriez-vous retracer les graphes ci-dessous ?).



Sur la boîte à moustaches, on perd l'information "qui est qui", et on peut dire que, globalement, les arbres abattus ont l'air un peu plus grands que ceux mesurés debout. Sur le graphe à droite, par contre, on conserve cette information, et la ligne pointillée nous indique où devraient se trouver les points si les hauteurs mesurées étaient les mêmes : le fait que les points se trouvent presque tous du même coté indique que *chaque* arbre a une hauteur supérieure mesurée abattu que mesuré debout.

Les tests "appariés" et "non appariés" ne vérifient pas la même chose, et donnent des résultats très différents. Le test non apparié examine les hypothèses :

H_0 : les hauteurs de deux populations d'abres, mesurées debout ou abattu sont *les mêmes*.

H_1 : les hauteurs de deux populations d'abres, mesurées debout ou abattu sont *différentes*.

Notez que l'hypothèse alternative n'a biologiquement pas grand sens ici, quand on oublie l'information cruciale que les arbres sont les mêmes dans les deux échantillons.

```
t.test(debout,abattu)
```

```

        Welch Two Sample t-test

data:  debout and abattu
t = -0.8246, df = 21.937, p-value = 0.4185
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -3.779124  1.629124
sample estimates:
mean of x mean of y
 27.68333  28.75833
```

Le test apparié étudie les hypothèses :

H_0 : la hauteur de chaque arbre mesuré debout ou abattu est *la même*.

H_1 : la hauteur de chaque arbre mesuré debout ou abattu est *différente*.

```
t.test(debout, abattu,paired=T)
```

```

        Paired t-test

data:  debout and abattu
t = -3.2343, df = 11, p-value = 0.007954
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.8065536 -0.3434464
sample estimates:
mean of the differences
                 -1.075
```

Concluez biologiquement dans ce cas. Attention, ici, pour simplifier l'exposé on n'a pas vérifié l'égalité des variances avant de faire le test du t de Student, mais il faudrait le faire...

Notez que cette démarche peut s'appliquer à toutes les procédures de tests déjà vues et à venir.

3 Exemples et exercices de comparaison de proportions sous

3.1 Latéralisation et sport d'opposition

Les études statistiques portant sur la latéralisation permettent d'estimer à environ 10% de la population la proportion d'individus qui, dans nos sociétés, utilisent préférentiellement leur main gauche dans les tâches motrices et dans l'écriture en particulier. Parmi les meilleurs tennismen et escrimeurs mondiaux, on a dénombré 18 gauchers sur 64. Peut-on admettre que les gauchers sont plus nombreux dans ces sports d'opposition que dans la population totale ? Il existe une procédure dans pour comparer une proportion observée à une proportion théorique.

```
prop.test(18,64,0.1)
```

```
	1-sample proportions test with continuity correction

data:  18 out of 64, null probability 0.1
X-squared = 21.3906, df = 1, p-value = 3.746e-06
alternative hypothesis: true p is not equal to 0.1
95 percent confidence interval:
 0.1794298 0.4095387
sample estimates:
      p
0.28125
```

La fonction donne la proportion observée dans l'échantillon (`sample estimates`).

```
18/64
```

```
[1] 0.28125
```

3.2 Etudiants du campus de la Doua et groupes sanguins

Les centres de transfusion sanguine ont diffusé le tableau des répartitions, en France, des principaux groupes sanguins. La proportion de O^-et O^+ est de 0.44. Une collecte de sang est organisée sur le campus de la Doua. Sur 356 étudiants, 148 sont de type O. Quelle conclusion pouvez-vous tirer de cette étude ?

3.3 Tabac et Cancer

On a suivi, sur une période de 20 ans, deux cohortes : 200 sujets fumeurs et 200 sujets non fumeurs. On a noté le nombre d'apparition de cancer dans chacune des cohortes : 40 chez les fumeurs ; 20 chez les non fumeurs. La différence d'apparition de cancer dans les deux cohortes est-elle significative ?
La procédure dans pour comparer deux proportions observées est :

```
prop.test(c(40,20),c(200,200))
```

```
        2-sample test for equality of proportions with continuity correction

data:  c(40, 20) out of c(200, 200)
X-squared = 7.0784, df = 1, p-value = 0.007802
alternative hypothesis: two.sided
95 percent confidence interval:
 0.02570481 0.17429519
sample estimates:
prop 1 prop 2
   0.2    0.1
```

3.4 Lecture et Apprentissage

Deux institutrices spécialisées désirent comparer deux méthodes d'apprentissage de lecture différentes M1 et M2. Elles choisissent chacune une méthode qu'elles appliquent dans leurs classes respectives. Pour la méthode M1, sur 25 enfants, l'institutrice compte 60 % d'enfants ayant bien appris à lire ; pour la méthode M2, sur 20 enfants, on en a compté 50%. Que peut-on conclure sur l'efficacité de ces deux méthodes ?

Que peut-on conclure pour ces donneés ?

4 Et si on veut des intervalles de confiance ?

Au lieu de tester une valeur par rapport à une référence, on peut vouloir mesurer un intervalle de confiance sur la moyenne ou sur une proportion, c-à-d l'intervalle dans lequel la vraie moyenne (ou la vraie proportion) de la population se situe avec de grandes chances – ces chances étant choisies par l'utilisateur.

Il n'existe pas de fonction particulière en R pour calculer des intervalles de confiance, parce que les fonctions de test le font déjà. Par exemple, en reprenant l'exemple de la partie 1 :

```
t.test(poids)
```

```
        One Sample t-test

data:  poids
t = 18.2217, df = 14, p-value = 3.785e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 3.980326 5.042341
sample estimates:
mean of x
 4.511333
```

On obtient l'intervalle à 95% de confiance de la moyenne du poids de la population des chats des Ardennes : $[3.98, 5.04]$. Pour obtenir un intervalle avec un niveau de confiance différent, il faut le préciser dans la commande du test :

```
t.test(poids,conf.level=0.98)
```

```

        One Sample t-test

data:  poids
t = 18.2217, df = 14, p-value = 3.785e-11
alternative hypothesis: true mean is not equal to 0
98 percent confidence interval:
 3.861560 5.161107
sample estimates:
mean of x
 4.511333
```

Notez que l'intervalle de confiance s'élargit bien quand on veut être "plus sûr" de ce que l'on dit : en statistiques, moins on veut prendre de risques, moins on est précis. Notez également que dans ces commandes, le résultat contient une p-valeur : c'est la p-valeur du test avec ses *paramètres par défaut*, ici (voir l'aide) la comparaison à une moyenne nulle. Cette valeur ne nous intéresse pas si on veut calculer un intervalle de confiance.

4.1 Taille des basketteurs américains

On connait la taille (en m) de 12 basketteurs américains (in *Mondial Basket*, juillet-août 1994).

```
taibask <- c(2.08, 2.01, 2.03, 2.10, 1.98, 2.08, 1.85, 2.03, 2.16, 2.01, 1.91, 1.88)
```

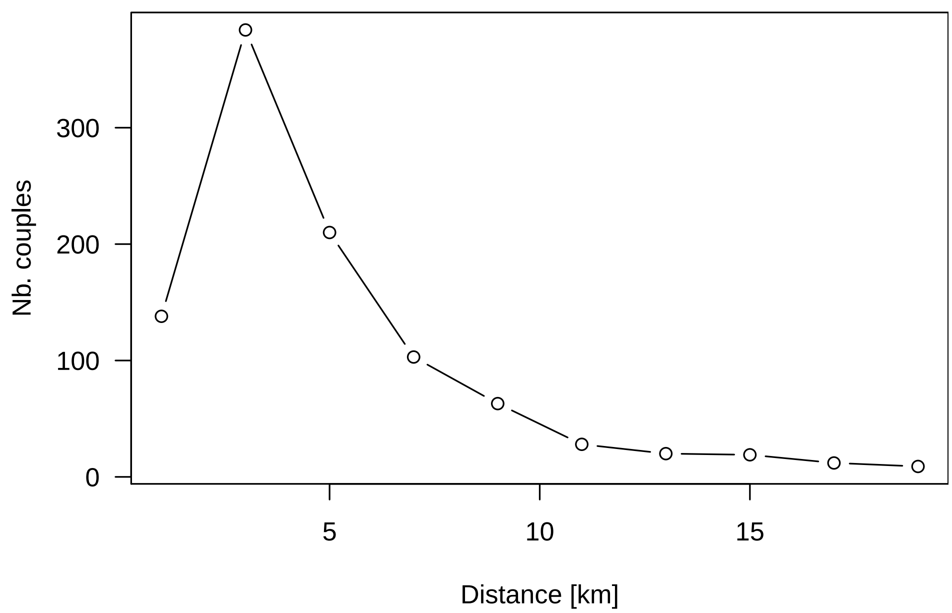
1. Donner la moyenne, la variance estimées de la population à partir de l'échantillon.
2. Donner les intervalles de confiance de la moyenne de la taille des basketteurs américains aux niveaux de confiance 0.95 et 0.99.

4.2 Distance entre domiciles des époux

Une enquête concernant les distances (en km) entre domiciles des époux au moment du mariage a donné, dans le Finistère, les résultats suivants.

Distances	Nombre de couples
0-2	138
2-4	384
4-6	210
6-8	103
8-10	63
10-12	28
12-14	20
14-16	19
16-18	12
18-20	9

1. Représenter graphiquement les données (exemple ci-dessous) et commenter.



2. Calculer la moyenne et la variance estimées de la distance entre les domiciles des époux au moment du mariage.
3. Donner l'intervalle de confiance de la moyenne de la population au niveau de confiance 0.95.

4.3 Présence d'une personne handicapée dans l'entourage

Après une enquête sur un échantillon de 45 étudiants de licence et maîtrise APA, on a constaté que 29 d'entre eux avaient, dans leur entourage, une personne présentant un handicap. Estimer, par intervalle de confiance au niveau 0.9, la proportion d'étudiants ayant, dans leur entourage, une personne handicapée.

5 Et les conditions d'application dans tout ca ?

Pour tous les tests précédents, la seule fois où nous avons vérifié les conditions d'application d'un test, c'est pour choisir le `t.test` adapté en fonction du fait que la variance des deux populations desquelles provenaient les deux échantillons à comparer étaient égales. En pratique, pour réaliser ces tests, il faut que certaines conditions soient respectées.

5.1 Indépendance

Quand on compare deux échantillons, les donnés à comparer doivent être indépendantes – sauf dans le cas des données appariées. L'indépendance signifie qu'aucun individu ne sert dans plusieurs échantillons à la fois ; le cas particulier des données appariées est celui où l'on effectue différentes mesures sur un même individu, pour les comparer ensuite. Cette hypothèse n'est pas vérifiable *a posteriori* : il faut que, dans le choix du plan expérimental, les individus soient bien attribués à un seul traitement ou effet pour que l'on puisse analyser les données correctement par la suite.

5.2 Taille de l'échantillon

On verra ci-dessous que la taille de l'échantillon n'est pas une condition absolue pour les tests de comparaisons de moyennes, même si elle a une influence. Par contre, dans le cas des tests de comparaison de proportion (à une valeur de référence ou entre elles), il est nécessaire que l'échantillon soit assez grand pour que l'on puisse effectuer le test tel quel ; si ce n'est pas le cas, on effectuera la même démarche statistique, mais avec un test non paramétrique nommé test de Fisher – qui sera développé dans la feuille de TP suivante.

5.3 Homoscédasticité

L'homoscédasticité est la condition selon laquelle, pour pouvoir comparer des moyennes, il faut que les variances des populations desquelles proviennent les échantillons soient égales – sans quoi on pourrait confondre différence entre les variances et différences entre les moyennes. Il a été montré plus haut comment tester l'égalité de deux variances ; ceci sera *toujours* à faire avant de tester l'égalité de deux moyennes avec un `t.test`. Dans le cas d'un test comparant des proportions, ou plus généralement des comptes, cette vérification n'a pas de sens.

Un point important quand on vérifie l'égalité des variances par `var.test` est le fait que le test de vérification est lui-même peu puissant, et suppose par défaut (voir H'_0) que les variances sont égales, ce qui est arrangeant pour l'analyste. Un problème qui pourrait survenir est celui des faibles effectifs (typiquement moins d'une dizaine de points par échantillon) : dans ce cas, il est possible que le test de comparaison des variances ne permette pas de rejeter H'_0 uniquement à cause du faible effectif, et pas parce que les variances sont très proches. Il est donc recommandé, en particulier pour ces petits échantillons, d'observer les données pour estimer si les variances semblent égales et de ne pas se fier uniquement au résultat du test ; si l'impression visuelle (par exemple avec un boxplot) est que les variances semblent différentes, on pourra prendre l'initiative de les considérer différentes, même si le test semble ne pas être capable de le dire.

5.4 Normalité

Pour utiliser le test du t de Student, il faut également que les données soient réparties normalement dans chaque échantillon, c-à-d. selon une distribution normale. Une façon de le vérifier, avec de gros échantillons, est de faire *pour chacun* un test, dit test de Shapiro-Wilk, dont les hypothèses sont els suivantes :

H_0 : les données *sont* réparties suivant une distribution normale.
H_1 : les données *ne sont pas* réparties suivant une distribution normale.
Ce test, par exemple appliqué aux données de poids des chats des Ardennes, se fait ainsi :

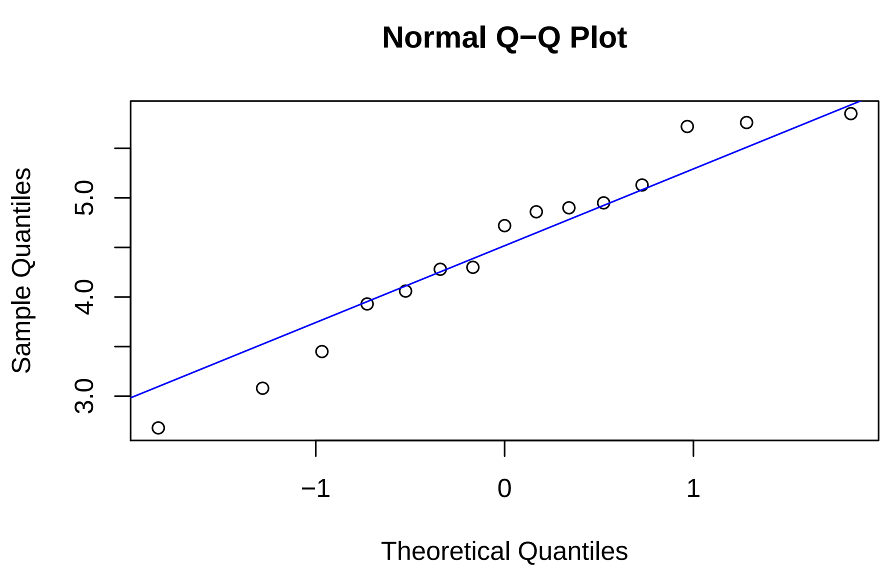
```
shapiro.test(poids)
```

```

        Shapiro-Wilk normality test

data:  poids
W = 0.9568, p-value = 0.6376
```

Que concluez-vous dans ce cas ?

On peut faire pour ce test la même remarque que pour le test de comparaison de variances : si l'effectif des échantillons est trop faible (moins d'une dizaine), le test est peu puissant, et sa conclusion peut être le non-rejet de H_0 simplement parce qu'il n'y a pas assez de données. Une manière visuelle de savoir si un échantillon est distribué normalement pourrait être de tracer son histogramme ; mais si l'échantillon est petit, l'histogramme sera probablement trop grossier pour que l'on puisse conclure. Un graphe qui permet de visualiser facilement si la distribution semble ou non normale est un graphe quantile-quantile :

```
qqnorm(poids)
qqline(poids,col="blue")
```



Sur ce type de graphe, les points devraient être parfaitement alignés sur la droite s'ils suivent une distribution normale. On voit qu'ici c'est presque le cas ; si on observe de gros écarts à cette répartition, en particulier des points sur les bords qui divergent de la droite, on pourra mettre en doute la normalité des données ; en effet ces points correspondent à des valeurs extrêmes trop présentes ou trop peu représentées dans les données.

Que faire si les données ne sont pas normales ? Ne pas paniquer, et tout simplement, aller chercher un autre test que le t de Student, à savoir le *test des rangs de Wilcoxon Mann-Whitney*, un test non-paramétrique. Ce test est moins puissant que le t de Student, mais il fonctionne aussi bien sur des données normales que non-normales ; il convient donc de l'employer dans le cas des données non-normales, ou dans celui des petits effectifs pour lesquels la normalité est discutable. Sa commande est `wilcox.test` ; son application est strictement la même que celle du `t.test`, y compris les options `alternative` et `paired`. L'option `var.equal` n'a aucun intérêt ici, puisque l'on n'a pas besoin de tester cette condition dans le test de Wilcoxon Mann-Whitney.

Un exemple :

```
wilcox.test(poids,mu=5)
```

```
        Wilcoxon signed rank test with continuity correction

data:  poids
V = 29.5, p-value = 0.08834
alternative hypothesis: true location is not equal to 5
```

Vous voyez qu'ici, la conclusion obtenue avec ce test moins puissant n'est pas la même : il vaut dont toujours mieux, si possible, employer un test de Student plutôt qu'un test de Wilcoxon si les conditions d'application sont réunies ; d'une manière générale, il vaut toujours mieux employer un test paramétrique qu'un test non paramétrique (si les conditions d'application sont remplies).

Reprenez les `t.test` que vous avez fait précedemment, et refaites les analyses avec un `wilcox.test` pour bien vérifier le comportement de ce test. Si vous comparez les résultats obtenus, que constatez-vous ?

6 Choisir le bon test

Dans les exercices suivants, vous devez transposer la question biologique posée en un test statistique, et l'appliquer. Attention aux conditions d'applications ! N'hésitez pas à explorer les résultats en regardant si vous vous tromperiez en faisant de "mauvais" tests, et essayez à chaque fois de représenter les données de la manière qui vous semble la plus adaptée.

6.1 Fécondité

On a étudié 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$^{\text{er}}$ au 35$^{\text{ème}}$ jour), à ceux du lot B à partir du 6$^{\text{ème}}$ jour seulement (du 6$^{\text{ème}}$ au 35$^{\text{ème}}$ jour). Les nombre d'œufs pondus par chaque insecte sont les suivantes :

Lot A (13 insectes) : `98, 84, 63, 75, 84, 66, 56, 48, 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.

6.2 Alcool

On a étudié 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 100$^{\text{ème}}$ 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 ?

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

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

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

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