

N.B. Le seuil de signification est par défaut 5 % sauf indication contraire.

Exercice 1 :(3 pts)

Un laboratoire d'agronomie a effectué une étude sur le maintien du pouvoir germinatif des graines de *Papivorus subquaticus* après une conservation de 3 ans. Le laboratoire souhaite vérifier si la probabilité de germination des graines de *Papivorus subquaticus* après trois ans de conservation est égale à p_0 . Il effectue le test d'hypothèse dont la sortie R est la suivante :

```
> prop.test(47, 80, p = NULL,  
+          alternative = "two.sided",  
+          conf.level = 0.95, correct = TRUE)
```

1-sample proportions test with continuity correction

```
data: 47 out of 80, null probability 0.5  
X-squared = 2.1125, df = 1, p-value = 0.1461  
alternative hypothesis: true p is not equal to 0.5  
95 percent confidence interval:  
 0.4718584 0.6946700  
sample estimates:  
      p  
0.5875
```

1. Donner la taille de l'échantillon utilisé pour le test. (0.5 pts)

.....
.....

2. Donner une estimation par intervalle de confiance de la proportion. (0.5 pts)

.....
.....
.....

3. Identifier les hypothèses du test qu'il effectue le laboratoire. (1 pts)

.....
.....
.....
.....

4. Le laboratoire peut-il affirmer que la probabilité de germination des graines de *Papivorus subquaticus* après trois ans de conservation est égale à p_0 . (1 pts)

.....
.....
.....

Exercice 2 :(2.5 pts)

1. On souhaite étudier la qualité de deux races différents des veaux. On mesure le poids à la naissance de 14 veaux du race A et 12 veaux du rache B. Identifier les tests d'hypothèse et interpréter les résultats des cas suivants

a) Test 1 : (1 pts)

```
> var.test(a,b,conf.level=0.95)

      F test to compare two variances

data:  a and b
F = 0.40482, num df = 13, denom df = 11, p-value = 0.1239
alternative hypothesis: true ratio of variances is not equal to 1
```

.....

.....

.....

.....

b) Test 2 : (1 pts)

```
> t.test(a,b,var.equal = T)

      Two Sample t-test

data:  a and b
t = 0.16358, df = 24, p-value = 0.8714
alternative hypothesis: true difference in means is not equal to 0
```

.....

.....

.....

.....

2. (0.5 pts) On souhaite étudier la variabilité d'une variable Y . On effectue deux modèles de régression linéaire. Le premier modèle **M1** en fonction de X_1 et le deuxième **M2** en fonction de X_2 . On note R_1^2 le coefficient de détermination du modèle **M1** et R_2^2 celui de **M2**. On admet le modèle **M1** si on a

- ☐ A) $R_1^2 < R_2^2$ & $AIC(M1) < AIC(M2)$
- ☐ B) $R_1^2 = R_2^2$ & $AIC(M1) > AIC(M2)$
- ☐ C) $R_1^2 = R_2^2$ & $AIC(M1) < AIC(M2)$

Exercice 4 :(5 pts)

On effectue une régression linéaire sur une base de données et on obtient le résultat suivant :

```

Call:
lm(formula = medv ~ chas + nox + rm + age + ptratio + lstat +
    indus + dis, data = Boston)

Residuals:
    Min       1Q   Median       3Q      Max
-15.4059  -2.8264  -0.6814   1.8507  28.3351

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  36.479713    4.581184   7.963 1.15e-14 ***
chas          3.309827    0.886715   3.733 0.000211 ***
nox         -16.881183    3.680144  -4.587 5.70e-06 ***
rm           4.083762    0.415798   9.822 < 2e-16 ***
age          -0.004199    0.013494  -0.311 0.755809
ptratio      -0.966587    0.117658  -8.215 1.85e-15 ***
lstat        -0.559438    0.050422 -11.095 < 2e-16 ***
indus        -0.059159    0.056946  -1.039 0.299373
dis          -1.215437    0.187701  -6.475 2.27e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.937 on 497 degrees of freedom
Multiple R-squared:  0.7164,    Adjusted R-squared:  0.7119
F-statistic: 157 on 8 and 497 DF,  p-value: < 2.2e-16

```

1. **(1 pts)** Identifier la variable à expliquer et les variables explicatives.
.....
.....
.....
.....
2. **(1.5 pts)** Donner les estimations numériques des paramètres du modèle.
.....
.....
.....
.....
3. **(1 pts)** Donner le coefficient de détermination et interpréter la qualité du modèle.
.....
.....
.....
.....
4. **(0.5 pts)** Donner le test d'hypothèses des paramètres du modèle(le test de significativité individuelle).
.....
.....
.....
.....
5. **(1 pts)** Identifier les variables les moins significatives et justifier votre choix.
.....
.....
.....
.....

Exercice 4 :(4 pts)

L'ensemble de données **"ToothGrowth"** inclus dans R décrit l'effet de la vitamine C sur la croissance des dents chez les cobayes. Le fichier contient 3 variables :

- **len** : Longueur des dents
- **dose** : Dose en milligrammes (0.5, 1, 2)
- **supp** : Type de supplément (VC ou OJ)

Des statistiques inférentielles sont utilisées pour déterminer si la longueur des dents est influencée par les doses de vitamine C. On souhaite appliquer l'analyse de la variance à un facteur.

1. (1 pts) Citer les conditions nécessaire d'application de l'Analyse de la variance.

.....

.....

.....

.....

.....

2. (2 pts) Interpréter les outputs R suivants

```
data: X[[i]]  
W = 0.76709, p-value = 1.742e-05
```

```
> tapply(len,dose,shapiro.test)  
$`0.5`
```

Shapiro-Wilk normality test

```
data: X[[i]]  
W = 0.94065, p-value = 0.2466
```

```
$`1`
```

Shapiro-Wilk normality test

```
data: X[[i]]  
W = 0.93134, p-value = 0.1639
```

```
$`2`
```

Shapiro-Wilk normality test

```
data: X[[i]]  
W = 0.97775, p-value = 0.9019
```

Output 1

```
> bartlett.test(len~dose)
```

Bartlett test of homogeneity of variances

```
data: len by dose  
Bartlett's K-squared = 0.66547, df = 2, p-value = 0.717
```

Output 2

Output 2 :

.....

```

Df Sum Sq Mean Sq F value Pr(>F)
dose      1    2224    2224.3    105.1 1.23e-14 ***
Residuals 58    1228     21.2
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

-
- This image shows a full page of a document template designed for handwritten notes or essays. It features approximately 30 evenly spaced, horizontal grey lines across the entire width of the page. The lines are thin and light, providing a guide for writing without being distracting. There are no margins, headers, footers, or other markings present on the page.