

Mini Projet 3 : Greffes d'organes

GUNDUZ Maxime

2023-11-23

1) Créer des récapitulatif statistiques pour chaque organe (mean, median, quartiles, min, max, ...) sous forme de

tableau.

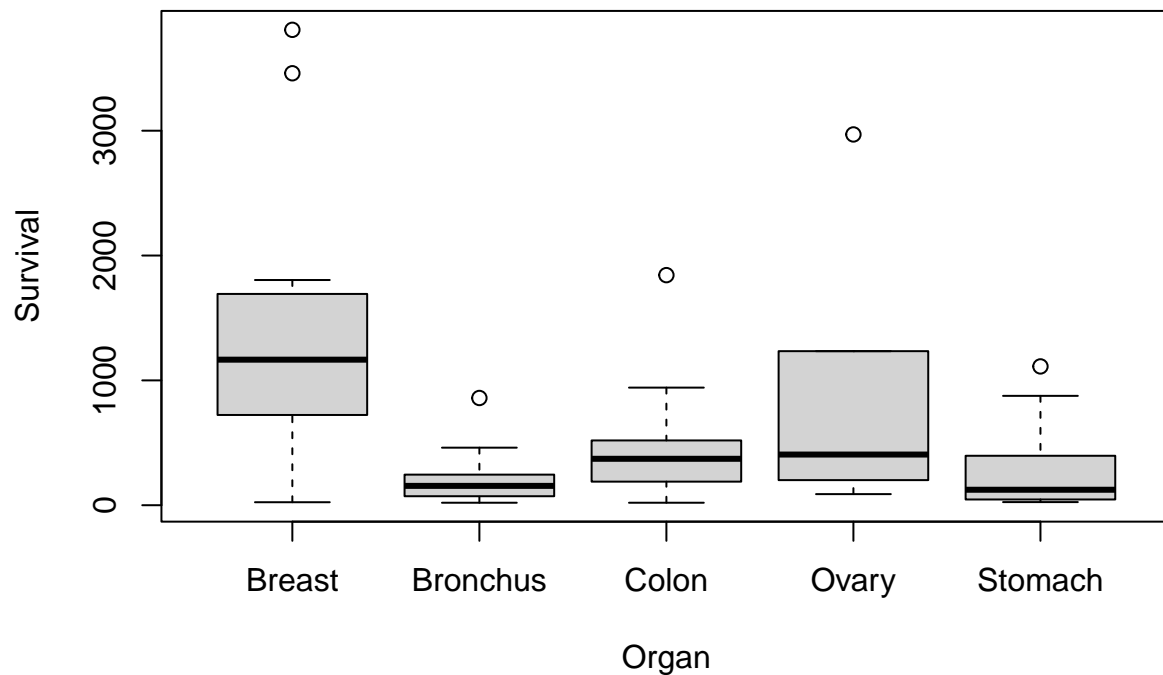
```
summaryTab <- split(tab, Organ)
lapply(summaryTab,summary)
```

```
## $Breast
##      Survival      Organ
##  Min.   : 24      Length:11
## 1st Qu.: 723      Class :character
##  Median:1166      Mode  :character
##   Mean  :1396
## 3rd Qu.:1692
##   Max.  :3808
##
## $Bronchus
##      Survival      Organ
##  Min.   : 20.0      Length:17
## 1st Qu.: 72.0      Class :character
##  Median:155.0      Mode  :character
##   Mean  :211.6
## 3rd Qu.:245.0
##   Max.  :859.0
##
## $Colon
##      Survival      Organ
##  Min.   : 20.0      Length:17
## 1st Qu.: 189.0      Class :character
##  Median: 372.0      Mode  :character
##   Mean  : 457.4
## 3rd Qu.: 519.0
##   Max.  :1843.0
##
## $Ovary
##      Survival      Organ
##  Min.   : 89.0      Length:6
## 1st Qu.: 239.8      Class :character
##  Median: 406.0      Mode  :character
##   Mean  : 884.3
```

```
## 3rd Qu.:1039.5
## Max.    :2970.0
##
## $Stomach
##      Survival      Organ
## Min.   : 25      Length:13
## 1st Qu.: 46      Class :character
## Median :124      Mode  :character
## Mean   : 286
## 3rd Qu.: 396
## Max.   :1112
```

2) Vérifier la présence de valeurs extrêmes avec des boxplots.

```
boites <- boxplot(Survival~Organ)
```



3) Vérifier la normalité et l'égalité des variances

```
# Test de Normalité

t = c("Stomach", "Bronchus", "Colon", "Ovary", "Breast")
```

```
for(i in 1:5){
  print(shapiro.test(Survival[Organ==t[i]]))
}
```

```
##
## Shapiro-Wilk normality test
##
## data: Survival[Organ == t[i]]
## W = 0.75473, p-value = 0.002075
##
##
## Shapiro-Wilk normality test
##
## data: Survival[Organ == t[i]]
## W = 0.76596, p-value = 0.0007186
##
##
## Shapiro-Wilk normality test
##
## data: Survival[Organ == t[i]]
## W = 0.76056, p-value = 0.0006134
##
##
## Shapiro-Wilk normality test
##
## data: Survival[Organ == t[i]]
## W = 0.76688, p-value = 0.029
##
##
## Shapiro-Wilk normality test
##
## data: Survival[Organ == t[i]]
## W = 0.86857, p-value = 0.07431
```

```
# Test de Variance
result <- oneway.test(Survival ~ Organ, data = tab)
print(result)
```

```
##
## One-way analysis of means (not assuming equal variances)
##
## data: Survival and Organ
## F = 3.5152, num df = 4.000, denom df = 19.862, p-value = 0.02514
```

```
p_value <- result$p.value
paste("La p-value du test d'égalité des variances est :", p_value)
```

```
## [1] "La p-value du test d'égalité des variances est : 0.0251366707831244"
```

```

if (p_value < 0.05) {
  paste("La p-value est inférieure à 0.05, l'hypothèse nulle d'égalité des variances est rejeté.")
} else {
  paste("La p-value est supérieure à 0.05, l'hypothèse nulle d'égalité des variances n'est pas rejeté.")
}

```

```
## [1] "La p-value est inférieure à 0.05, l'hypothèse nulle d'égalité des variances est rejeté."
```

4) Transformer les données (log(survival)) et vérifier la normalité et l'égalité des variances

```

# Appliquer le logarithme

tabLog <- tab
tabLog$Survival <- log(tabLog$Survival)

# Test de Normalité

for(i in 1:5){
  print(shapiro.test(tabLog$Survival[tabLog$Organ==t[i]]))
}

```

```

##
##  Shapiro-Wilk normality test
##
## data:  tabLog$Survival[tabLog$Organ == t[i]]
## W = 0.92837, p-value = 0.3245
##
##
##  Shapiro-Wilk normality test
##
## data:  tabLog$Survival[tabLog$Organ == t[i]]
## W = 0.98047, p-value = 0.9613
##
##
##  Shapiro-Wilk normality test
##
## data:  tabLog$Survival[tabLog$Organ == t[i]]
## W = 0.92636, p-value = 0.1891
##
##
##  Shapiro-Wilk normality test
##
## data:  tabLog$Survival[tabLog$Organ == t[i]]
## W = 0.983, p-value = 0.9655
##
##
##  Shapiro-Wilk normality test
##
## data:  tabLog$Survival[tabLog$Organ == t[i]]
## W = 0.802, p-value = 0.009995

```

```
# Test de Variance
resultLog <- oneway.test(tabLog$Survival ~ tabLog$Organ, data = tab)
print(resultLog)

##
## One-way analysis of means (not assuming equal variances)
##
## data: tabLog$Survival and tabLog$Organ
## F = 3.4103, num df = 4.000, denom df = 21.658, p-value = 0.02615

p_valueLog <- resultLog$p.value

paste("La p-value du test d'égalité des variances est :", p_valueLog)

## [1] "La p-value du test d'égalité des variances est : 0.0261465147399908"

if (p_valueLog < 0.05) {
  paste("La p-value est inférieure à 0.05, l'hypothèse nulle d'égalité des variances est rejeté.")
} else {
  paste("La p-value est supérieure à 0.05, l'hypothèse nulle d'égalité des variances n'est pas rejeté.")
}

## [1] "La p-value est inférieure à 0.05, l'hypothèse nulle d'égalité des variances est rejeté."
```

5) Exécuter le test d'ANOVA suivi du test post-hoc de Tukey

```
resAov <- aov(Survival ~ Organ)
resTukey <- TukeyHSD(resAov)

summary(resAov)

##              Df    Sum Sq Mean Sq F value    Pr(>F)
## Organ         4 11535761 2883940    6.433 0.000229 ***
## Residuals    59 26448144  448274
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(resTukey)
```

```
##           Length Class  Mode
## Organ 40      -none- numeric
```

6) Interpréter les résultats

Il y'a une différence significative entre la survie et les organes (0.001). D'après le test de Tukey, ici des différences significatives entre certains groupes d'organes sont présente, et ceci indique des variations spécifiques entre les moyennes de survie.