

Rapport de projet de statistique descriptive et inférentielle utilisant le langage R

Table of contents

1 Exercice 1 :	1
1.0.1 Calculer la moyenne et la variance d'une série statistique des données entrées par un utilisateur	1
1.0.2 Résultat de la première partie du code	2
1.0.3 Résultat du code	3
2 Exercice 2 : Influence de l'alcool sur le temps de réaction au volant	4
3 Fonction de répartition empirique	4
4 Test d'hypothèse pour comparer les deux groupes	6

1 Exercice 1 :

1.0.1 Calculer la moyenne et la variance d'une série statistique des données entrées par un utilisateur

- Première du code :
 - On demande à l'utilisateur de saisir l'effectif de la série statistique
 - On initialise un vecteur ‘numeric’ de la longueur spécifiée
 - On demande de saisir les différents éléments de la série statistique

```

1#=====
2#                               Exercice 1
3#
4#=====
5
6
7# Demander à l'utilisateur la taille de la série
8n <- as.integer(readline("Entrez l'effectif de la série : "))
9
10# Initialiser un vecteur vide
11serie <- numeric(n)
12
13# Lire les éléments un par un
14for (i in 1:n) {
15  serie[i] <- as.numeric(readline(paste("Entrez la valeur numéro", i, ": ")))
16}
17

```

1.0.2 Résultat de la première partie du code

```

> # Demander à l'utilisateur la taille de la série
> n <- as.integer(readline("Entrez l'effectif de la série : "))
Entrez l'effectif de la série : 5
> # Initialiser un vecteur vide
> serie <- numeric(n)
> # Lire les éléments un par un
> for (i in 1:n) {
+   serie[i] <- as.numeric(readline(paste("Entrez la valeur numéro", i, ": ")))
+ }
Entrez la valeur numéro 1 : 0.60
Entrez la valeur numéro 2 : 0.80
Entrez la valeur numéro 3 : 0.78
Entrez la valeur numéro 4 : 0.45
Entrez la valeur numéro 5 : 0.90
>

```

- Seconde partie du code :
 - On calcule la moyenne de la série puis on l'affiche
 - On calcule la variance de l'échantillon ()

```
18
19 |
20 # Calcul de La moyenne
21 moyenne <- mean(serie)
22
23 # Affichage du résultat
24 cat("La moyenne de la série est :", moyenne, "\n")
25
26 # Variance de L'échantillon
27 variance_echantillon = var(serie)
28 cat("La variance de l'échantillon est de: ", variance_echantillon, "\n")
29
30 # Calcul de la variance de la population ( $S^2$ )
31 variance_population <- var(serie) * (n - 1) / n
32 # Affichage de la valeur de la variance de la série statistique
33 cat("La variance de la population est de : ", variance_population, "\n")
34
```

1.0.3 Résultat du code

```
> # Calcul de la moyenne
> moyenne <- mean(serie)
> # Affichage du résultat
> cat("La moyenne de la série est :", moyenne, "\n")
La moyenne de la série est : 0.706
> cat("La variance de l'échantillon est de: ", variance_echantillon, "\n")
La variance de l'échantillon est de: 3.2
> # Calcul de la variance de la population ( $S^2$ )
> variance_population <- var(serie) * (n - 1) / n
> # Affichage de la valeur de la variance de la série statistique
> cat("La variance de la population est de : ", variance_population, "\n")
La variance de la population est de : 0.025744
> |
```

2 Exercice 2 : Influence de l'alcool sur le temps de réaction au volant

```
# Les données
# Le vecteur sans_alcool
sans_alcool <- c(0.68, 0.64, 0.68, 0.82, 0.58, 0.80, 0.72, 0.65, 0.84, 0.73,
                 0.65, 0.59, 0.78, 0.67, 0.65)
# Le vecteur avec_vecteur
avec_alcool <- c(0.73, 0.62, 0.66, 0.92, 0.68, 0.87, 0.77, 0.70, 0.88, 0.79,
                  0.72, 0.60, 0.78, 0.66, 0.68)

ecdf_sans_alcool <- ecdf(sans_alcool)
ecdf_avec_alcool <- ecdf(avec_alcool)

ecdf_sans_alcool
```

```
Empirical CDF
Call: ecdf(sans_alcool)
x[1:12] = 0.58, 0.59, 0.64, ..., 0.82, 0.84
```

```
ecdf_avec_alcool
```

```
Empirical CDF
Call: ecdf(avec_alcool)
x[1:13] = 0.6, 0.62, 0.66, ..., 0.88, 0.92
```

3 Fonction de repartition empirique

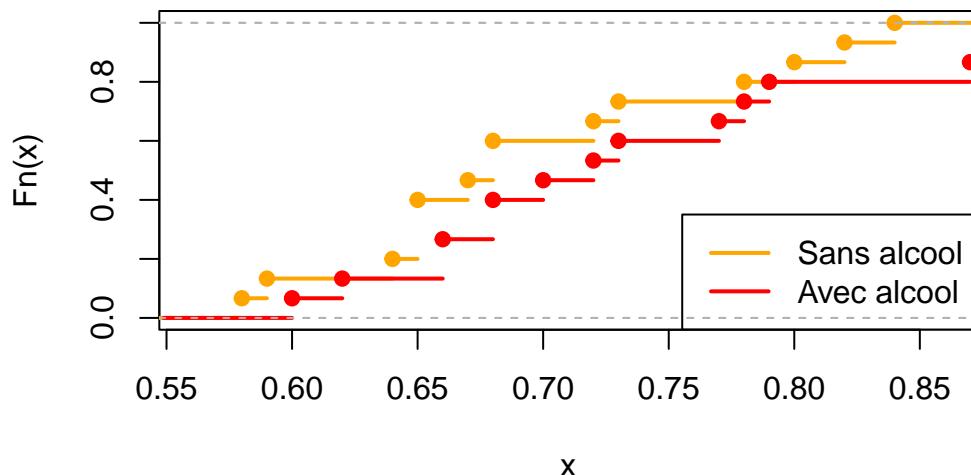
1. Tracer de la fonction de repartition empirique correspondant aux deux situations (utilisation de simple plot())

```
#Tracer les fonctions de répartitions
#temps de reaction sans consommation d'alcool
plot(ecdf_sans_alcool, main = "Fonctions de répartition empirique",
      col = "orange", lwd = 2)

#temps de consommation sans alcool
lines(ecdf_avec_alcool, ,col = "red", lwd = 2)
```

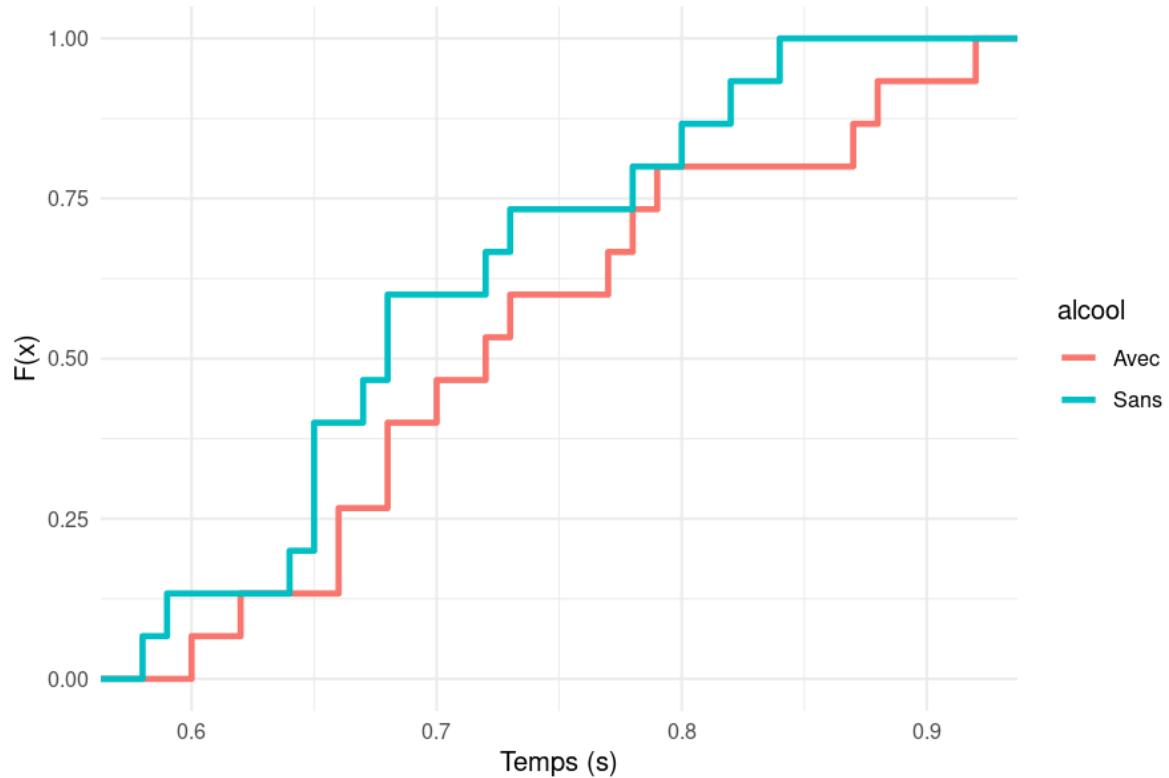
```
#afficher une legende  
legend("bottomright", legend = c("Sans alcool", "Avec alcool"),  
col = c("orange", "red"), lwd = 2)
```

Fonctions de répartition empirique



2. Tracer de la fonction de repartition empirique correspondant aux deux situations (utilisation de simple ggplot2())

Fonctions de répartition du temps de réaction



4 Test d'hypothèse pour comparer les deux groupes

Il s'agit de montrer l'influence de m'alcool sur le temps de réaction au seuil de risque $\alpha = 5\%$ soit un seuil de confiance de 95%. Pour ce faire on utilisera le test de student car les données sont distribuées dans chaque groupe.

1. Formulation des hypothèses - Hypothèse nulle (H_0) :