Atelier 5

Ioos Sterenn

19/11/2019

## Indice de masse corporelle (IMC) chez des enfants

Un échantillon de dossiers d’enfants a ete saisi. Ce sont des enfants vus lors d’une visite en 1er section de maternelle en 1996-1997 dans des ecoles de Bordeaux (Gironde, France). L’échantillon est constitue de 152 enfants ages de 3 ou 4 ans.

### Variables et codage

Vous avez de telecharcher le fichier de le site:

imcenfant <- read.csv2("~/Downloads/Atelier5-master/imcenfant.csv")  
names(imcenfant)

## [1] "SEXE" "zep" "poids" "an" "mois" "taille"

et de répondre a toutes les questions suivantes:

1. Calculer l’IMC de tous les enfants et l’âge et rajoutez une colonne pour l’âge et une autre pour l’ IMC.

imc\_enfant <- imcenfant$poids / (imcenfant$taille/100)^2  
age\_enfant <- imcenfant$an + imcenfant$mois/12  
  
imcenfant <- data.frame(imcenfant, age\_enfant, imc\_enfant)

Commentaire: Voir tableau imcenfant.

1. Extrayez les enfants ayant un IMC < 15 et un age <= 3.5 ans.

enfant\_ex <- imcenfant$SEXE[imcenfant$imc\_enfant <15 & imcenfant$age\_enfant <=3.5]  
table(enfant\_ex)

## enfant\_ex  
## F G   
## 4 4

Commentaire: Il y a 4 garcons et 4 filles qui possèdent un IMC inferieur à 15 et de moins ou égal à 3 ans et demi.

1. Donnez le nombre d’enfants vérifiant les conditions ci-dessus.

length(enfant\_ex)

## [1] 8

Commentaire: Il y en a 8 au total.

1. Est-ce que la plus part des enfants ayant un IMC < 15 et un age <= 3.5 ans sont dans une école située en zone d’éducation prioritaire?

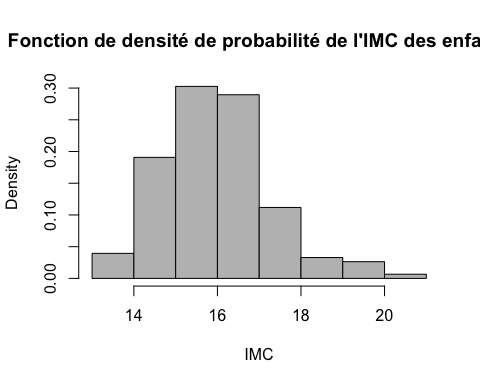
enfant\_ex\_ZEP <- imcenfant$zep[imcenfant$imc\_enfant <15 & imcenfant$age\_enfant <=3.5]  
table(enfant\_ex\_ZEP)

## enfant\_ex\_ZEP  
## N O   
## 3 5

Commentaire: On retrouve 3 qui ne sont pas en ZEP et 5 qui y sont.

1. Représentez la fonction de densité de probabilité de l’IMC chez l’échantillon.

hist(imcenfant$imc\_enfant, freq=FALSE, col="grey", xlab="IMC", main="Fonction de densité de probabilité de l'IMC des enfants")

 Commentaire: voir l’histogramme.

1. Est-ce on peut modelé l’IMC avec une variable aléatoire normal?

6.1. Montrer que la différence entre la médian et la moyenne est très petite et alors on peut supposer que les deux sont égal.

med\_IMC <- median(imcenfant$imc\_enfant)  
moy\_IMC <- mean(imcenfant$imc\_enfant)  
  
med\_IMC

## [1] 15.9838

moy\_IMC

## [1] 16.00498

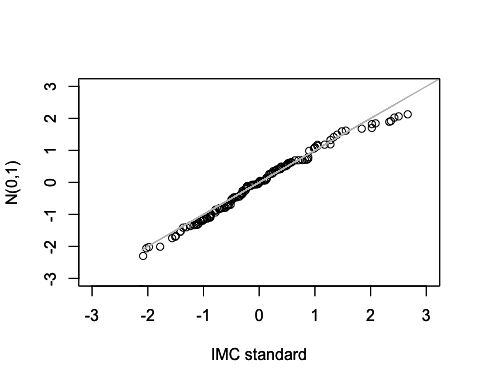
differenceMM <- moy\_IMC - med\_IMC  
differenceMM

## [1] 0.02117762

Commentaire: La différence est de 0,02 donc on peut dire que la médiane est égale à la moyenne.

6.2 Utilisez la fonction pour étudier si l’IMC a le même fonction de distribution que une variable alèatoire normal.

z<- (imcenfant$imc\_enfant - mean(imcenfant$imc\_enfant))/sd(imcenfant$imc\_enfant)  
n <-length(z)  
N <- rnorm(n)  
qqplot(z,N,xlim=c(-3,3),ylim=c(-3,3), xlab="IMC standard", ylab="N(0,1)")  
par(new=TRUE)  
plot(z,z,type="l",col="grey",xlim=c(-3,3),ylim=c(-3,3),xlab="IMC standard", ylab="N(0,1)")

 Commentaire: Voir le graphique.

6.3 Est-qu’on peut quantifier la différence entre les deux fonctions de distributions?

# Utiliser pour repondre la question 6.3

Commentaire:

1. Est-ce que l’IMC chez les enfants dans une école située en zone d’éducation prioritaire est différent de les enfants qui ne sont pas dans une école située en zone d’éducation prioritaire?

7.1 Donnez le résumé statistique de l’IMC chez les enfants dans une école située en zone d’éducation prioritaire.

summary\_IMC\_ZEP <- summary(imcenfant$imc\_enfant[imcenfant$zep=="O"])  
summary\_IMC\_ZEP

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 13.33 15.10 16.00 16.09 16.70 20.29

Commentaire: Le résumé est de : Min. 1st Qu. Median Mean 3rd Qu. Max. 13.33 15.10 16.00 16.09 16.70 20.29

7.2 Donnez le résumé statistique de l’IMC chez les enfants que ne sont pas dans une école située en zone d’éducation prioritaire.

summary\_IMC\_zep <- summary(imcenfant$imc\_enfant[imcenfant$zep=="N"])  
summary\_IMC\_zep

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 13.30 15.30 15.84 15.76 16.50 17.65

Commentaire: Le résumé est de : Min. 1st Qu. Median Mean 3rd Qu. Max. 13.30 15.30 15.84 15.76 16.50 17.65

1. Quelle est la probabilité de trouver une fille dans dans une école située en zone d’éducation prioritaire?

proba\_FZEP <- length(imcenfant$taille[imcenfant$SEXE=="F" & imcenfant$zep=="O"])/length(imcenfant$taille[imcenfant$SEXE=="F"])  
proba\_FZEP

## [1] 0.7464789

Commentaire: La probabilité d’avoir une fille dans une école ZEP est de 0,746.

1. Est-ce qu’on peut modelé la taille des enfants avec un variable aléatoire normal?

med\_IMC <- median(imcenfant$taille)  
moy\_IMC <- mean(imcenfant$taille)  
  
med\_IMC

## [1] 101

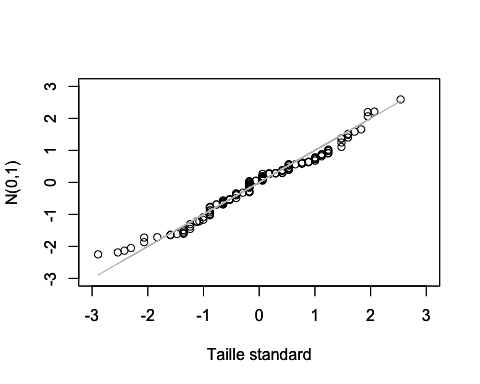
moy\_IMC

## [1] 100.748

differenceMM <- moy\_IMC - med\_IMC  
differenceMM

## [1] -0.2519737

z<- (imcenfant$taille - mean(imcenfant$taille))/sd(imcenfant$taille)  
n <-length(z)  
N <- rnorm(n)  
qqplot(z,N,xlim=c(-3,3),ylim=c(-3,3), xlab="Taille standard", ylab="N(0,1)")  
par(new=TRUE)  
plot(z,z,type="l",col="grey",xlim=c(-3,3),ylim=c(-3,3),xlab="Taille standard", ylab="N(0,1)")

 Commentaire: La différence entre la moyenne et la médiane est très faible donc on peut comparer avec une variable aléatoire normale. Alors on peut observer le graphique.

1. Est-ce qu’on peut modelé le poids des enfants avec un variable aléatoire normal?

med\_IMC <- median(imcenfant$poids)  
moy\_IMC <- mean(imcenfant$poids)  
  
med\_IMC

## [1] 16

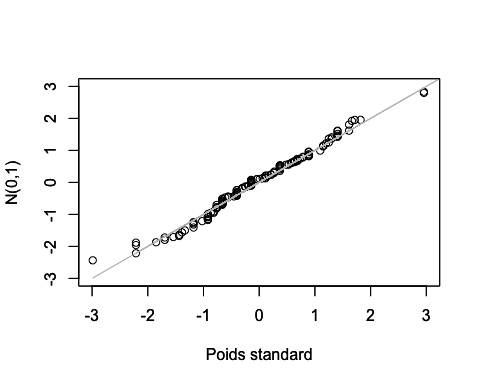
moy\_IMC

## [1] 16.28026

differenceMM <- moy\_IMC - med\_IMC  
differenceMM

## [1] 0.2802632

z<- (imcenfant$poids - mean(imcenfant$poids))/sd(imcenfant$poids)  
n <-length(z)  
N <- rnorm(n)  
qqplot(z,N,xlim=c(-3,3),ylim=c(-3,3), xlab="Poids standard", ylab="N(0,1)")  
par(new=TRUE)  
plot(z,z,type="l",col="grey",xlim=c(-3,3),ylim=c(-3,3),xlab="Poids standard", ylab="N(0,1)")



Commentaire: La différence entre la moyenne et la médiane est très faible donc on peut comparer avec une variable aléatoire normale. Alors on peut observer le graphique.