Modélisation: premières manipulations

Commentaires sur les versions de R, RStudio versions et R notebook

- RStudio Version 1.1.383 with a R version 3.4.3.
- Ceci est un document R Markdown documen qui intercale du code R et du texte. Permet de faire un rapport. Le code R est situ'e entre des "chunks"
- Vous pouvez ex'ecuter ces "chunks" en cliquant sur le bouton *Run* 'a l int'erieur d un chunk ou en se plaçant 'a l'interieur du chunk *Ctrl+Alt+Enter* ou *Cmd+Shift+Enter*.
- Ajouter un nouveau chunk en cliquant sur *Insert Chunk* dans la barre d'outil ou en cliquant sur *Ctrl+Alt+I* or *Cmd+Option+I*.
- Si vous cliquez sur le bouton **Knit** cela va g'en'erer un document html incluant 'a la fois les commentaires et le code R associ'e aux chunks.
- Quelques r'ef'erences de bases 'a l adresse http://rmarkdown.rstudio.com/authoring_basics.html.
- En cas de probl'eme **d'encodage**: aller dans le menu principal et cliquer sur "r'eouvrir avec l'encodage UTF-8".

Commentaires sur l"etude

- Le r'esultat de votre 'etude doit être envoy'e 'a l adresse marie-luce.taupin@univ-ervy.fr
- **Je n'accepte que les documents .html** file (obtenu avec le bouton *Knit*) **pas de .Rmd**

Exercice 1: manipulations de R et Rmarkdown

0. Charger les packages R

```
knitr::opts_chunk$set(fig.height=5, fig.width=7)
library(ggfortify) # for nice plots
```

```
## Loading required package: ggplot2
```

```
library(ggplot2) # for nice plots
library("gridExtra")
```

```
#setwd("~/Users/marieluce.taupin/HOME/WORK/ENSEIGNEMENT/ENS17-18/L2 Math/MODELISATION L2 MATH/Introduct
getwd()
```

```
## [1] "/Users/marieluce.taupin/HOME/WORK/ENSEIGNEMENT/ENS17-18/L2 Math/MODELISATION L2 MATH/Introductio
```

```
dir()
```

```
##  [1] "old"                   "slides0-intro.pdf"
##  [3] "TP Intro R.aux"        "TP Intro R.log"
##  [5] "TP Intro R.pdf"        "TP Intro R.synctex.gz"
##  [7] "TP Intro R.tex"        "TP_intro.pdf"
##  [9] "TP_intro.Rmd"          "TP1_gen.pdf"
## [11] "TP1_gen.r"
```

```
ls()
```

```
## character(0)
```

1. Charger un package

```
library(MASS)
```

2. Charger des donn'ees

```
data()
help(birthwt)
head(birthwt)
```

```
##    low age lwt race smoke ptl ht ui ftv  bwt
## 85   0  19 182    2     0   0  0  1   0 2523
## 86   0  33 155    3     0   0  0  0   3 2551
## 87   0  20 105    1     1   0  0  0   1 2557
## 88   0  21 108    1     1   0  0  1   2 2594
## 89   0  18 107    1     1   0  0  1   0 2600
## 91   0  21 124    3     0   0  0  0   0 2622
```

```
?birthwt
str(birthwt)
```

```
## 'data.frame':    189 obs. of  10 variables:
##  $ low  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ age  : int  19 33 20 21 18 21 22 17 29 26 ...
##  $ lwt  : int  182 155 105 108 107 124 118 103 123 113 ...
##  $ race : int  2 3 1 1 1 3 1 3 1 1 ...
##  $ smoke: int  0 0 1 1 1 0 0 0 1 1 ...
##  $ ptl  : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ ht   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ ui   : int  1 0 0 1 1 0 0 0 0 0 ...
##  $ ftv  : int  0 3 1 2 0 0 1 1 1 0 ...
##  $ bwt  : int  2523 2551 2557 2594 2600 2622 2637 2637 2663 2665 ...
```

```
ls.str(birthwt)
```

```
## age :  int [1:189] 19 33 20 21 18 21 22 17 29 26 ...
## bwt :  int [1:189] 2523 2551 2557 2594 2600 2622 2637 2637 2663 2665 ...
## ftv :  int [1:189] 0 3 1 2 0 0 1 1 1 0 ...
## ht :  int [1:189] 0 0 0 0 0 0 0 0 0 0 ...
## low :  int [1:189] 0 0 0 0 0 0 0 0 0 0 ...
## lwt :  int [1:189] 182 155 105 108 107 124 118 103 123 113 ...
## ptl :  int [1:189] 0 0 0 0 0 0 0 0 0 0 ...
## race :  int [1:189] 2 3 1 1 1 3 1 3 1 1 ...
## smoke :  int [1:189] 0 0 1 1 1 0 0 0 1 1 ...
## ui :  int [1:189] 1 0 0 1 1 0 0 0 0 0 ...
```

```
summary(birthwt)
```

```
##       low              age              lwt             race      
##  Min.   :0.0000   Min.   :14.00   Min.   : 80.0   Min.   :1.000  
##  1st Qu.:0.0000   1st Qu.:19.00   1st Qu.:110.0   1st Qu.:1.000  
##  Median :0.0000   Median :23.00   Median :121.0   Median :1.000  
##  Mean   :0.3122   Mean   :23.24   Mean   :129.8   Mean   :1.847  
##  3rd Qu.:1.0000   3rd Qu.:26.00   3rd Qu.:140.0   3rd Qu.:3.000  
##  Max.   :1.0000   Max.   :45.00   Max.   :250.0   Max.   :3.000  
##      smoke            ptl               ht                ui        
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.0000  
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.0000  
```

```
##  Median :0.0000   Median :0.0000   Median :0.00000   Median :0.0000
##  Mean   :0.3915   Mean   :0.1958   Mean   :0.06349   Mean   :0.1481
##  3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.0000
##  Max.   :1.0000   Max.   :3.0000   Max.   :1.00000   Max.   :1.0000
##       ftv              bwt
##  Min.   :0.0000   Min.   : 709
##  1st Qu.:0.0000   1st Qu.:2414
##  Median :0.0000   Median :2977
##  Mean   :0.7937   Mean   :2945
##  3rd Qu.:1.0000   3rd Qu.:3487
##  Max.   :6.0000   Max.   :4990
```

4. Op'erations 'el'ementaires

```
?mean
help.search("mean")
ls()
```

```
## character(0)
```

```
y<-c(1,2,6,7,0,-1)
print(y)
```

```
## [1]  1  2  6  7  0 -1
```

```
c(5:7,y)
```

```
## [1]  5  6  7  1  2  6  7  0 -1
```

```
a<-1:3
b<-4:6
a+b
```

```
## [1] 5 7 9
```

```
a*b
```

```
## [1]  4 10 18
```

```
a/b
```

```
## [1] 0.25 0.40 0.50
```

```
a+1
```

```
## [1] 2 3 4
```

```
a*2
```

```
## [1] 2 4 6
```

```
1/a
```

```
## [1] 1.0000000 0.5000000 0.3333333
```

```
rm(list=c("a"))
a<-1:3
numeric(5)
```

```
## [1] 0 0 0 0 0
```

```
numeric(5)->a
a[3]<-2
a
```

```
## [1] 0 0 2 0 0
```

```
rep(1,4)
```

```
## [1] 1 1 1 1
```

```
rep(1:3,4)
```

```
##  [1] 1 2 3 1 2 3 1 2 3 1 2 3
```

```
rep(1:3,each=2)
```

```
## [1] 1 1 2 2 3 3
```

```
a<-c(5,3,6,NA)
is.na(a)
```

```
## [1] FALSE FALSE FALSE  TRUE
```

```
!is.na(a)
```

```
## [1]  TRUE  TRUE  TRUE FALSE
```

```
a[!is.na(a)]
```

```
## [1] 5 3 6
```

```
mean(a)
```

```
## [1] NA
```

```
mean(a,na.rm=TRUE)
```

```
## [1] 4.666667
```

```
mean(a[!is.na(a)])
```

```
## [1] 4.666667
```

```
min(a)
```

```
## [1] NA
```

```
max(a)
```

```
## [1] NA
```

```
which.max(a)
```

```
## [1] 3
```

```
sort(a)
```

```
## [1] 3 5 6
```

```
sort(a,decreasing=FALSE)
```

```
## [1] 3 5 6
```

```
3+2
```

```
## [1] 5
```

```
3>2
```

```
## [1] TRUE
```

```r
(3>2)+(5>8)
```

```
## [1] 1
```

```r
n=3
x=c(3.1,1,4,7,9)
y=c(0,1,1,0)
# voir une partie et taille
x[2:4]
```

```
## [1] 1 4 7
```

```r
x[x>3]
```

```
## [1] 3.1 4.0 7.0 9.0
```

```r
x[y==1]
```

```
## [1] 1 4
```

```r
length(x)
```

```
## [1] 5
```

```r
# creation automatique
x=rep(0,1000)
x=seq(1,5000,5)
```

5. Les matrices

```r
# matrices
M=matrix(1,nrow=2,ncol=3)
v=seq(1,6,1)
M1=matrix(v,nrow=2,ncol=3,byrow=TRUE)
M2=matrix(v,nrow=2,ncol=3,byrow=FALSE)
I<-diag(1,nrow=4)
```

6. Encore quelques op'erations

```r
labs<-paste(c("X","Y"),1:10,sep="")
x[1:10]
```

```
##  [1]  1  6 11 16 21 26 31 36 41 46
```

```r
c("x","y")[rep(c(1,2,2,1),times=4)]
```

```
##  [1] "x" "y" "y" "x" "x" "y" "y" "x" "x" "y" "y" "x" "x" "y" "y" "x"
```

```r
y<-x[-(1:5)]
fruit<-c(5,10,1,20)
names(fruit)<-c("orange","banane","pomme","p?che")
lunch<-fruit[c("pomme","orange")]
x<-c(10.4,5.6,3.1,6.4,21.7)
assign("x",c(10.4,5.6,3.1,6.4,21.7))
y<-c(x,0,x)
z<-y[-(1:5)]
v<- 2*x+y+1
```

```
## Warning in 2 * x + y: la taille d'un objet plus long n'est pas multiple de
## la taille d'un objet plus court
```

```r
z<- 0:9
digits<- as.character(z)
d<-as.integer(digits)
x<- array(1:20,dim=c(4,5))
i<- array(c(1:3,3:1),dim=c(3,2))
x[i]
```

```
## [1] 9 6 3
```

```r
x[i]<-0
x
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    5    0   13   17
## [2,]    2    0   10   14   18
## [3,]    0    7   11   15   19
## [4,]    4    8   12   16   20
```

Exo2

1. Créer un vecteur contenant la suite des entiers de 1 à 12 de deux façons différentes.
2. Créer le vecteur $U = (0.5\,1\,1.5\,2\,2.5\,3\,3.5\,4\,4.5\,5)$ de trois manières différentes.
3. Quelle est la longueur de U, la somme des coordonnées de U?

```r
1:12
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12
```

```r
seq(from=1,to=12,by=1)
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12
```

```r
U=seq(from=0.5,to=5,len=10)
U=seq(from=0.5,to=5,by=0.5)
U=1:10/2
length(U)
```

```
## [1] 10
```

```r
sum(U)
```

```
## [1] 27.5
```

Exo3

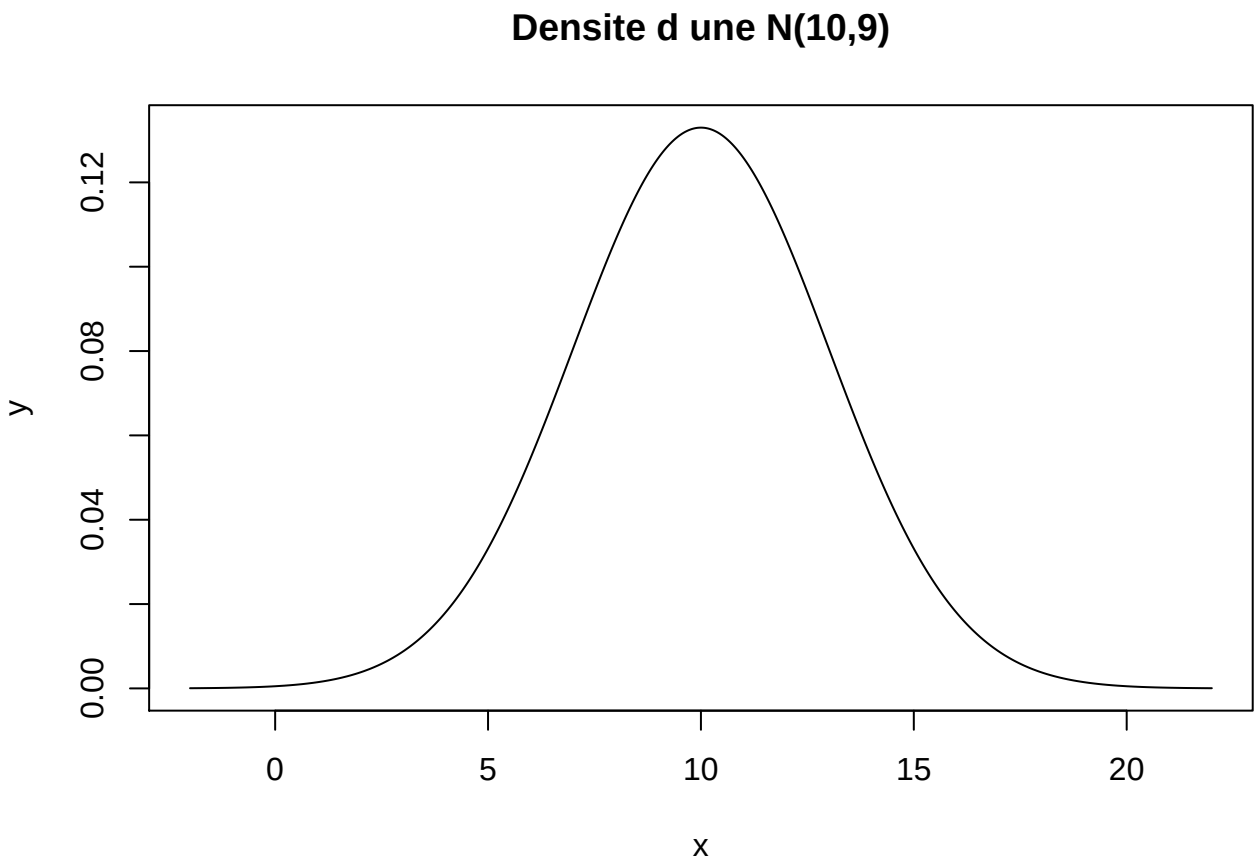
On veut mimer un jeu de pile ou face. On note 1 si le r'esultat obtenu est ''pile''et 0 si c'est''face''. Simuler 1000 lancers de pi'eces. On gagne 1 euro si le r'esultat est''pile''et on perd un euro si c'est''face''. Combien d'argent avez-vous gagn'e 'a l'issue des 1000 lancers.

```r
x<- rbinom(1000,1,0.5)
s=sum(rbinom(1000,1,0.5))
v=sum(rbinom(1000,1,0.5))-(1000-sum(rbinom(1000,1,0.5)))
```

Exo4

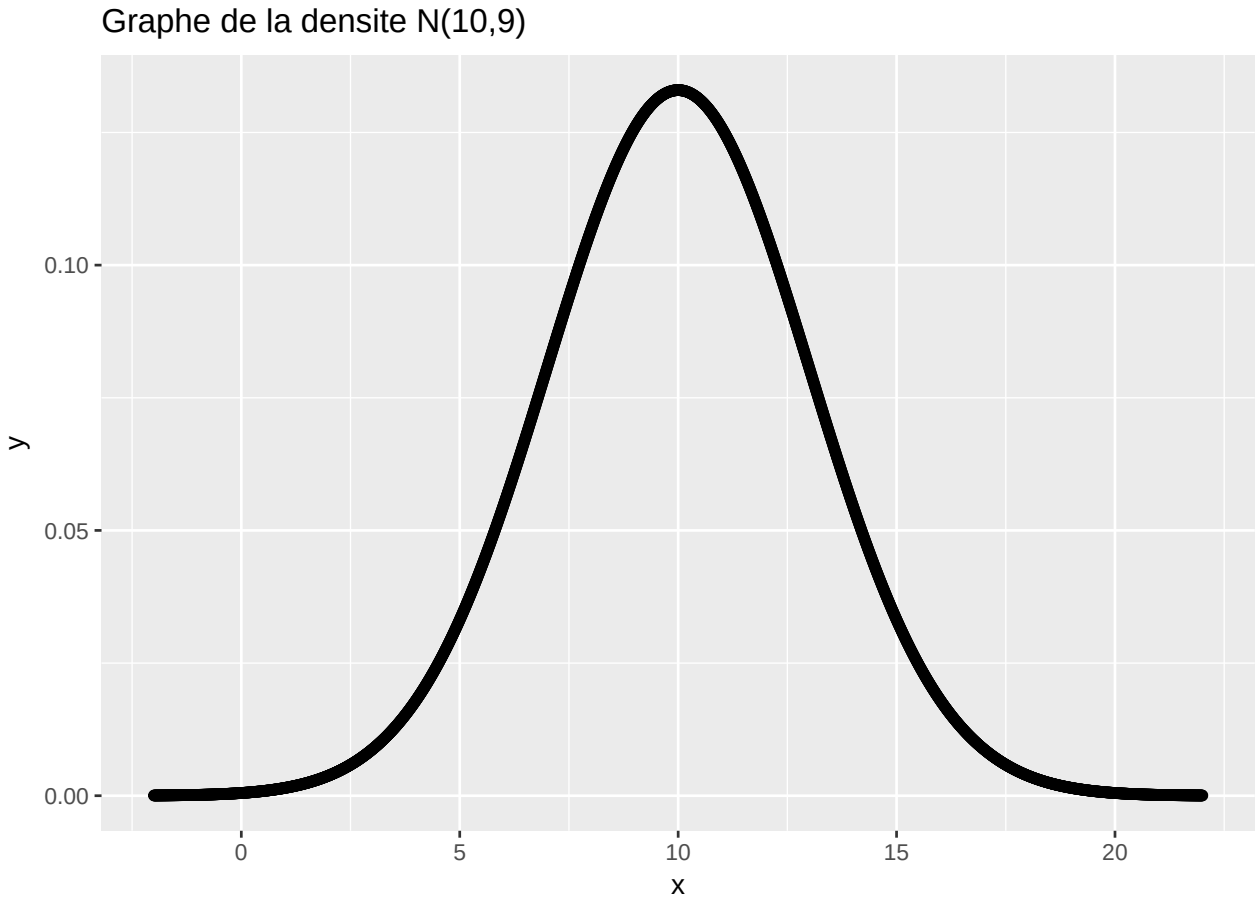
1. Tracer la densité de la gaussienne $\mathcal{N}(0,1)$.
2. Tracer la densité de la gaussienne $\mathcal{N}(10,9)$.
3. Mettre sur un même graphique les graphes des deux densités ci-dessus.
4. Mettre sur une même fenêtre les graphes des densités d'une $\mathcal{X}(10)$ et $\mathcal{F}(11,14)$.
5. Déterminer $\mathbb{P}(|\mathcal{N}(0,1)| \leq 1.96)$.
6. Déterminer x tel que $\mathbb{P}(\mathcal{N}(0,1) \leq x) = 0.90$.
7. Simuler un échantillon de $n = 250$ variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{N}(0,1)$. Tracer leur histogramme.

```r
# Graphe de N(10,9) (Deux m\'ethodes)
x=seq(-2,22,0.01)
y=dnorm(x,mean=10,sd=3)
plot(x,y,type="l",main="Densite d une N(10,9)")
```



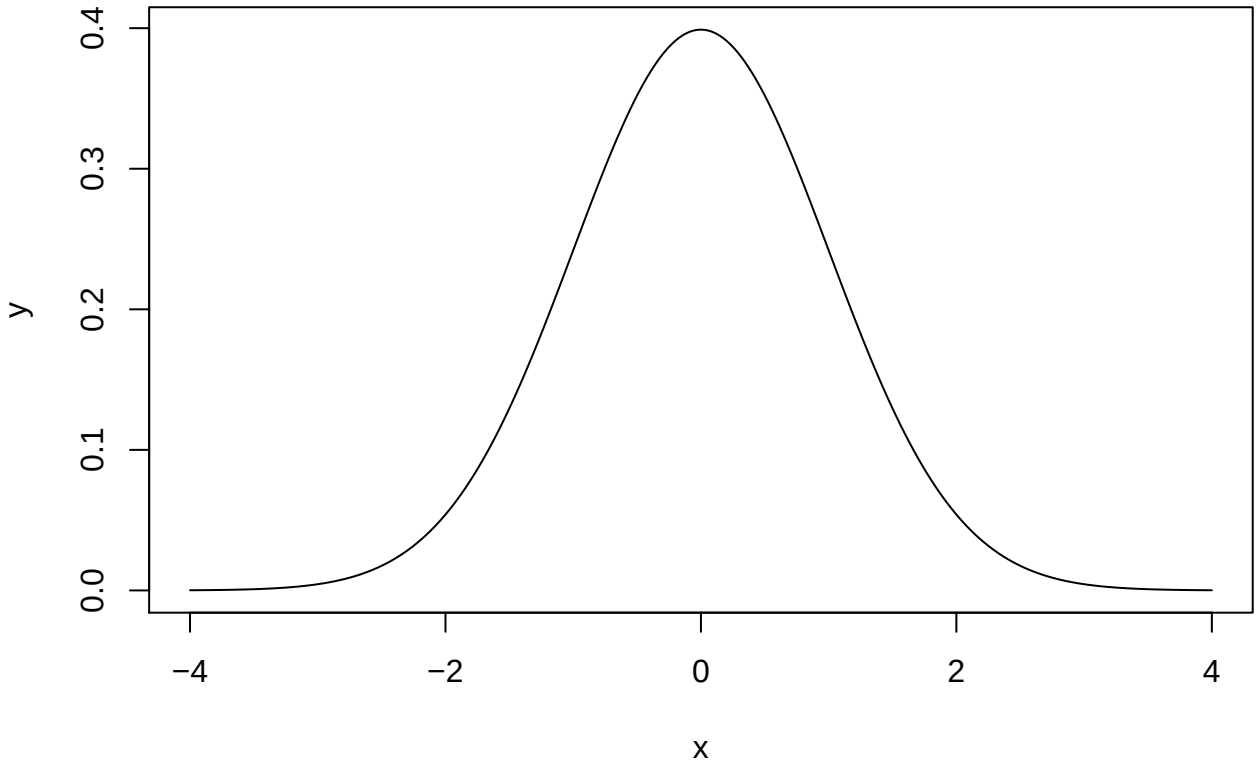
Autre m'ethode

```r
ggplot(data=NULL, aes(x,y)) +
    geom_point()+
    labs(title="Graphe de la densite N(10,9)",x= "x", y = "y")
```



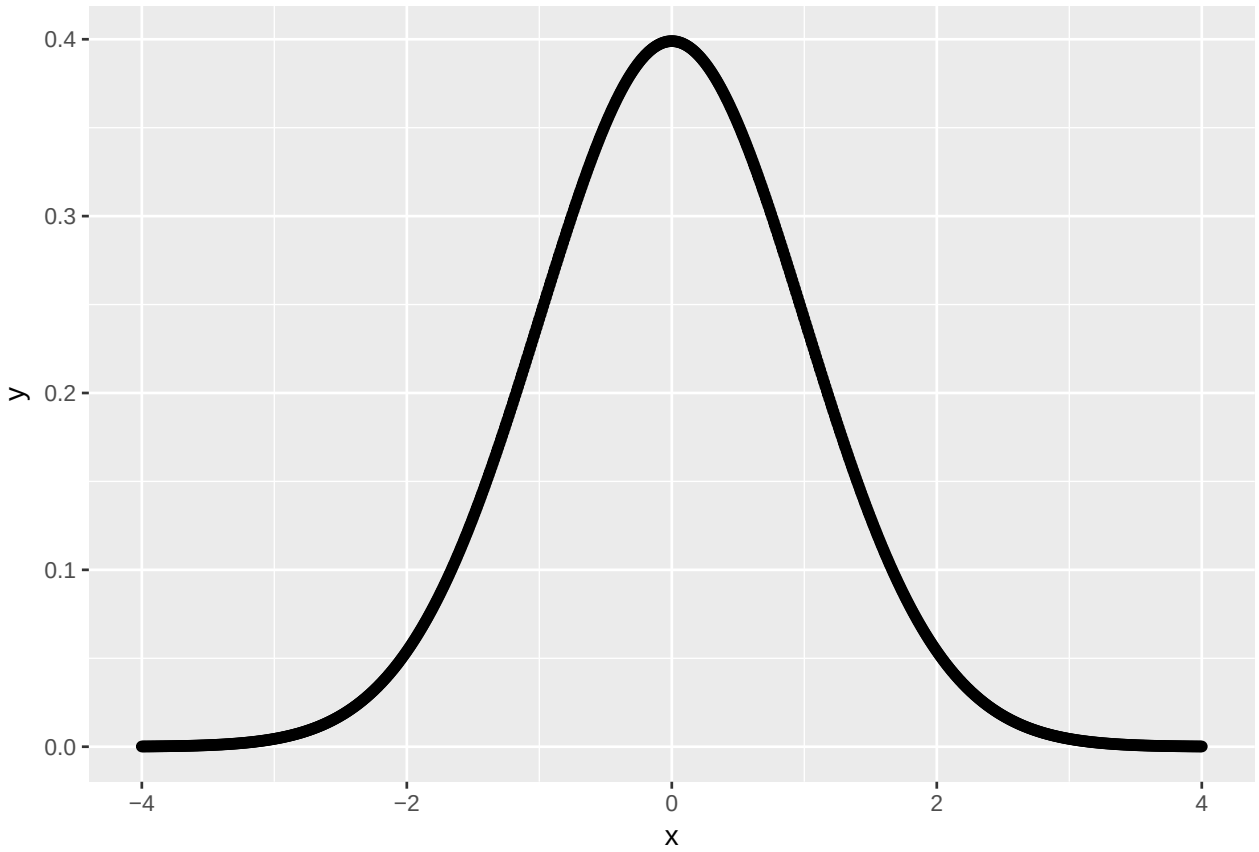
Graphe de la densit'e N(0,1) (Premi'ere m'ethode)

```
x=seq(-4,4,0.01)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",main="Densite de la  N(0,1)")
```



Graphe de la densit'e N(0,1) (Deuxi‘eme m'ethode)

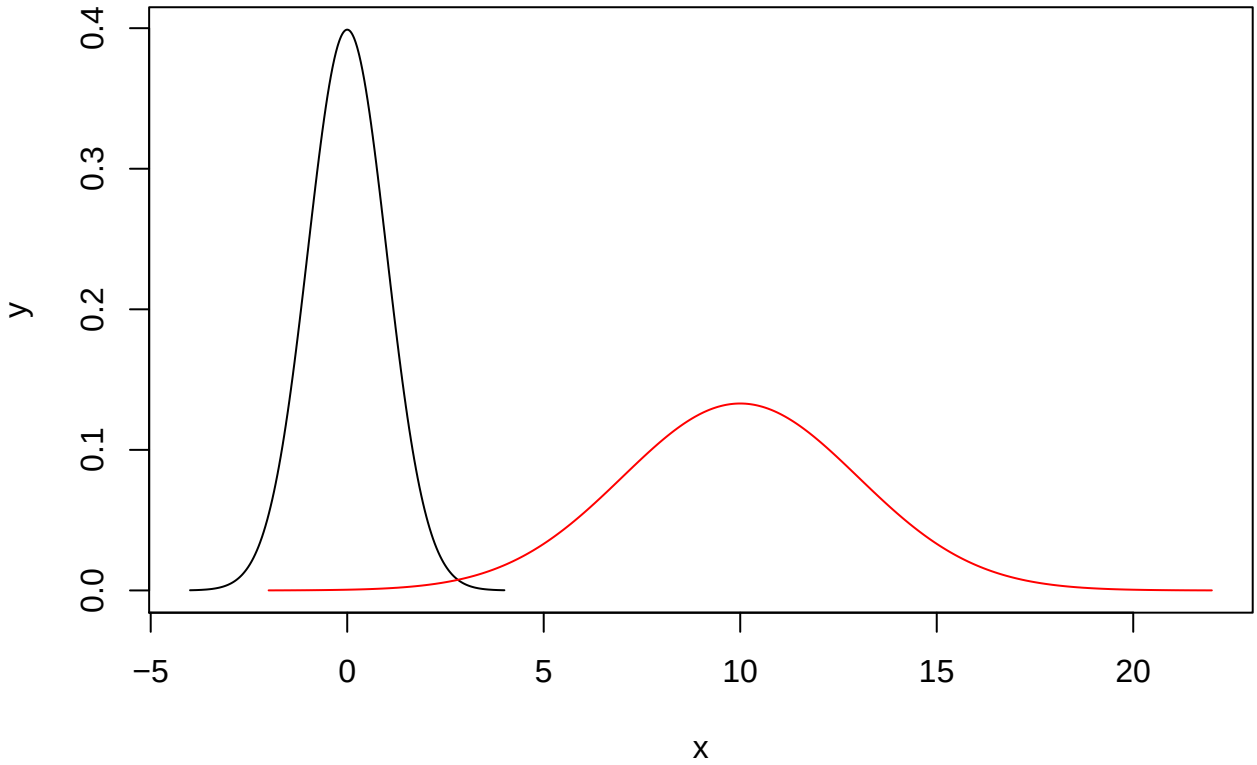
```
x=seq(-4,4,0.01)
y=dnorm(x,mean=0,sd=1)
ggplot(data=NULL, aes(x,y)) +
    geom_point()+
    labs(title="Graphe de la densite N(0,1)",x= "x", y = "y")
```



Graphe des deux densites

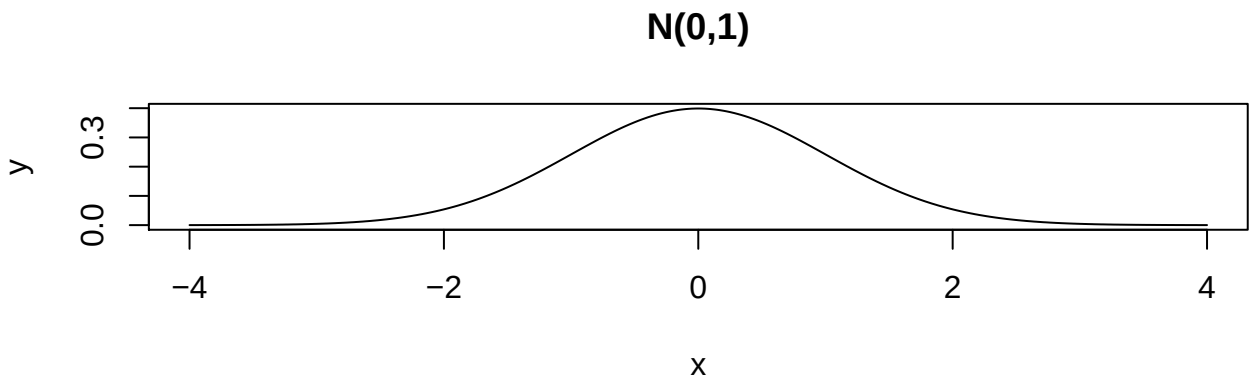
```
x=seq(-4,4,0.01)
y=dnorm(x,mean=0,sd=1)
x1=seq(-2,22,0.01)
y1=dnorm(x1,mean=10,sd=3)
```

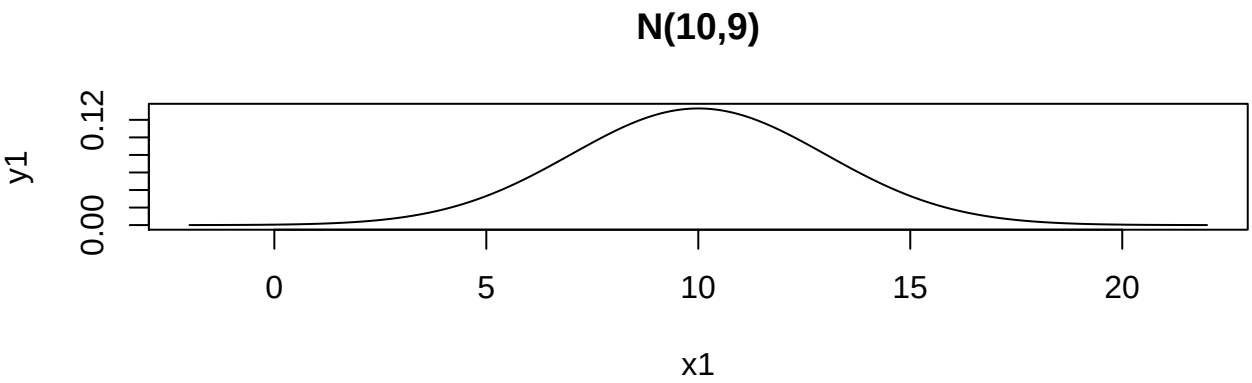
```
plot(x,y,type="l",main="Graphe des densites N(0,1) et N(10,9)",xlim=c(-4,22))
lines(x1,y1,col="red")
```



2 graphes sur une meme fenetre

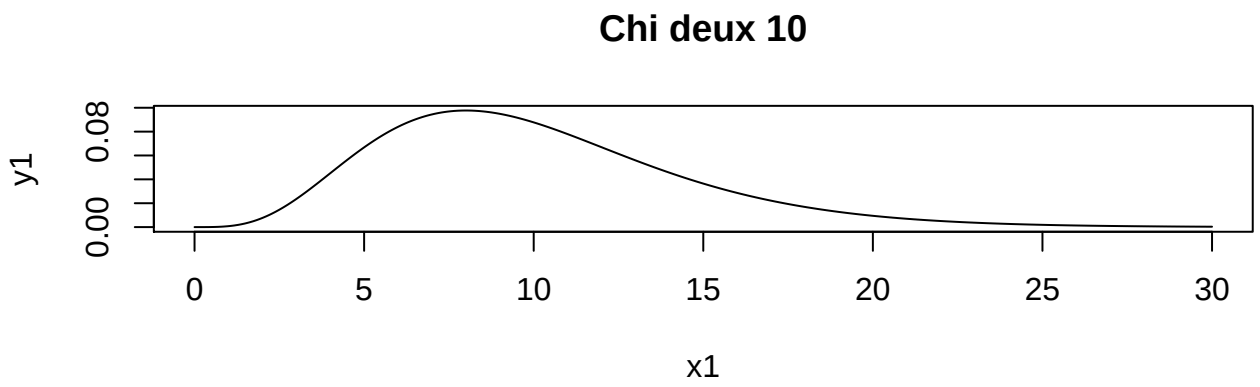
```
par(mfrow=c(2,1))
plot(x,y,type="l",main="N(0,1)")
plot(x1,y1,type="l",main="N(10,9)")
```

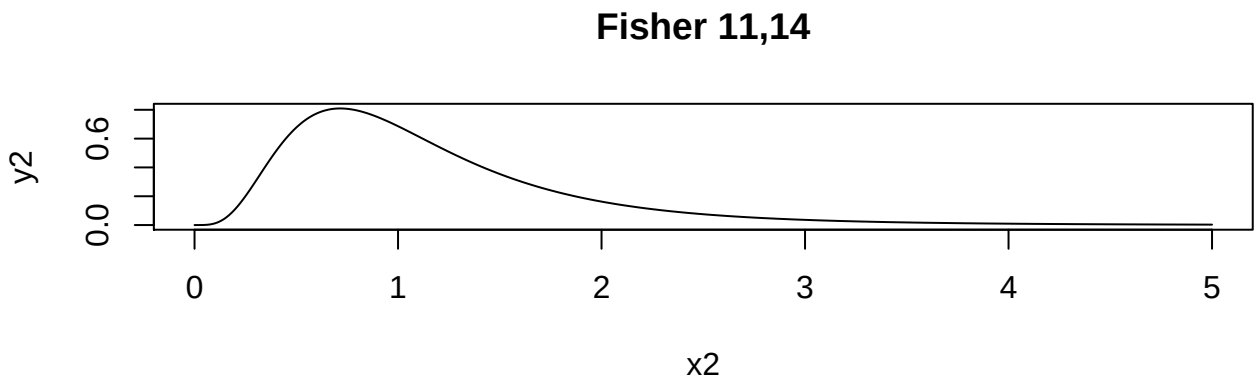




2 graphes sur une meme fenetre # Chi deux 10 et Fisher 11,14

```
par(mfrow=c(2,1))
x1=seq(0,30,0.01)
z=dnorm(500,2,3)
y1=dchisq(x1,df=10)
x2=seq(0,5,0.01)
y2=df(x2,df1=11,df2=14)
plot(x1,y1,type="l",main="Chi deux 10")
plot(x2,y2,type="l",main="Fisher 11,14")
```






```
pnorm(1.96,mean=0,sd=1)
```

```
## [1] 0.9750021
```

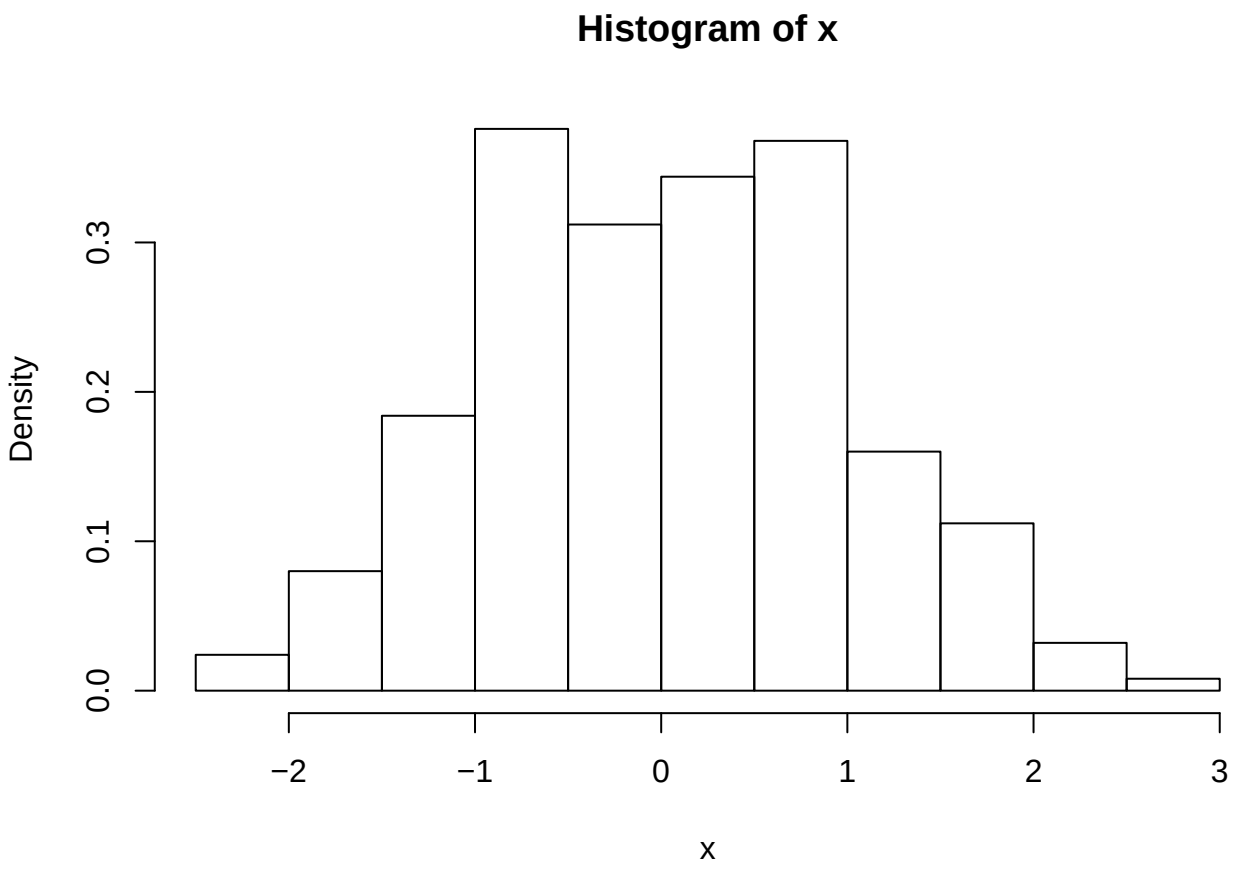
```
qnorm(0.05, mean=0,sd=1)
```

```
## [1] -1.644854
```

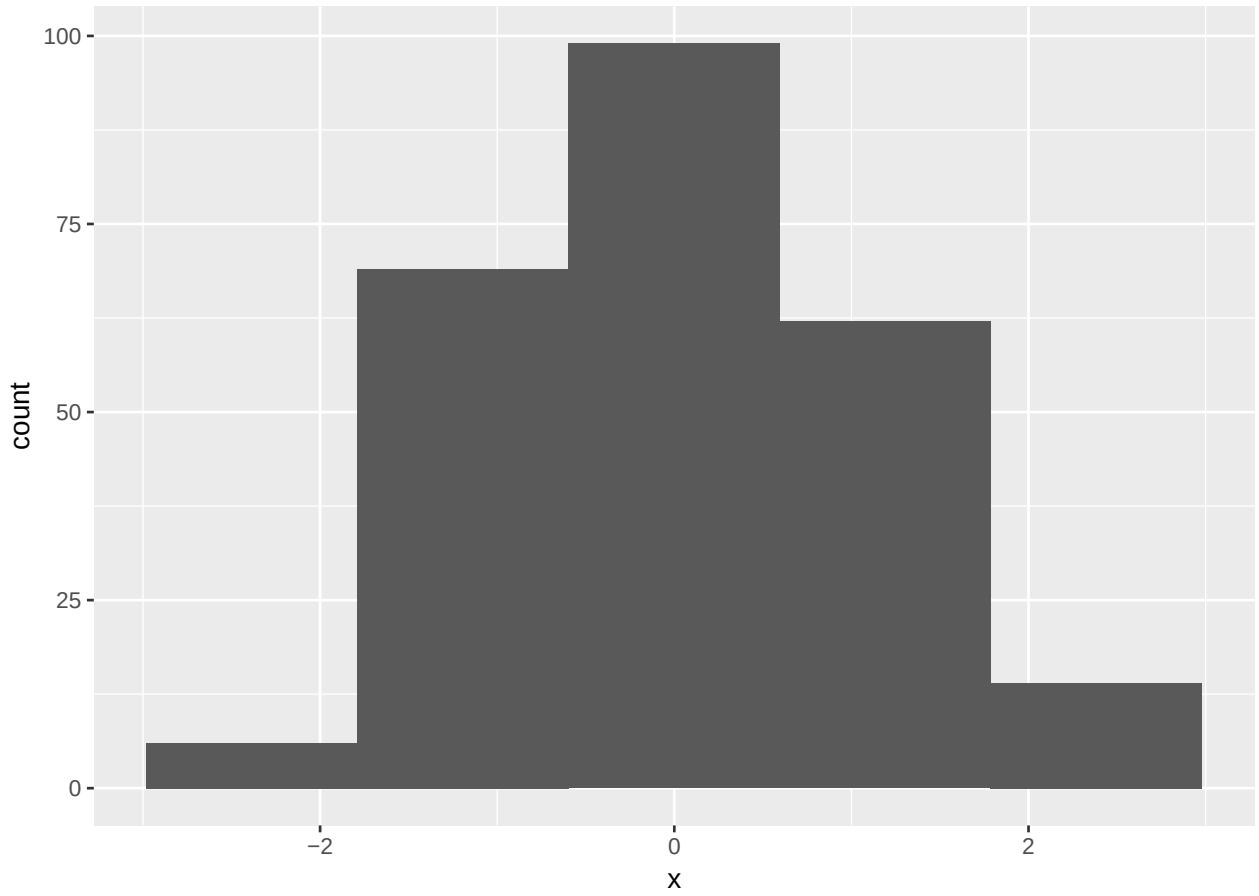
```
x<-rnorm(n=250,mean=0,sd=1)
mean(x)
```

```
## [1] 0.02870056
```

```
hist(x,freq=FALSE)
```




```
qplot(x, geom="histogram",bins=5)
```


100
75
50
25
0
count
−2
0
2
x