MAD_TP1

You ZUO

2019/9/16

Exo1

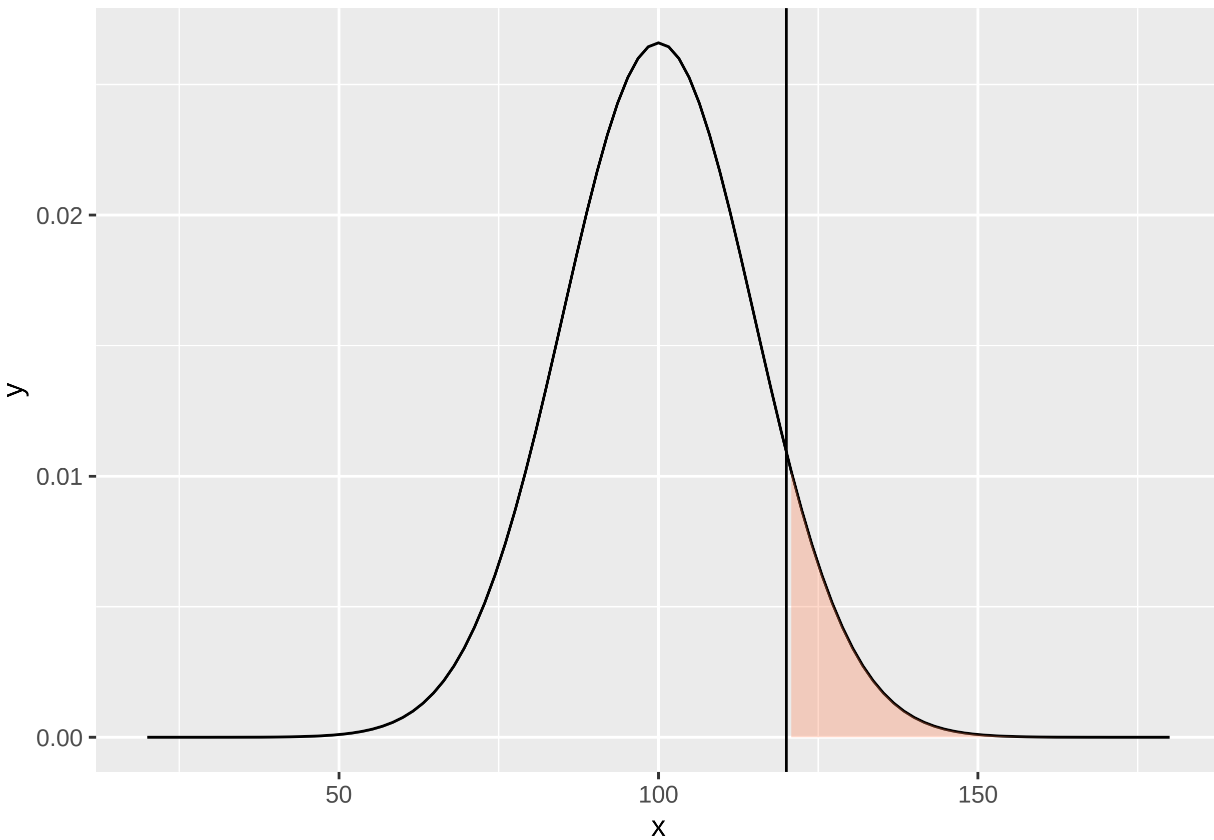
```r
1-pnorm(120,mean = 100,sd = 15)
```

```
## [1] 0.09121122
```

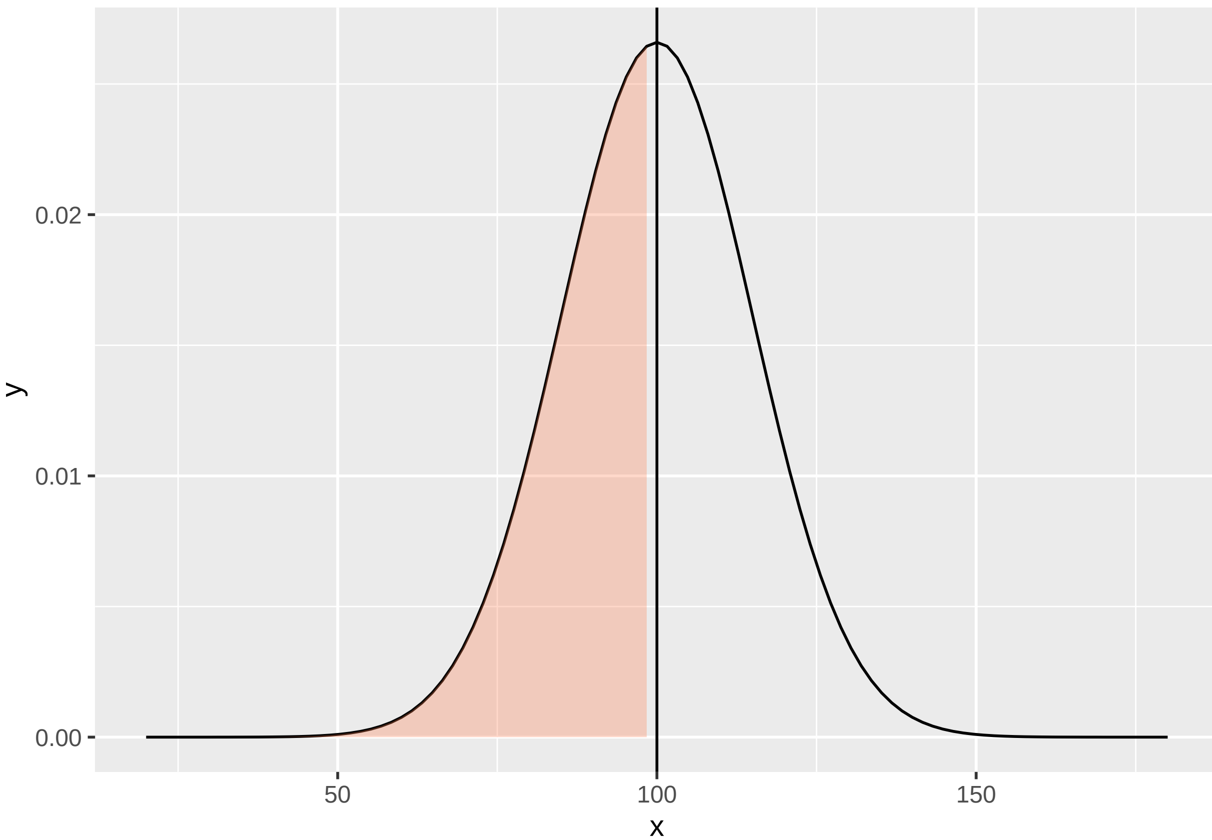
```r
pnorm(100,mean = 100,sd = 15)
```

```
## [1] 0.5
```

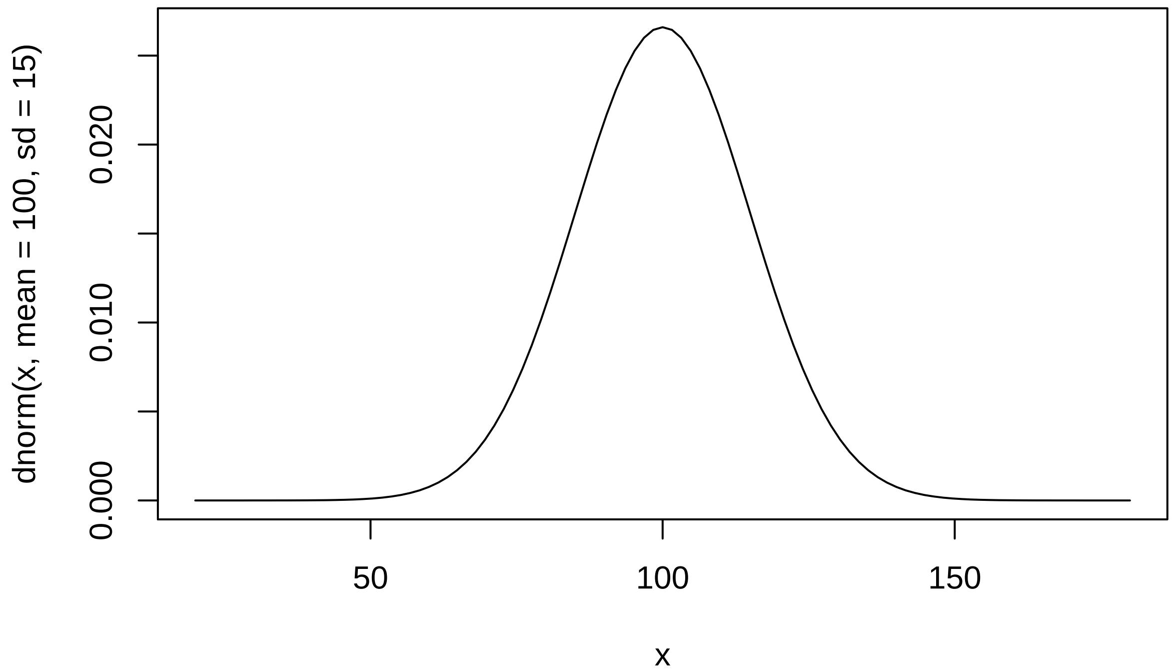
```r
QI.sup120 <- function(x){
  ifelse(x>120,dnorm(x,mean = 100, sd = 15), NA)
}

library(ggplot2)
ggplot(data.frame(x=c(20,180)),aes(x)) +
  stat_function(fun = dnorm, args = list(mean=100,sd=15)) +
  stat_function(fun = QI.sup120, geom = "area", fill = "coral", alpha = 0.3) +
  geom_vline(xintercept = 120)
```



```
QI.inf100 <- function(x){
  ifelse(x<100,dnorm(x,mean = 100,sd = 15), NA)
}
ggplot(data.frame(x=c(20,180)),aes(x)) +
  stat_function(fun = dnorm, args = list(mean=100,sd=15)) +
  stat_function(fun = QI.inf100, geom = "area", fill = "coral", alpha = 0.3) +
  geom_vline(xintercept = 100)
```




```
curve(dnorm(x,mean = 100,sd = 15),20,180)
```

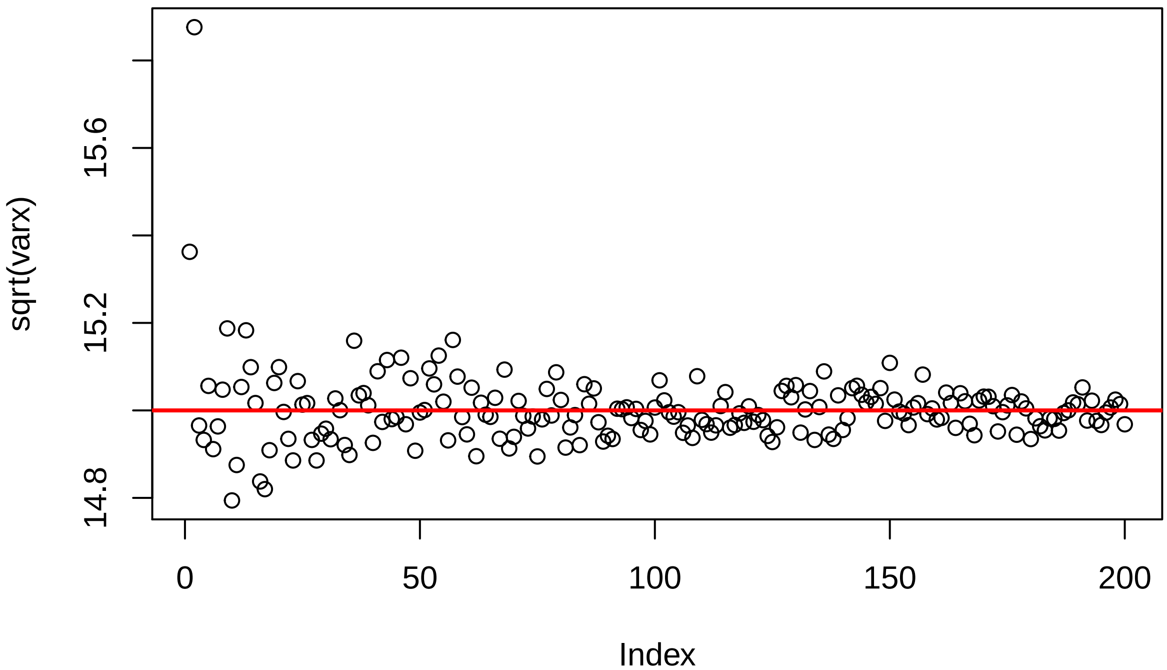


Exo2

$$\sigma_{\hat{ML}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

with $\hat{x} = \frac{1}{n}\sum x_i$

```
n <- seq(from = 100, to = 100000, by = 500)
varx <- sapply(n,function(m){
  x <- rnorm(m,mean = 100,sd = 15)
  var(x)
})
plot(sqrt(varx))
abline(h = 15, col = "red", lwd = 2)
```




```r
n <- 10
x <- rnorm(n = 10, mean = 100, sd = 15)
sigma2ML <- mean((x-mean(x))^2)
sigma2stat <- n/(n-1)*sigma2ML
var(x)
```

```
## [1] 301.2664
```

```r
sigma2ML
```

```
## [1] 271.1397
```

```r
sigma2stat
```

```
## [1] 301.2664
```

Exo3

```r
data("iris")
str(iris)
```

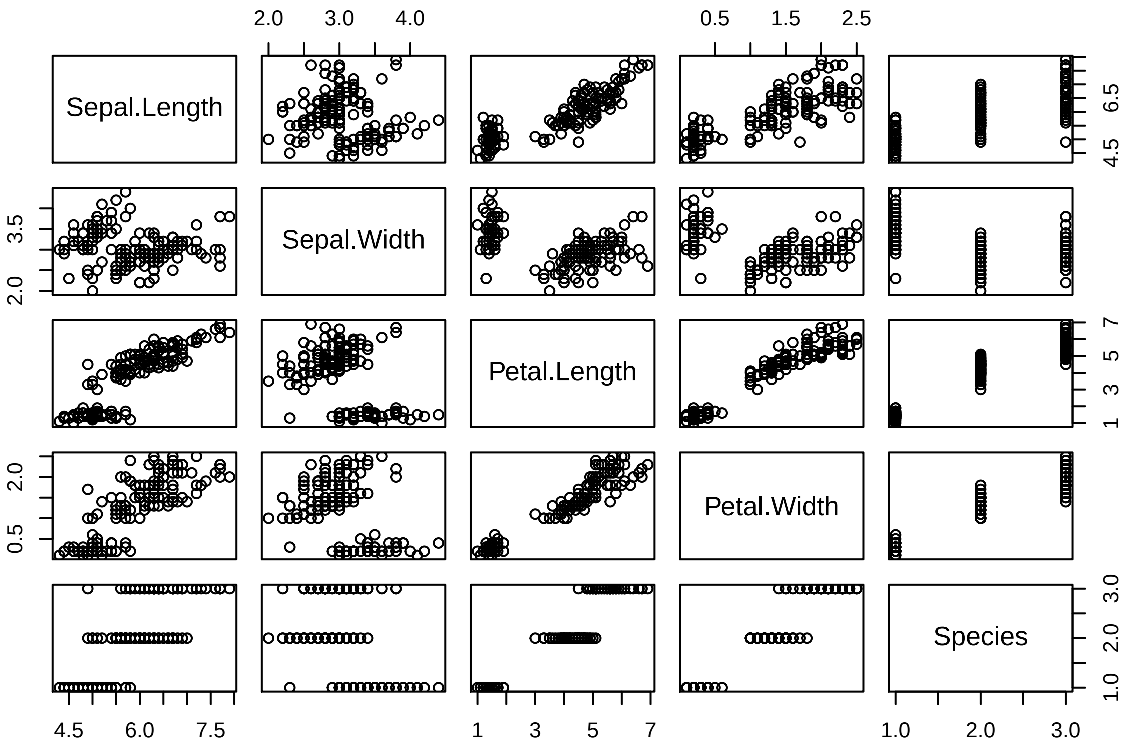
```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
##  $ Species     : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```r
help("iris")
summary(iris)
```

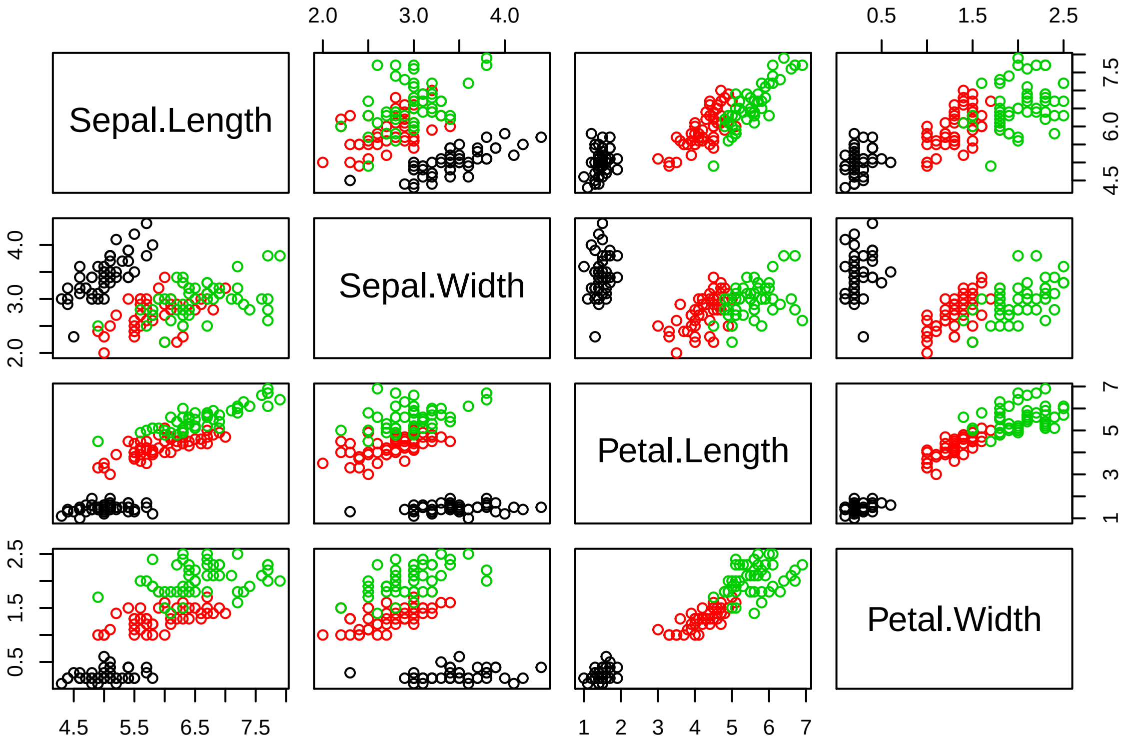
```
##   Sepal.Length    Sepal.Width     Petal.Length    Petal.Width
##  Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100
##  1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300
```

```
##  Median :5.800   Median :3.000   Median :4.350   Median :1.300
##  Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
##  3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
##  Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
##        Species
##  setosa    :50
##  versicolor:50
##  virginica :50
##
##
##
```

```
plot(iris)
```



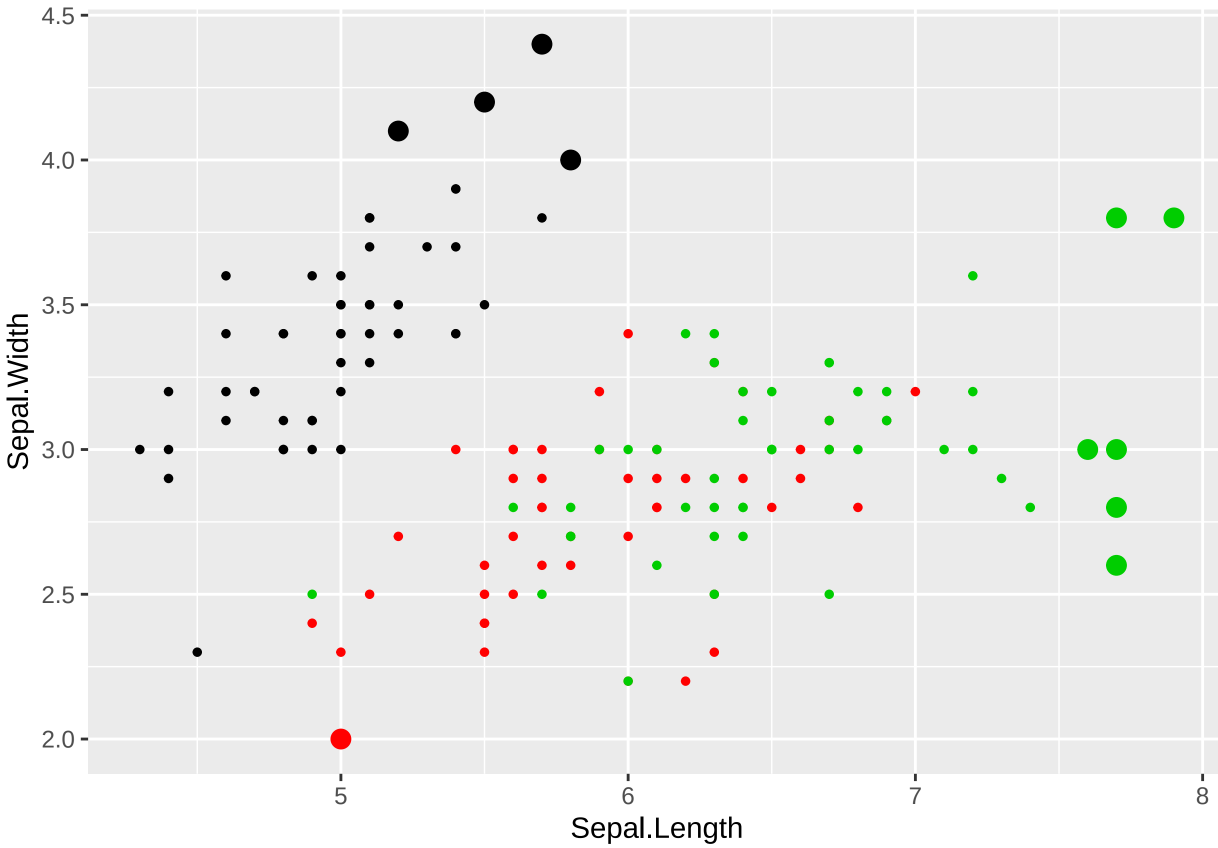

```
pairs(iris[,1:4], col = iris$Species)
```




```
library(tidyverse)
```

```
## -- Attaching packages ----------------------- tidyverse 1.2.1 --

## v tibble  2.1.3     v purrr   0.3.2
## v tidyr   0.8.3     v dplyr   0.8.3
## v readr   1.3.1     v stringr 1.4.0
## v tibble  2.1.3     v forcats 0.3.0

## -- Conflicts ------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
parameters <-
  as.tibble(iris) %>%
  select(-"Species") %>%
  gather(factor_key = TRUE) %>%
  group_by(key) %>%
  summarise(mean = mean(value), sd = sd(value)) %>%
  mutate(min=mean - 2*sd, max = mean + 2*sd)

flower.outliers <- (apply(t((t(iris[,1:4]) < parameters$min) + (t(iris[,1:4]) > parameters$max)),1,sum)
ggplot(iris, aes(x=Sepal.Length,y=Sepal.Width)) +
  geom_point(colour = as.numeric(iris$Species), size = flower.outliers*2+1)
```

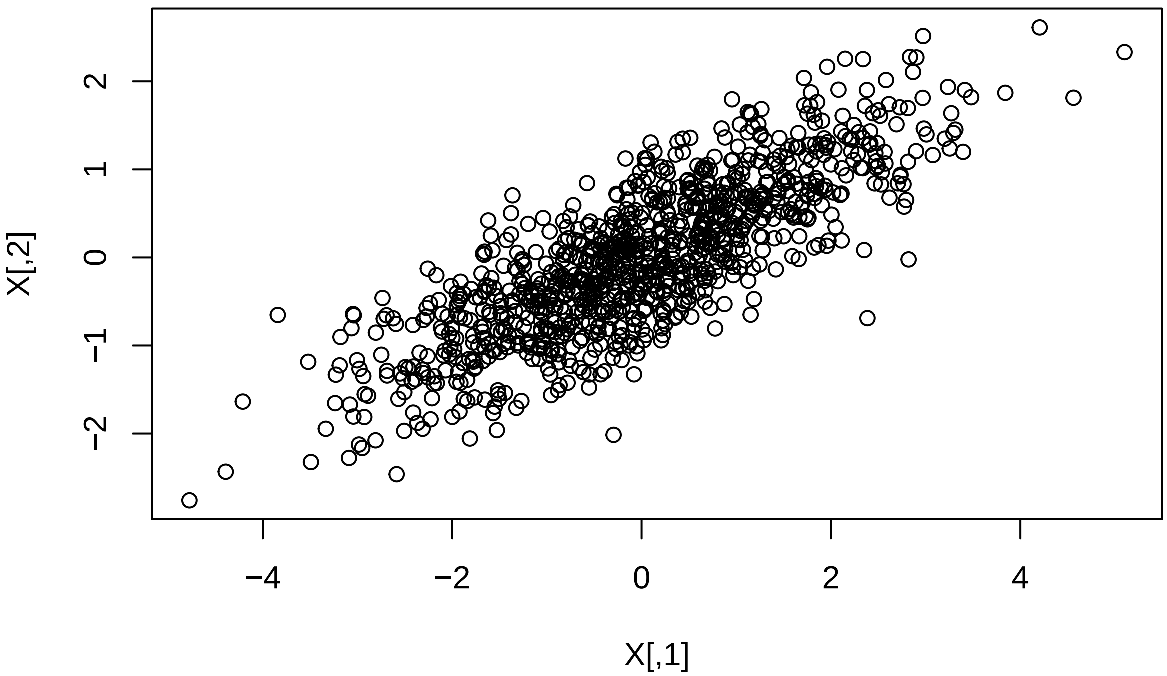


exo4

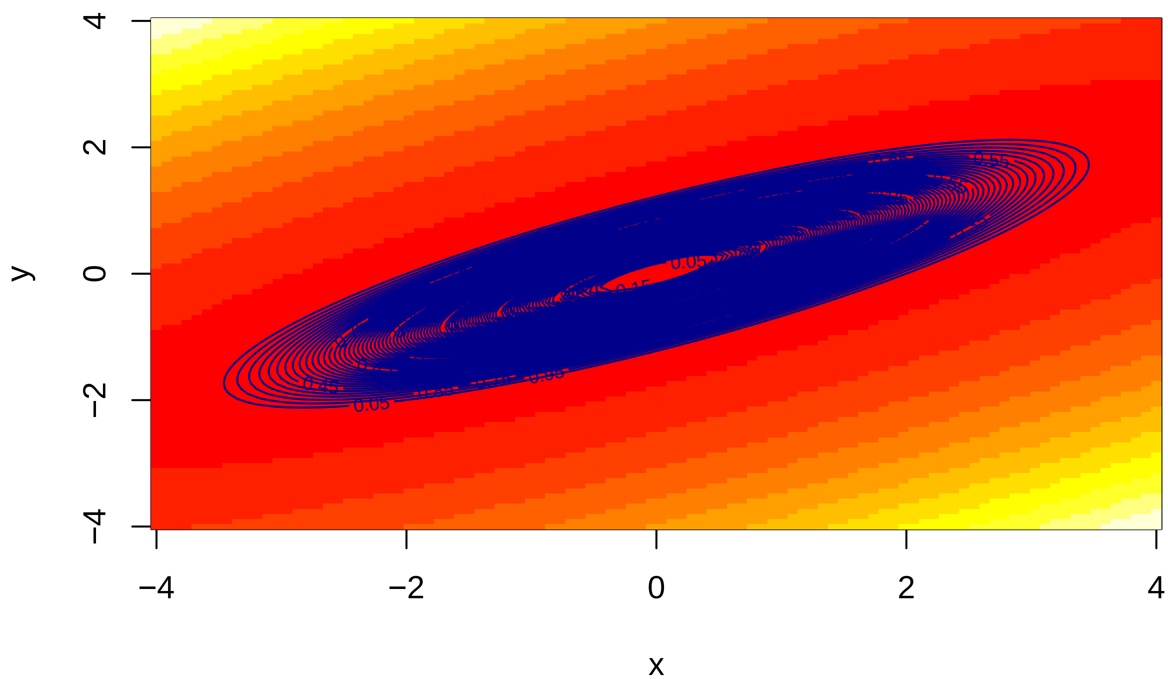
```
library(mvtnorm)
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

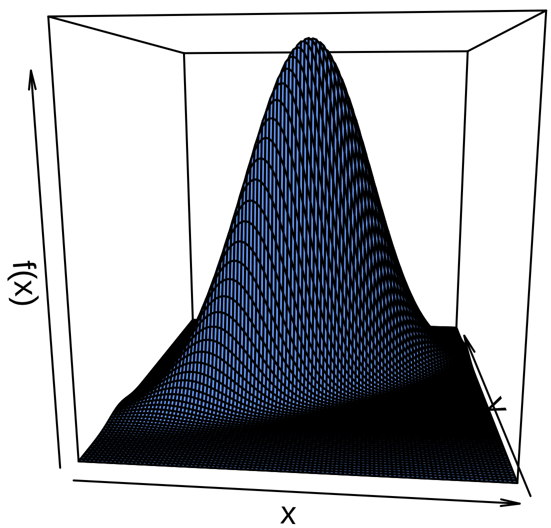
```
sigma = matrix(nrow = 2, ncol = 2, data = c(2,1,1,0.75))
mu = matrix(c(0,0),2,1)

X <- mvrnorm(n = 1000, mu = mu, Sigma = sigma)
plot(X)
```




```
Q <- qchisq(p = seq(0.05,0.95,by = 0.01), df = 2)
x <- seq(-4,4,length=100)
y <- seq(-4,4,length=100)
sigmainv <- solve(sigma)
a <- sigmainv[1,1]
b <- sigmainv[2,2]
c <- sigmainv[1,2]
z <- outer(x,y,function(x,y) (a*x^2+b*y^2+2*c*x*y))
image(x,y,z)
contour(x,y,z,col = "blue4", levels = Q, labels = seq(0.05,0.95,0.1),add=T)
```



```
persp(x,y,1/(2*pi)*det(sigmainv)^(-1/2)*exp(-0.5*z),col = "cornflowerblue",theta = 5, phi = 10, zlab =
```



$$logf(x_1, x_2) = -\frac{1}{2}(x_1, x_2)\Sigma^{-1}(x_1, x_2)^t$$