# 24-12-05

December 22, 2024

[1]:
```r
dir_tex <- "/Users/gianlucamastrantonio/Dropbox (Politecnico di Torino Staff)/
 ↪Didattica/statistica computazionale/dispense/mod3/"
setwd("/Users/gianlucamastrantonio/Dropbox (Politecnico di Torino Staff)/
 ↪Didattica/statistica computazionale/codice R/mixture wind")

library(ggplot2)
library(rstan)
library(nimble)

load("/Users/gianlucamastrantonio/Dropbox (Politecnico di Torino Staff)/
 ↪Didattica/statistica computazionale/datasets/wind.RData")

gg_theme <- theme(
    legend.title = element_text(size = 25),
    axis.title = element_text(size = 28), # Increase axis title size
    axis.text = element_text(size = 24),
    legend.position = "bottom",
    legend.key.size = unit(1., "cm"),
    legend.text = element_text(size = 24)
)

ls()
```

```
Caricamento del pacchetto richiesto: StanHeaders


rstan version 2.32.6 (Stan version 2.32.2)


For execution on a local, multicore CPU with excess RAM we recommend calling
options(mc.cores = parallel::detectCores()).
To avoid recompilation of unchanged Stan programs, we recommend calling
rstan_options(auto_write = TRUE)
For within-chain threading using `reduce_sum()` or `map_rect()` Stan functions,
change `threads_per_chain` option:
rstan_options(threads_per_chain = 1)
```

```
nimble version 1.2.1 is loaded.
For more information on NIMBLE and a User Manual,
please visit https://R-nimble.org.

Note for advanced users who have written their own MCMC samplers:
  As of version 0.13.0, NIMBLE's protocol for handling posterior
  predictive nodes has changed in a way that could affect user-defined
  samplers in some situations. Please see Section 15.5.1 of the User Manual.


Caricamento pacchetto: 'nimble'


Il seguente oggetto `e mascherato da 'package:stats':

    simulate


Il seguente oggetto `e mascherato da 'package:base':

    declare
```

1. 'dir_tex' 2. 'gg_theme' 3. 'wind'

[2]:
```
summary(wind)
```

```
     plot                lat             lon
 Length:81663       Min.   :43.34   Min.   :-113.1
 Class :character   1st Qu.:43.40   1st Qu.:-113.0
 Mode  :character   Median :43.40   Median :-113.0
                    Mean   :43.40   Mean   :-113.0
                    3rd Qu.:43.41   3rd Qu.:-113.0
                    Max.   :43.43   Max.   :-113.0
    datetime                       obs_speed         obs_dir
 Min.   :2010-06-13 08:00:00.00   Min.   : 0.000   Min.   :  0.0
 1st Qu.:2010-07-03 00:00:00.00   1st Qu.: 1.900   1st Qu.:111.0
 Median :2010-07-24 14:00:00.00   Median : 3.100   Median :196.0
 Mean   :2010-07-25 09:51:26.65   Mean   : 3.844   Mean   :183.8
 3rd Qu.:2010-08-16 17:00:00.00   3rd Qu.: 5.100   3rd Qu.:254.0
 Max.   :2010-09-10 20:00:00.00   Max.   :46.900   Max.   :360.0
```

Simuliamo da un modello mistura del tipo

$$y_i|z_i \sim N(\mu_{z_i}, \sigma^2_{z_i})$$

$$z_i \sim Discr(\pi_1, \pi_2)$$

# 1 esempio 1

$$\pi = (0.3, 0.7)$$

$$\sigma_1^2 = \sigma_2^2 = 2$$

$$\mu_1 = -10 \quad \mu_2 = 5$$

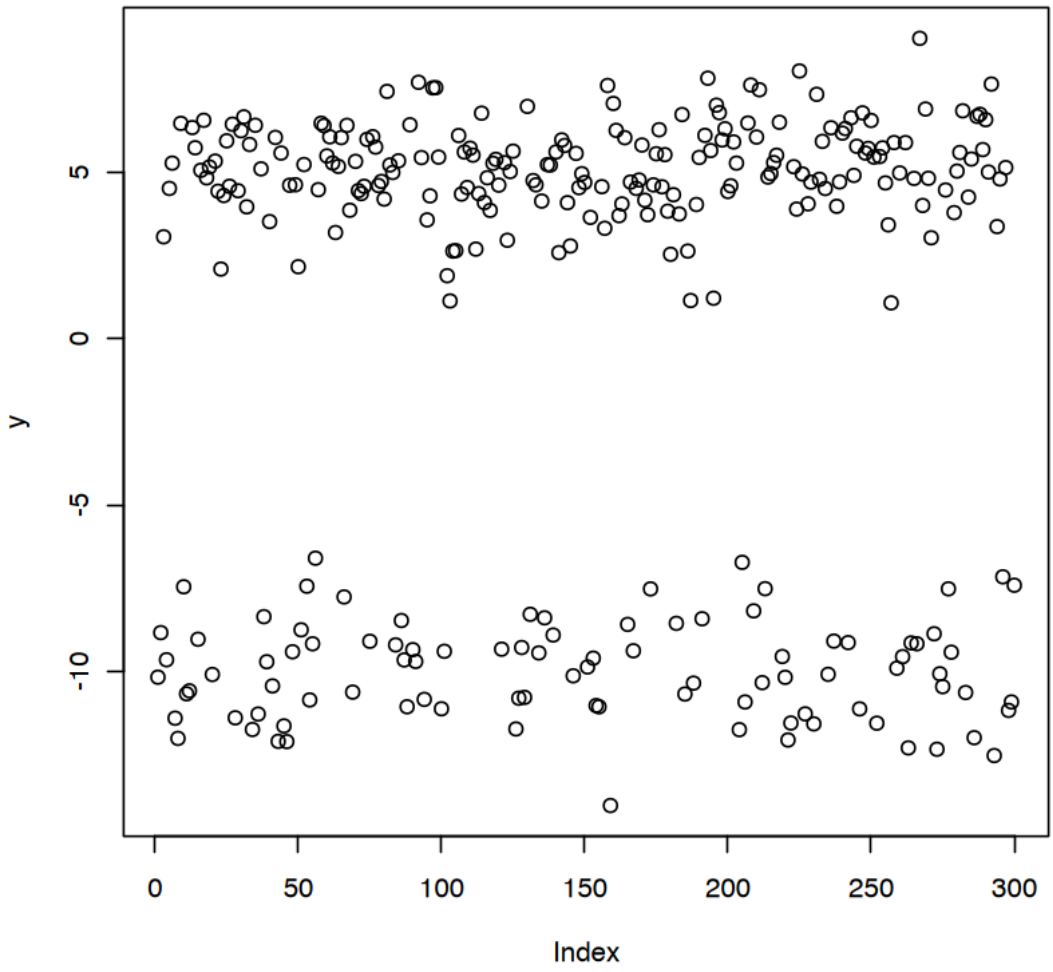
[24]:
```
pi_vec = c(0.3, 0.7)
mu_vec = c(-10,5)
sigma2_vec = c(2,2)
n = 300
```

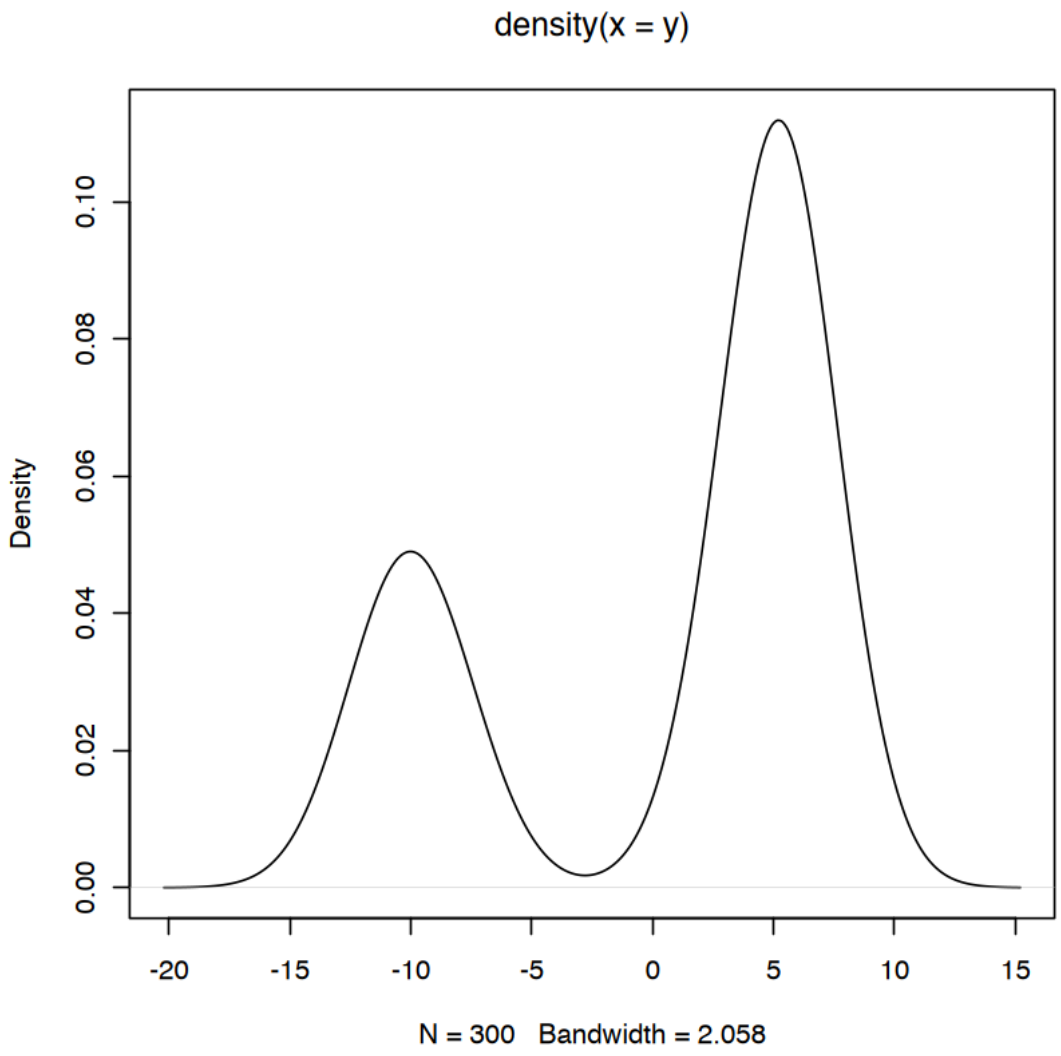
[ ]:
```
z = matrix(NA, ncol=1, nrow=n)
y <- matrix(NA, ncol = 1, nrow = n)
for(i in 1:n)
{
  z[i] = sample(1:2, 1, prob = pi_vec)
  y[i] = rnorm(1, mu_vec[z[i]], sigma2_vec[z[i]]^0.5)
}

plot(y)
plot(density(y))
```

[ ]:
```

```


y
5
0
-5
-10
0
50
100
150
200
250
300
Index

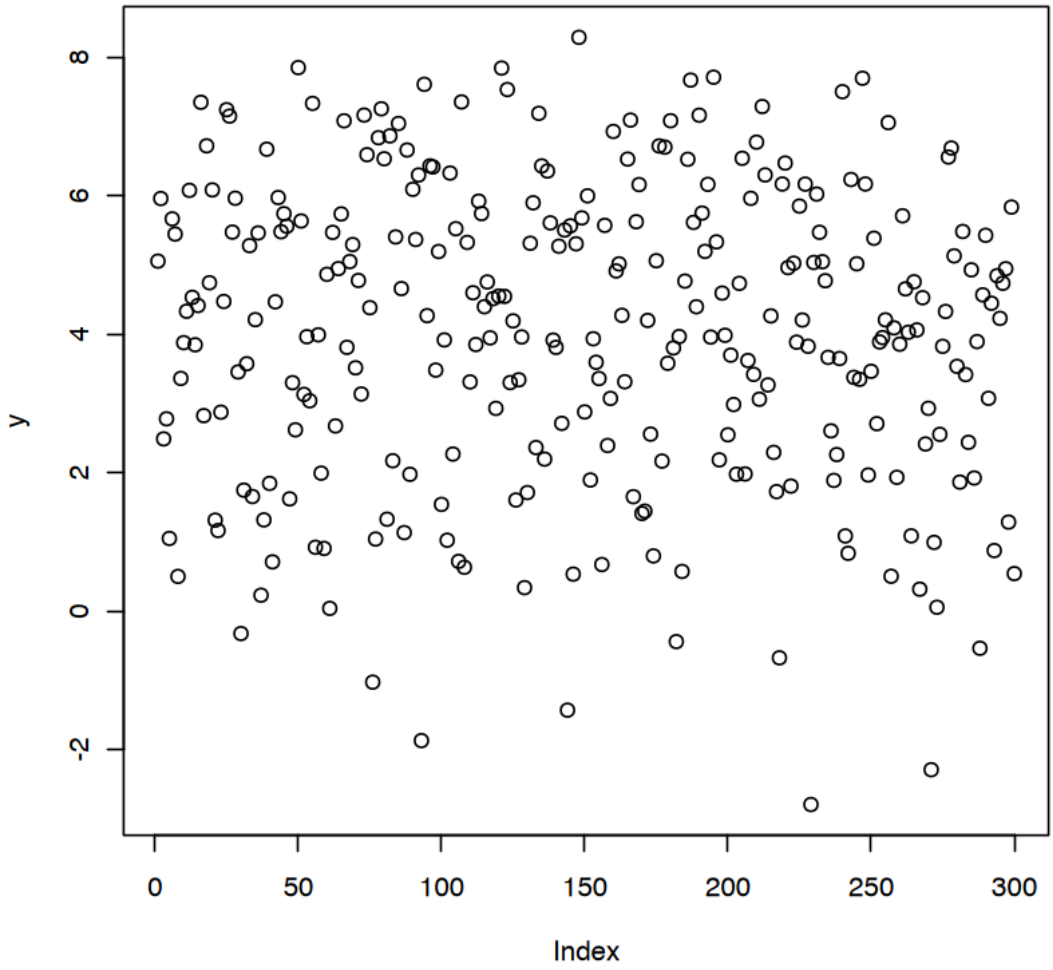


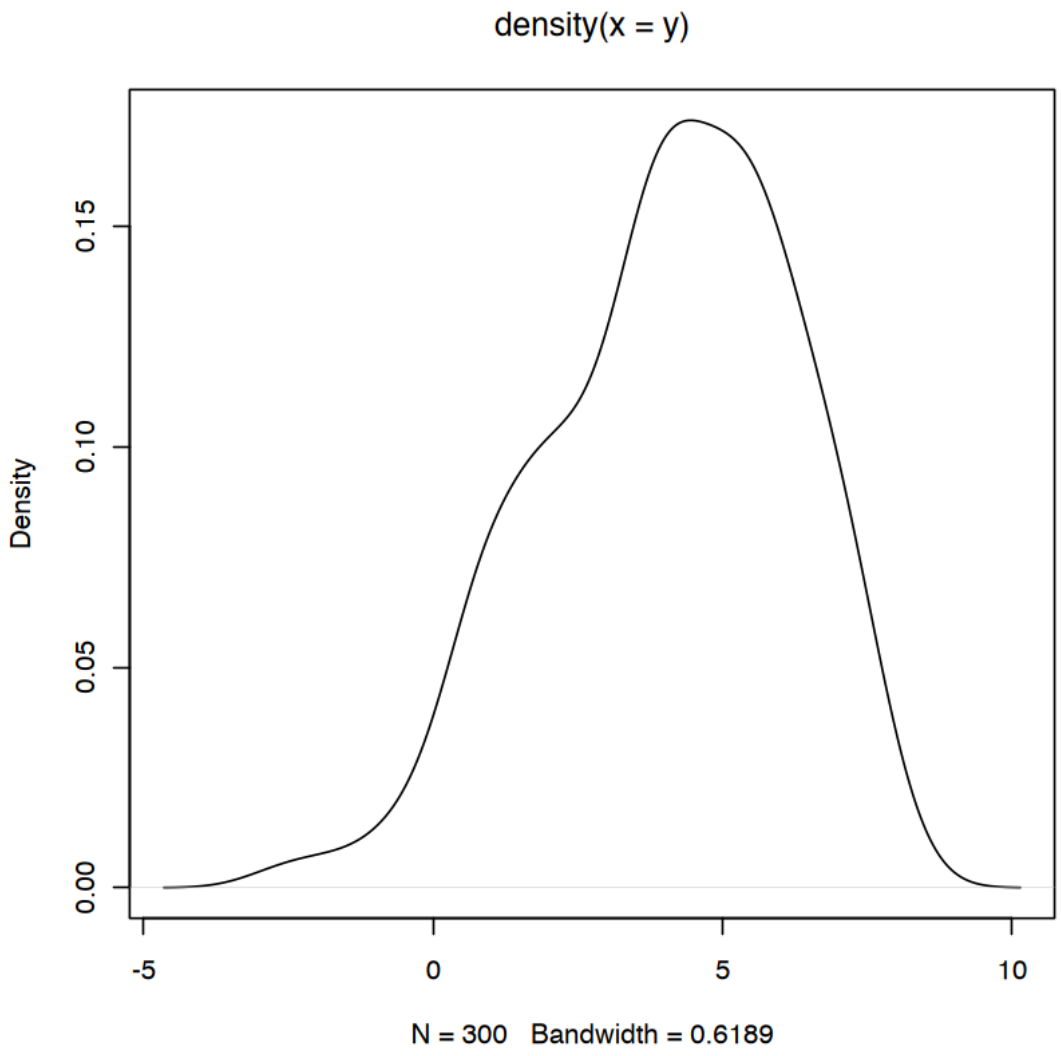

$$\mu_1 = +2 \quad \mu_2 = 5$$

[30]:
```
mu_vec <- c(2, 5)
z <- matrix(NA, ncol = 1, nrow = n)
y <- matrix(NA, ncol = 1, nrow = n)
for (i in 1:n)
{
  z[i] <- sample(1:2, 1, prob = pi_vec)
  y[i] <- rnorm(1, mu_vec[z[i]], sigma2_vec[z[i]]^0.5)
}

plot(y)
```

```
plot(density(y))
```




density(x = y)
Density
0.00
0.05
0.10
0.15
-5
0
5
10
N = 300 Bandwidth = 0.6189

[ ]: