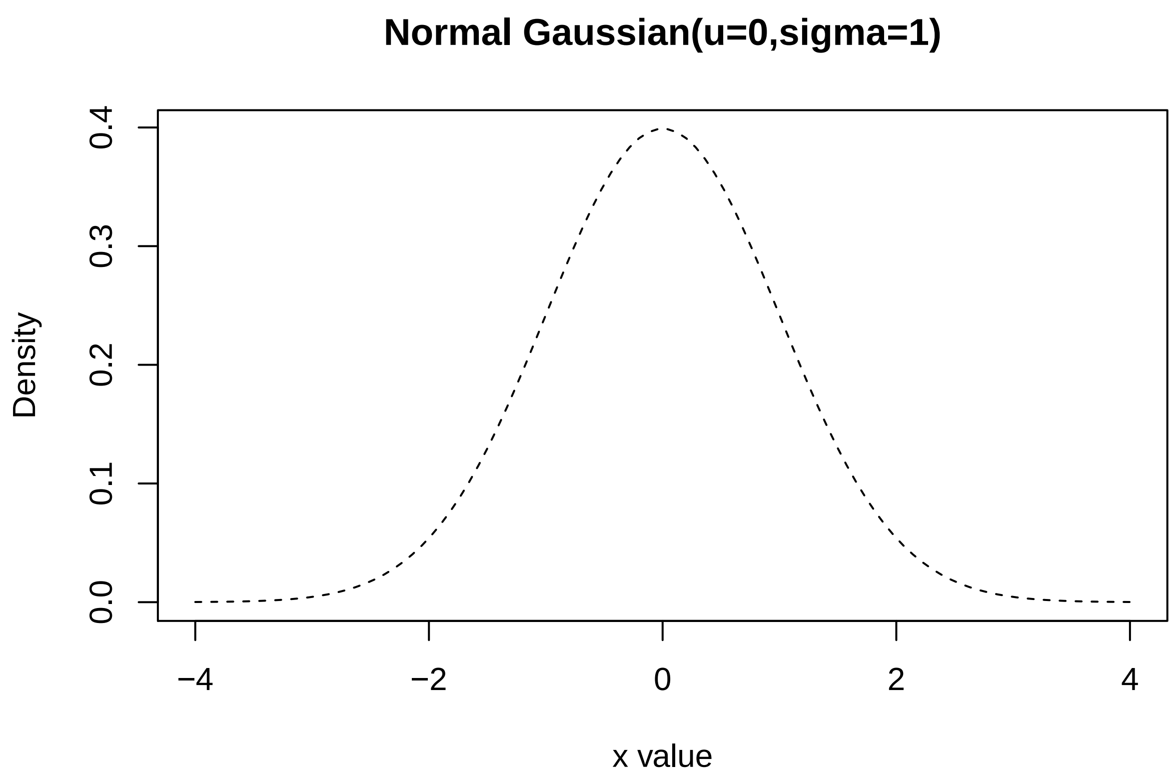
Maximum Likelihood

Daniel Frederico Lins Leite

19 March 2017

Gaussian

```
x <- seq(-4, 4, length=100)
hx <- dnorm(x)
plot(x, hx, type="l", lty=2, xlab="x value", ylab="Density", main="Normal Gaussian(u=0,sigma=1)")
```

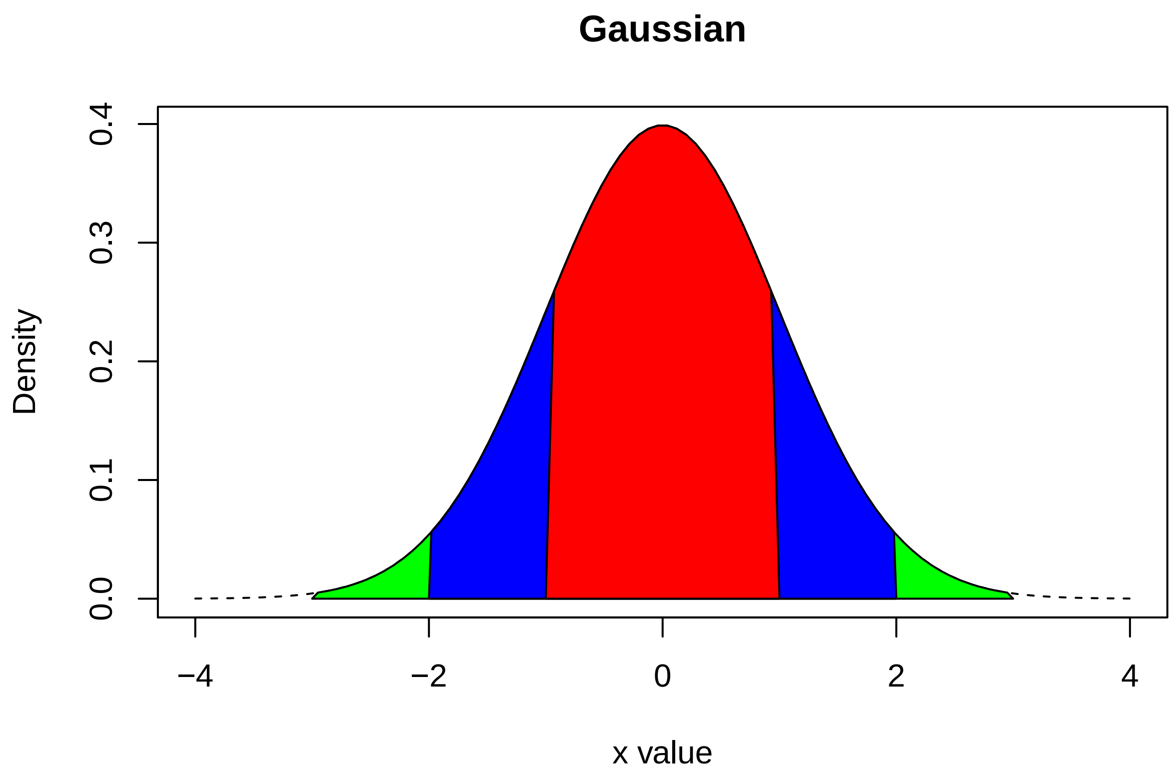


Percentage

```
plotGaussian <- function(u, sigma, color = "black"){
  x <- seq(-4, 4, length=100)
  hx <- dnorm(x, mean = u, sd = sigma)
  plot(x, hx, type="l", lty=2, xlab="x value", ylab="Density", main="Gaussian", col = color)
}
plotArea <- function(u, sigma, sigmaSize, color){
  x <- seq(-4, 4, length=100)
  hx <- dnorm(x, mean = u, sd = sigma)
  l <- -(sigma*sigmaSize)
```

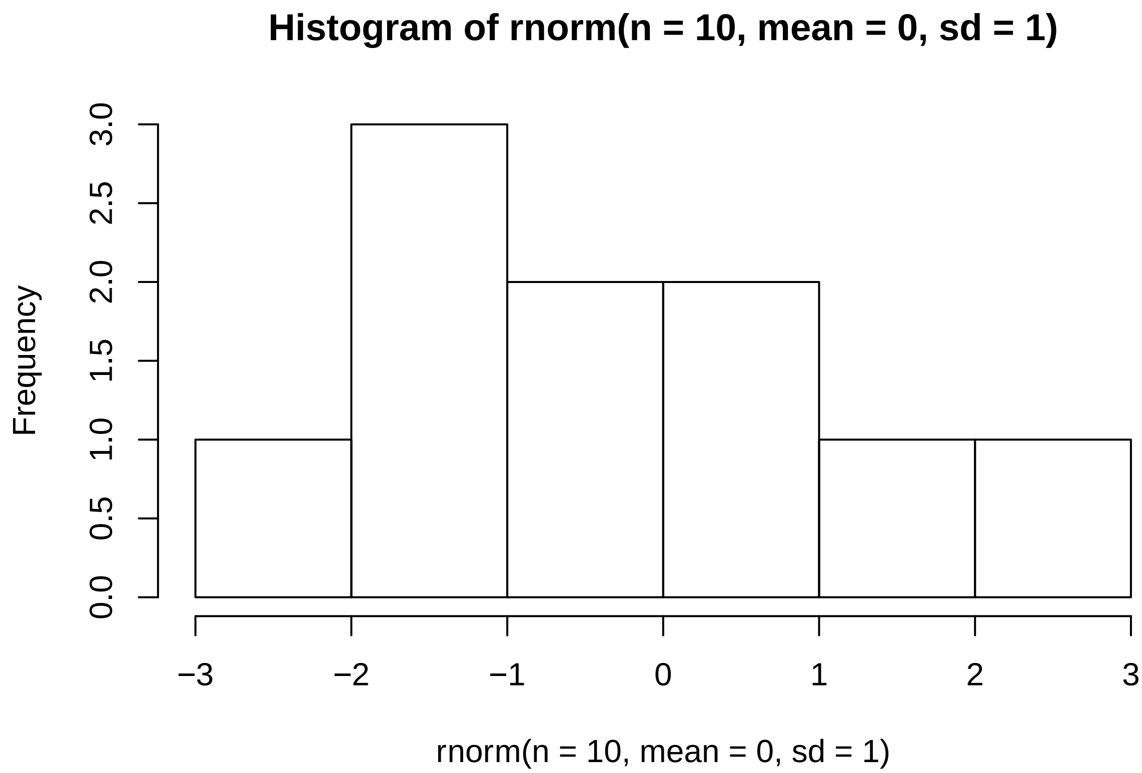
```
  r <- (sigma*sigmaSize)
  i <- x >= l & x <= r
  polygon(c(l,x[i],r), c(0,hx[i],0), col=color)
}

plot.new()
plotGaussian(0,1)
plotArea(0,1,3,"green")
plotArea(0,1,2,"blue")
plotArea(0,1,1,"red")
```

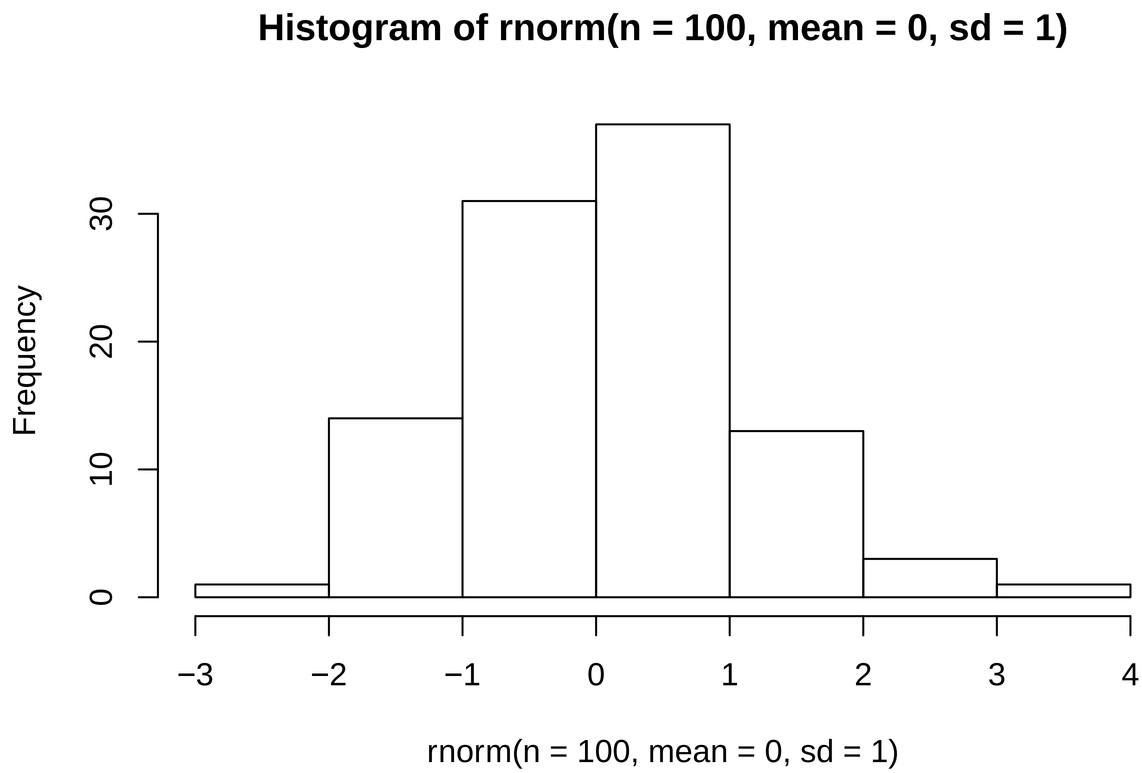


Density Simulation

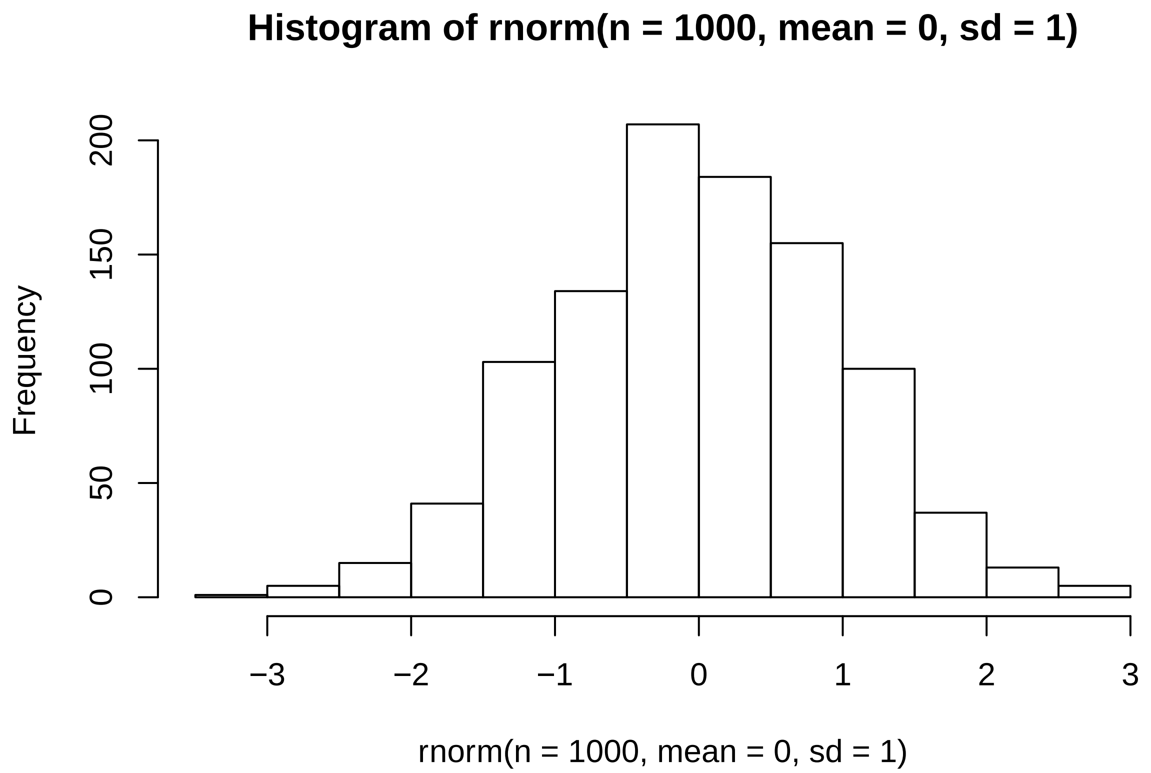
```
hist(rnorm(n = 10, mean = 0, sd = 1))
```



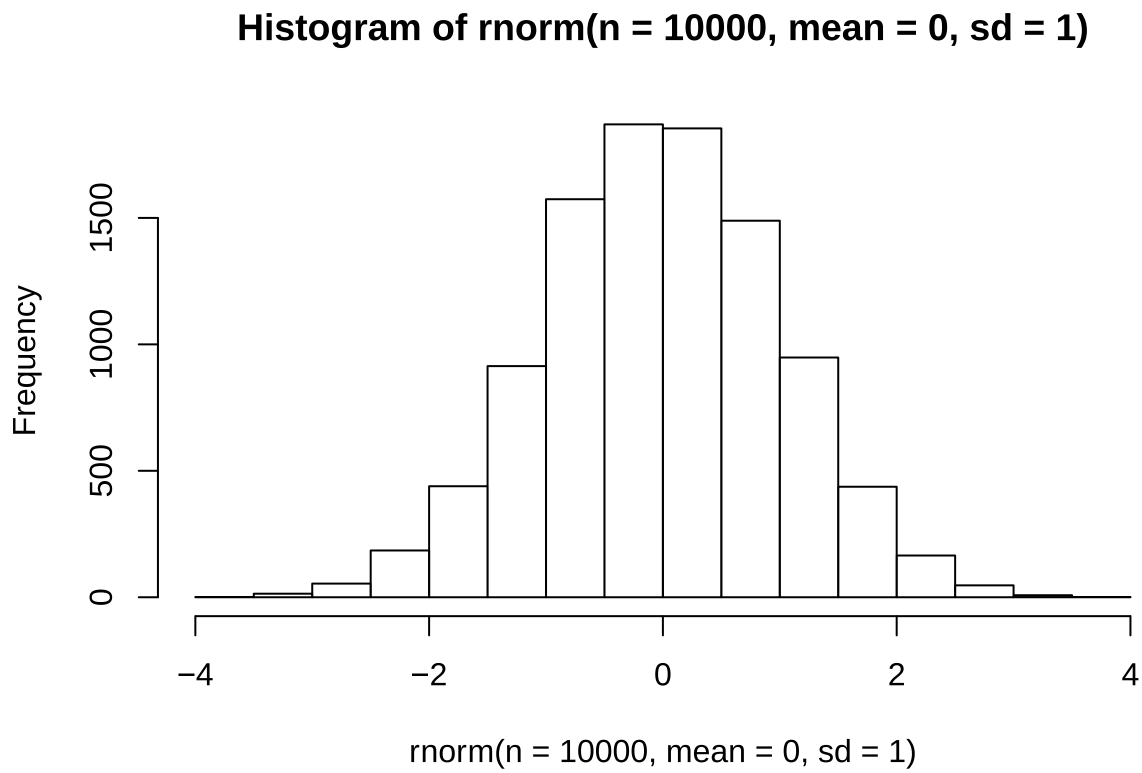

```r
hist(rnorm(n = 100, mean = 0, sd = 1))
```




```
hist(rnorm(n = 1000, mean = 0, sd = 1))
```




```
hist(rnorm(n = 10000, mean = 0, sd = 1))
```



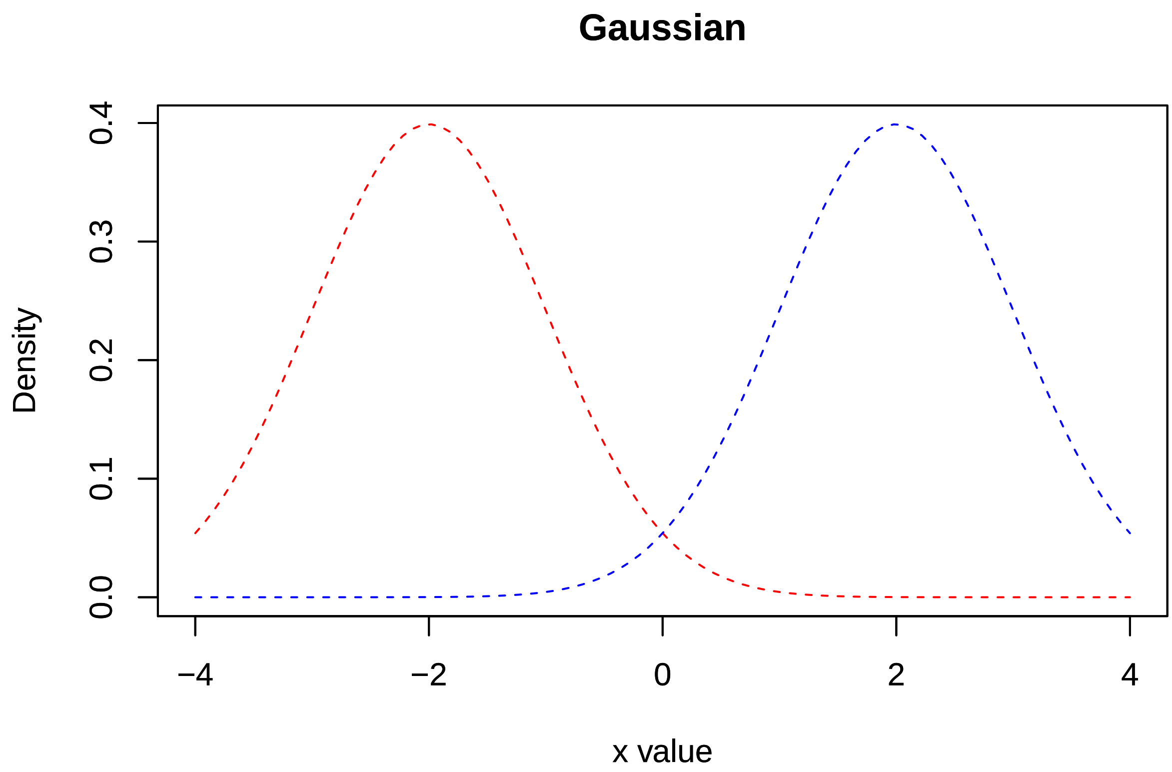
Likelihood

Suppose we have an observation and two possible distributions that can be considered as the source distributions of this observations. We want to choose the best option: in this case the most probable source distribution.

Options:
Gaussian #1
u = -2
sd = 1

Gaussian #2
u = 2
sd = 1

```
plotGaussian(-2,1,"red")
par(new=TRUE)
plotGaussian(2,1, "blue")
```



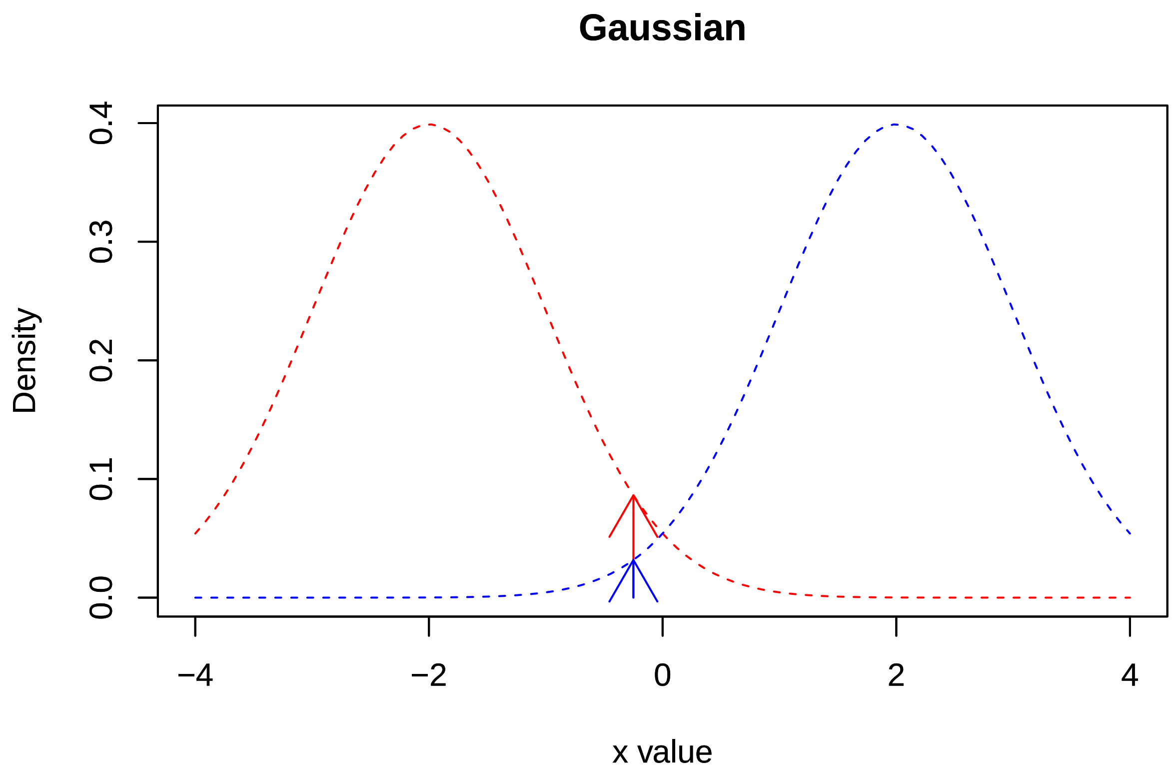
Observations
A = -0.25
B = -1
C = 0.45

To choose the best distribution we will choose the distribution whose density is bigger in that particular point. For example, for observation A:

Probability of observing A from Gaussian #1 (u-2,sd=1) = 0.0862773
Probability of observing A from Gaussian #2 (u=2,sd=1) = 0.0317397

In this case the best option is Gaussian #1. It is more likely/probable that the observation #A comes from Gaussian #1 than Gaussian #2.

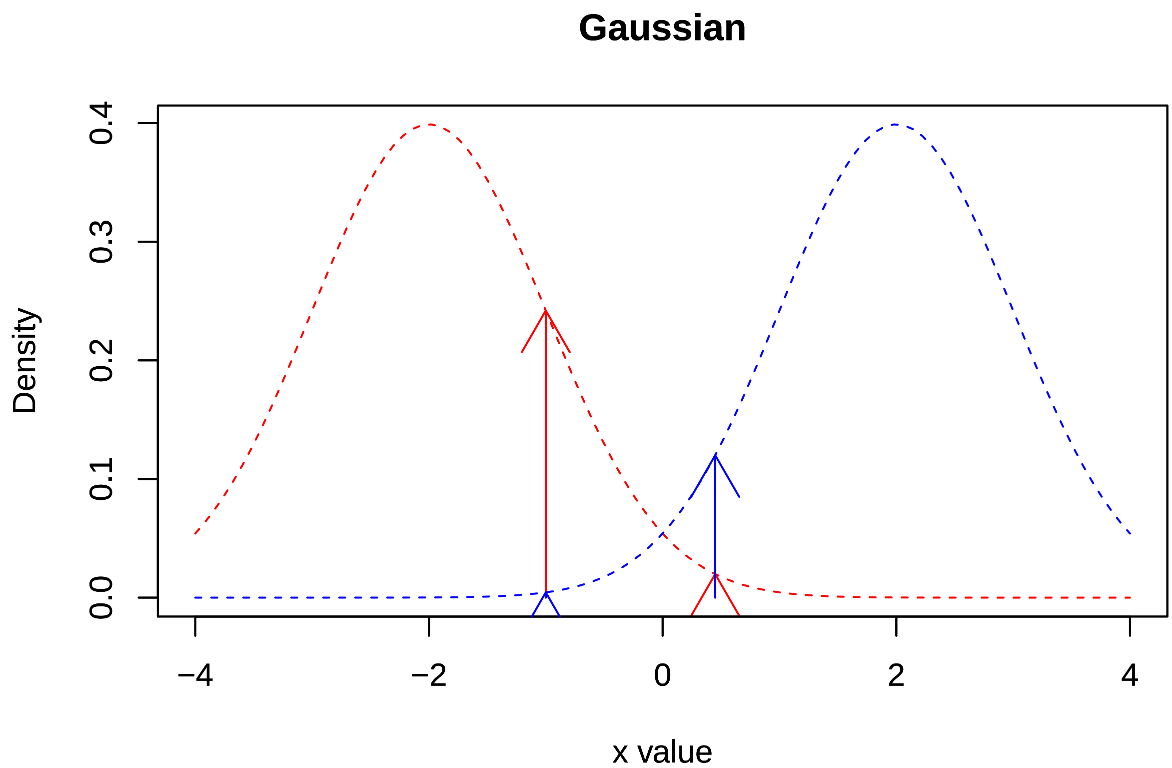
```r
plotGaussian(-2,1,"red")
arrows(-0.25,0,-0.25,dnorm(-0.25,-2,1), col = "red")
par(new=TRUE)
plotGaussian(2,1, "blue")
arrows(-0.25,0,-0.25,dnorm(-0.25,2,1), col = "blue")
```



We can use the same calculations to the others observations.

Probability of observing B from Gaussian #1 (u-2,sd=1) = 0.2419707
Probability of observing B from Gaussian #2 (u=2,sd=1) = 0.0044318
Probability of observing C from Gaussian #1 (u-2,sd=1) = 0.0198374
Probability of observing C from Gaussian #2 (u=2,sd=1) = 0.120009

```
plotGaussian(-2,1,"red")
arrows(-1,0,-1,dnorm(-1,-2,1), col = "red")
arrows(0.45,0,0.45,dnorm(0.45,-2,1), col = "red")
par(new=TRUE)
plotGaussian(2,1, "blue")
arrows(-1,0,-1,dnorm(-1,2,1), col = "blue")
arrows(0.45,0,0.45,dnorm(0.45,2,1), col = "blue")
```



So we can sau that:
The likelihood of u = -2 and sd = 1 for a observation -1 is 0.2419707 because the probability of seeing a value -1 from a Gaussian with u = -2 and sd = 1 is 0.2419707.
In others words:
The likelihood of theta, the gaussian parameters, given x is 0.2419707 because the probability of x given theta is 0.2419707.
Or:
L(theta|x) = p(x|theta)

Maximun Likelihood

Imagine now that we do not have two options predertimined. We have all possible Gaussians and we want to find the best gaussian to each observation, first one-by-one, and them as a set. The best gaussian for observation A is the Gaussian with the maximum likelihod. So our problem can be described as:

maximize theta in L(theta|x)

We saw that this is the same as

maximize theta in p(x | theta)

Se given x, which is a known value, we must find the theta that maximize the function. theta in this particular case is the set {u,sd}. So given X, we must find u and sd that will maximize p(x|u,sd).

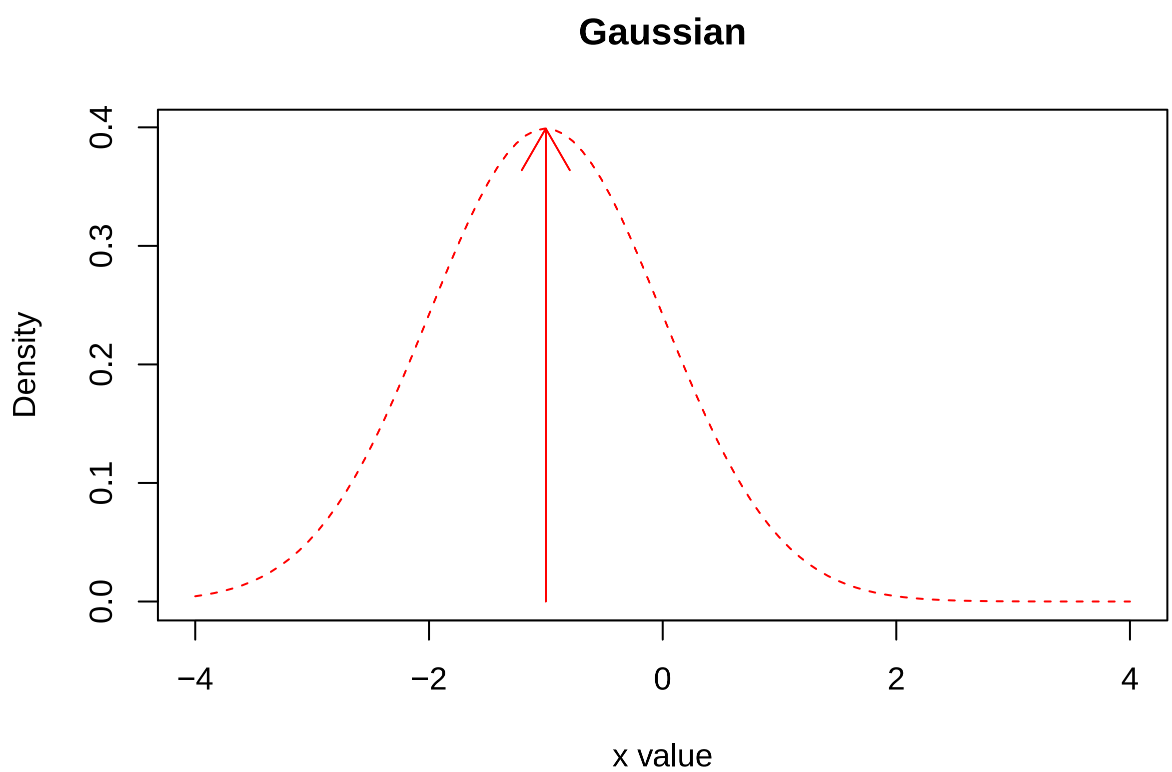
For this trivial case we will have a problem, that will be shown later. For now let fix sd =1 and try to maximize just for u. So we have:

maximize u in p(x | u, sd = 1)

Well… looking again to the Gaussian distribution we see that the maximum value is at u. So if we want to maximize p(x | u, sd = 1), we just have to make u = x.

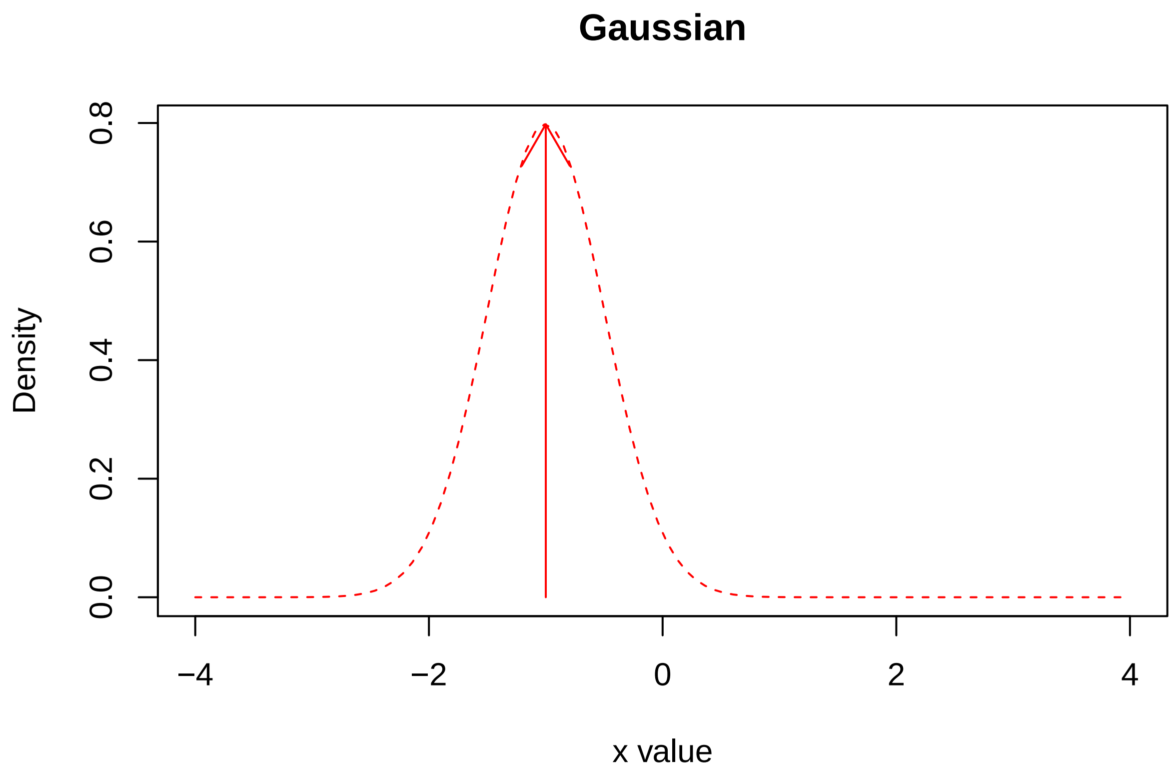
So in this case we have:

```
plotGaussian(-1,1,"red")
arrows(-1,0,-1,dnorm(-1,-1,1), col = "red")
```



But we still have one problem. We can improve the likelihood in this case decreasing the sd. If we chose a sd = 0.5, for example, we will have:

```
plotGaussian(-1,0.5,"red")
arrows(-1,0,-1,dnorm(-1,-1,0.5), col = "red")
```


Gaussian
Density
0.0
0.2
0.4
0.6
0.8
−4
−2
0
2
4
x value