P_M4_1

September 26, 2021

1 Module 4 Peer Review Assignment

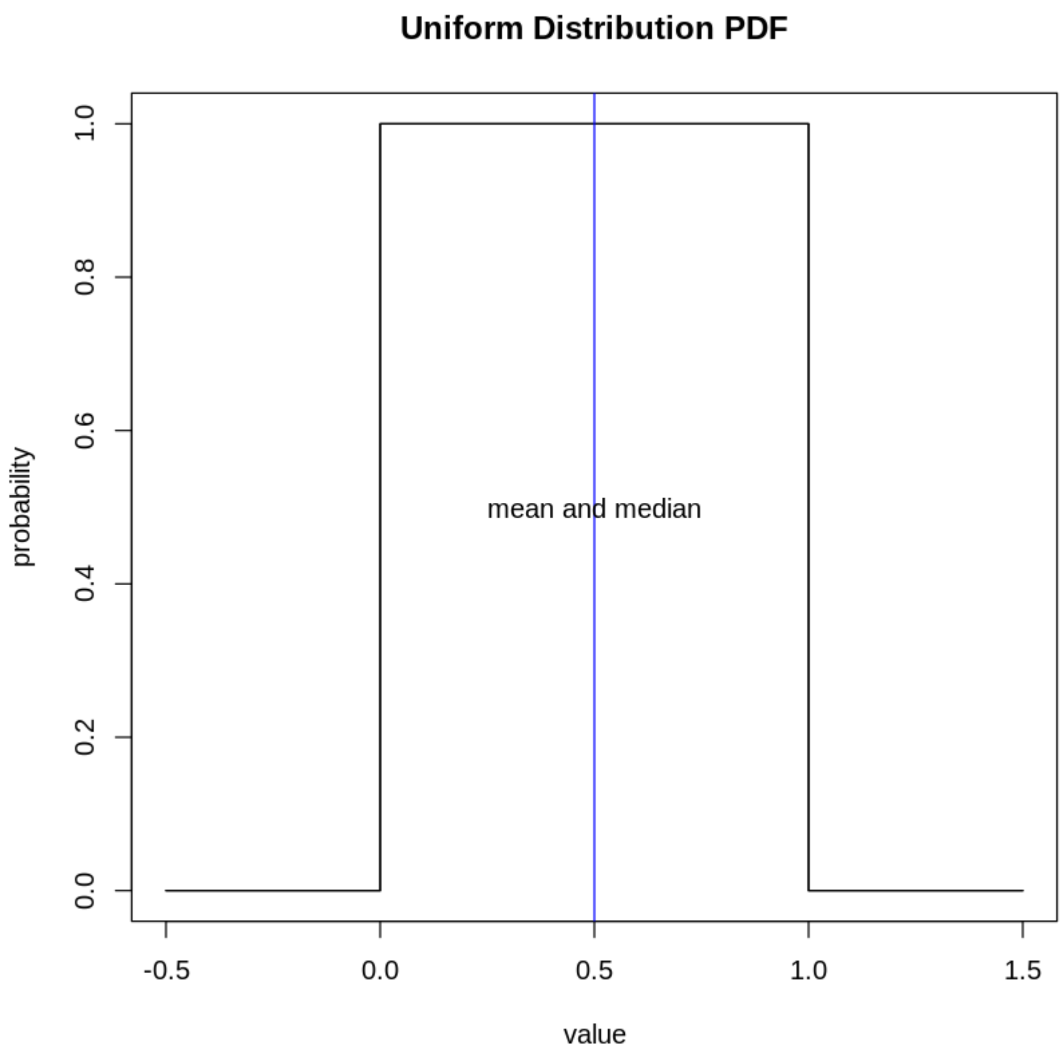
2 Problem 1

A continuous random variable with cumulative distribution function F has the median value m such that $F(m) = 0.5$. That is, a random variable is just as likely to be larger than its median as it is to be smaller. A continuous random variable with density f has the mode value x for which $f(x)$ attains its maximum. For each of the following three random variables, (i) state and graph the density function and then (ii) compute the median, mode and mean and show these values on the graph.

a) W which is uniformly distributed over the interval $[a, b]$, for some value $a, b \in \mathbb{R}$.

[49]:
```
# Your Code Here

a = 0
b = 1
mean = 0.5*(a+b)
median = 0.5*(a+b)
print(mean)
print(median)

data = c(0, 0, 1, 1, 0,0)
labs = c(-0.5,0,0,1,1, 1.5)
plot(x = labs, y = data, type='l', main='Uniform Distribution PDF',␣
↪xlab='value', ylab ='probability')
lines(x = labs, y = data, type='l')

abline(v=0.5, col='blue')
text(x=0.5, y=0.5, labels='mean and median')
```

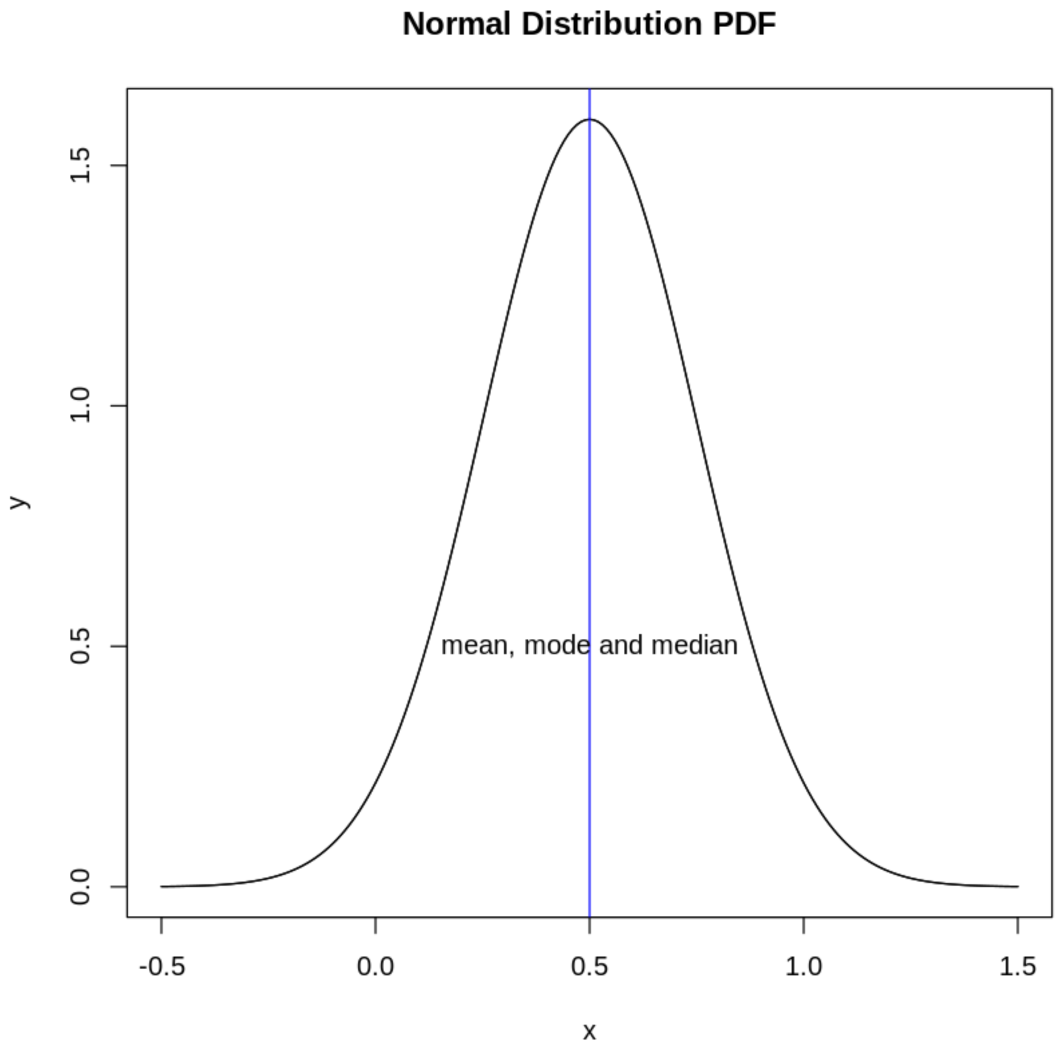
```
[1] 0.5
[1] 0.5
```



This is a uniform random variable between 0 and 1. The probability of any value < 0 or > 1 is 0. The mean and median are both 0.5. The mode of the uniform distribution is any value between the limits of the distribution since each value is equally likely.

b) X which is normal with parameters μ and σ^2, for some value $\mu, \sigma^2 \in \mathbb{R}$.

[47]:
```
# Your Code Here
x <- seq(-0.5, 1.5, by = .01)
y <- dnorm(x, mean = 0.5, sd = 0.25)
plot(x=x,y=y, type='l', main='Normal Distribution PDF')
lines(x=x,y=y, type='l')
abline(v=0.5, col='blue')
text(x=0.5, y=0.5, labels='mean, mode and median')
```



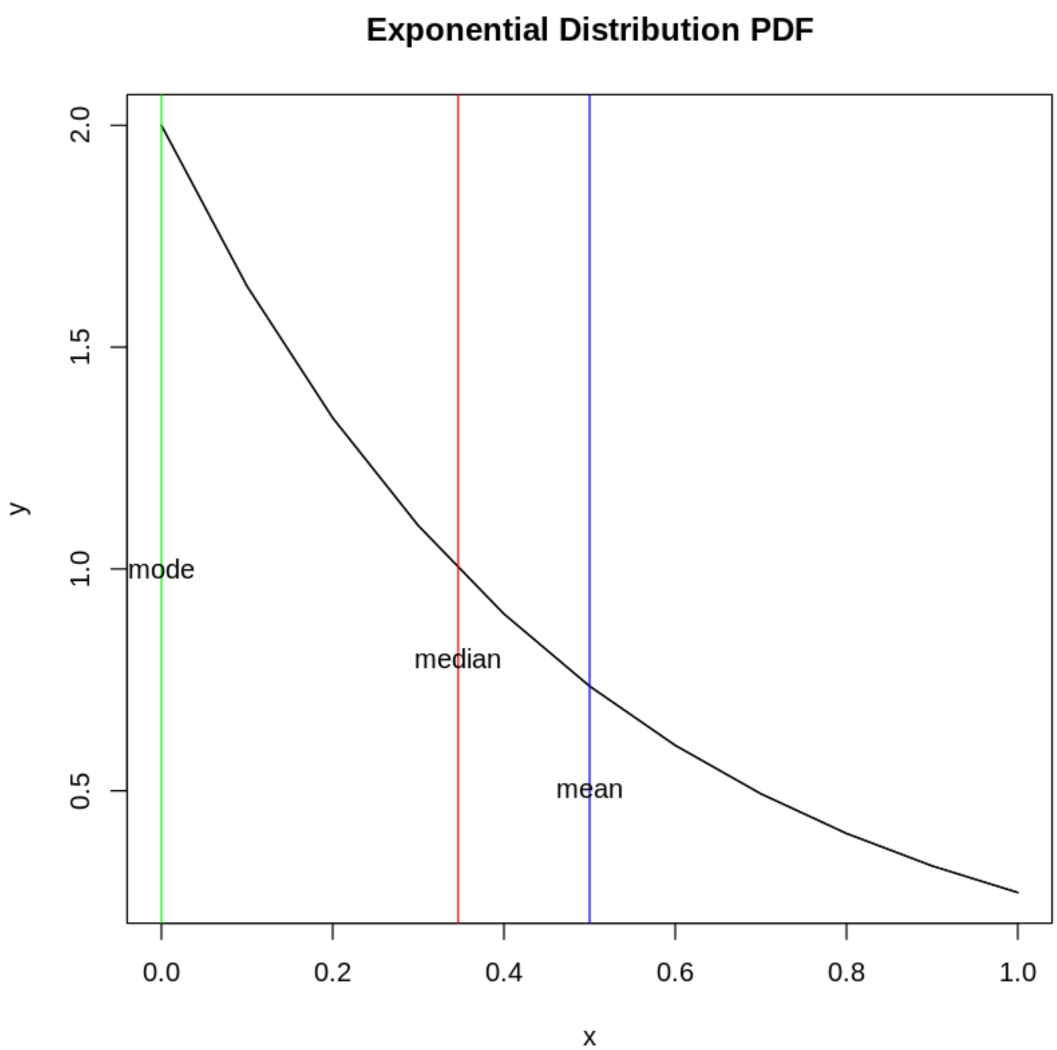
This is a normal or Gaussian random variable PDF. To satisfy the prompt any value of sigma squared is acceptable. In a normal distribution the mean, median and mode are all equal to mu or 0.5 in this case.

c) Y which is exponential with rate $\lambda \in \mathbb{R}$.

[46]:
```
# Your Code Here

lambda = 2
median = log(2)/lambda
print(median)
mean = 1/lambda
print(mean)
```

```
x <- seq(0, 1, by = .1)
y <- dexp(x, rate=lambda)
plot(x,y, type='l', main='Exponential Distribution PDF')
lines(x,y, type='l')

abline(v=0.5, col='blue')
text(x=0.5, y=0.5, labels='mean')
abline(v=median, col='red')
text(x=median, y=0.8, labels='median')
abline(v=0, col='green')
text(x=0, y=1, labels='mode')
```

```
[1] 0.3465736
[1] 0.5
```



This is an exponential random variable PDF. The mean is 1/lambda. For the mean to be 0.5, lambda = 2. The median is ln(2) / lambda or 0.347. The mode of an exponential distribution is 0.

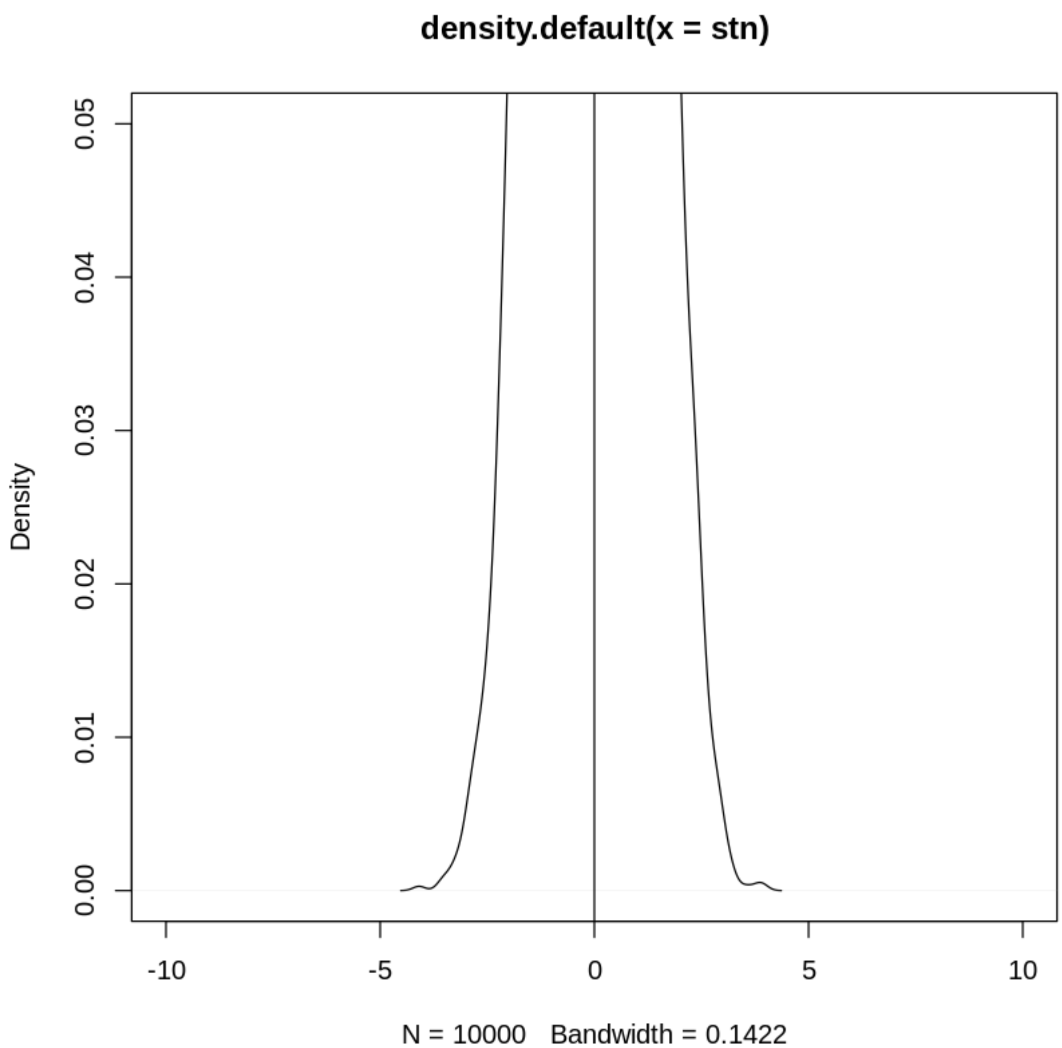
3 Problem 2

For this problem, we're going to visualize what's happening when we go between different normal distributions.

Part A)

Draw at least 10000 samples from the standard normal distribution $N(0, 1)$ and store the results. Make a density histogram of these samples. Set the x-limits for your plot to $[-10, 10]$ and your y-limits to $[0, 0.5]$ so we can compare with the plots we'll generate in **Parts B-D**.

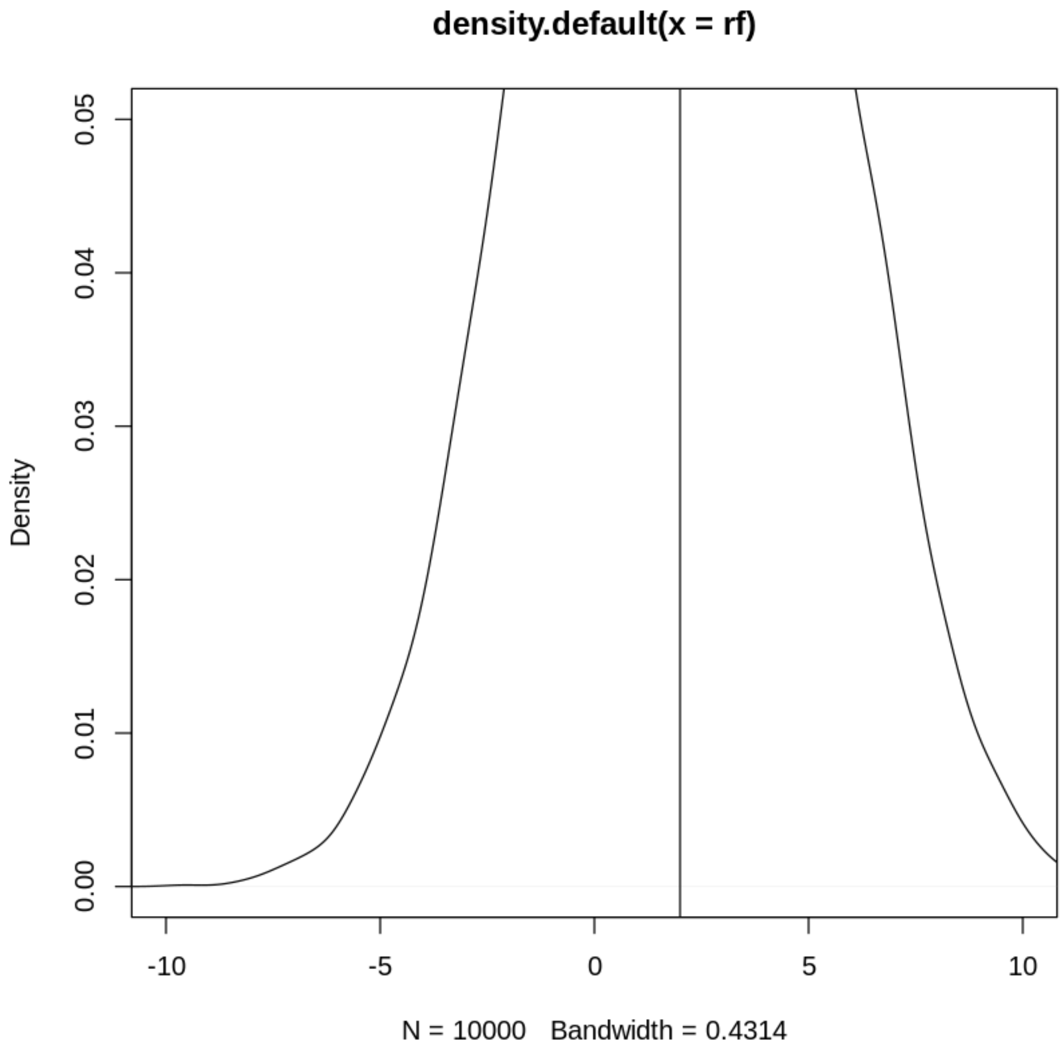
[86]:
```
# Your Code Here
stn <- rnorm(10000, mean =0, sd=1)

# Note I'm using a desnisty plot here, not a normal binned histogram.
d <- density(stn)
plot(d, xlim =c(-10, 10), ylim =c(0, 0.05))
abline(v=0)
```



Part b) Now generate 10000 samples from a $N(2, 3)$ distribution and plot a histogram of the results, with the same x-limits and y-limits. Does the histogram make sense based on the changes to parameters?

Note: Be careful with the parameters for `rnorm`. It may help to check the documentation.

[68]:
```
rf <- rnorm(10000, mean =2, sd=3)
d2 <- density(rf)
plot(d2, xlim =c(-10, 10), ylim =c(0, 0.05))
abline(v=2)
```



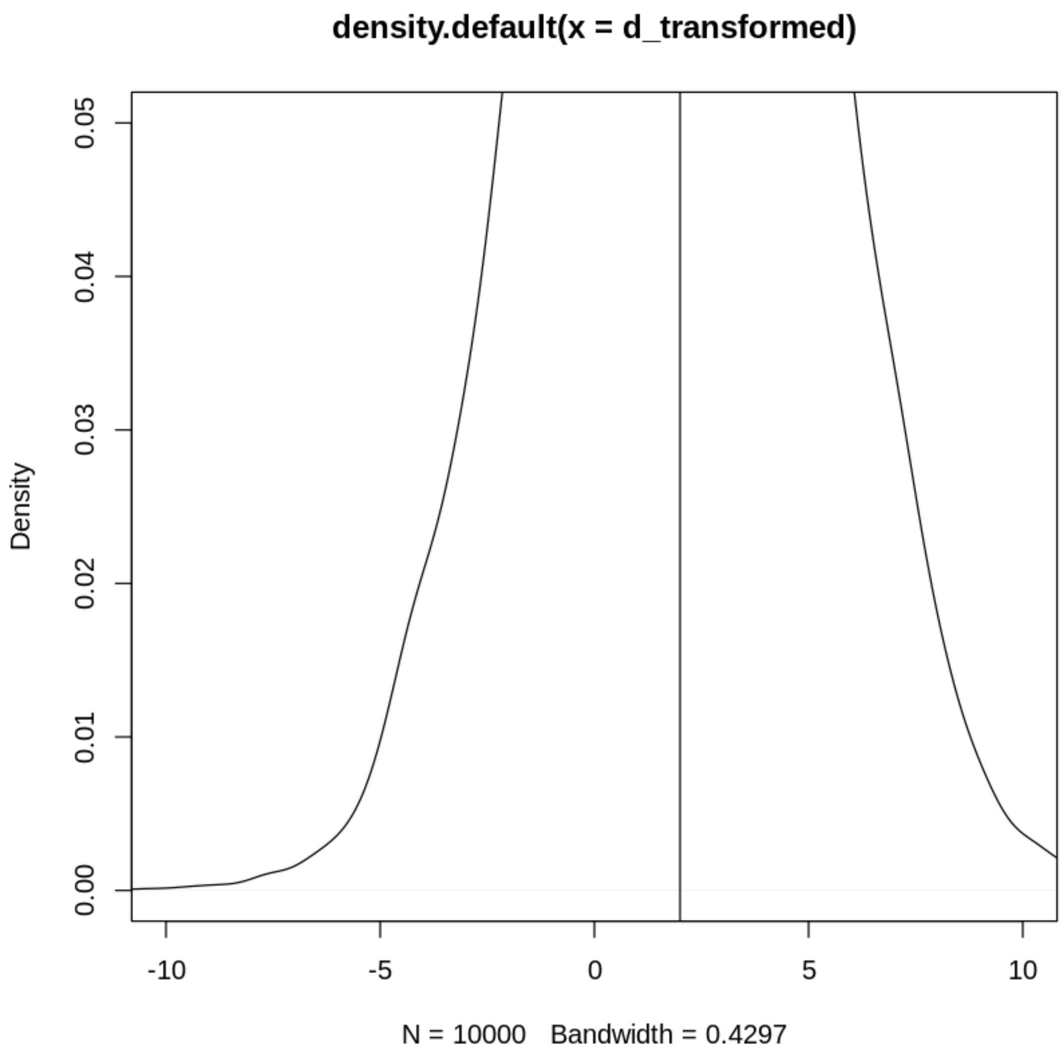
Since we increased the variance and moved the mean, it makes sense that the density plot would be wider and centered around the new mean of 2.

Part c)

Suppose we are only able to sample from the standard normal distribution $N(0, 1)$. Could we take those samples and perform a simple transformation so that they're samples from $N(2, 3)$? Try this, and plot another histogram of the transformed data, again with the same axes. Does your histogram based of the transformed data look like the histogram from **Part B**?

[85]:
```
# Your Code Here
# Standard Normal
stn <- rnorm(10000, mean =0, sd=1)
d_transformed = (stn*3)+2
```

```
d3 =density(d_transformed)
plot(d3, xlim =c(-10, 10), ylim =c(0, 0.05))
abline(v=2)
```

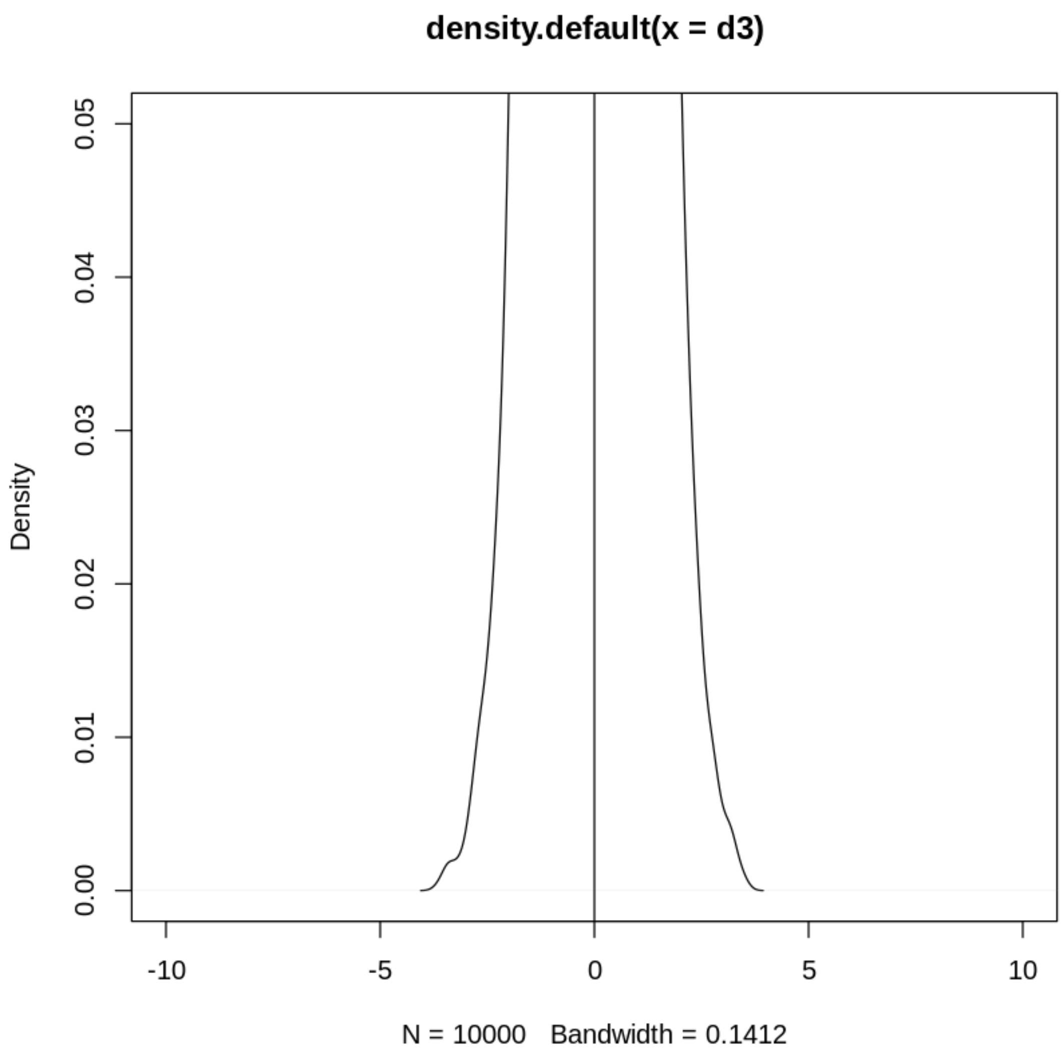


Yes, it is quite simple to transform normal distributions into another normal distribution. Multiplying the standard normal values by the new standard deviation and adding the new mean creates a new normal distribution. This plot of the transformed distribution here looks much like the plot of the N(2,3) distribution we created in Part B.

Part d)

But can you go back the other way? Take the $N(3, 4)$ samples from **Part B** and transform them into samples from $N(0, 1)$? Try a few transformations and make a density histogram of your transformed data. Does it look like the plot of $N(0, 1)$ data from **Part A**?

[87]:
```
mean = 3
sd = 4
rf <- rnorm(10000, mean =mean, sd=sd)

d3 = (rf-mean)/sd
d3 =density(d3)
plot(d3, xlim =c(-10, 10), ylim =c(0, 0.05))
abline(v=0)
```



Note what the question asks for (N(3,4)) is actually different than the distibution in part B, which was N(2,3). The whole concept of the Standard Normal Distribution and z-scores is based on our ability to transform any normal distribution into the standard normal. To transform any normal distribution to standard normal simply subtract the mean and divide by the stadard deviation, the inverse of what we did in Part C. This plot looks much like the standard normal plot we created in part A.