Untitled

Pegah Abedini

2023-12-22

## Whats the emprical model?

### Definition

* The empirical model is a discrete distribution based on a sample of size that assigns probability to each data point.

## Example 4.7

Consider a sample of size 8 in which the observed data points were 3, 5, 6, 6, 6, 7, 7, and 10. The empirical model then has probability function

c(1/8)

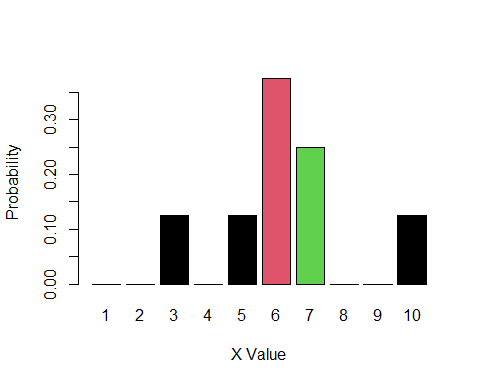
## [1] 0.125

The empirical distribution is . Each data point contributes probability to the probability function, so the parameters are the observations in the data set that produced the empirical distribution

## Visualizing Emperical Model in Example 4.7

now we are going to plotting the Probability mass function (emperical model) that we illustrated in last slide (example 4.7 ).

p = c(0,0,0.125,0,0.125,0.375,0.25,0,0,0.125)  
barplot(p , col = c(0,0,1,0,1,2,3,0,0,1) ,   
 names.arg = c(1:10) , xlab = "X Value" ,ylab = "Probability")



## Whats Kernel Smoothing Model?

Another example of a data-dependent model is the kernel smoothing model.

### Definition

Rather than placing a probability mass of at each data point, . This piece is centered at the data point so that this model follows the data, but not perfectly. It provides some smoothing when compared to the empirical distribution.

## Making a uniform Kernel Smoothing Model for Example 4.7 Data with bandwidth 2.

Now we are going to make a .

So we need to start with a good standpoint to understand the uniform Kernel smoothing models. We know that a . and we know that when we are telling a .

## Making a uniform Kernel Smoothing Model for Example 4.7 Data with bandwidth 2 for Point x = 3

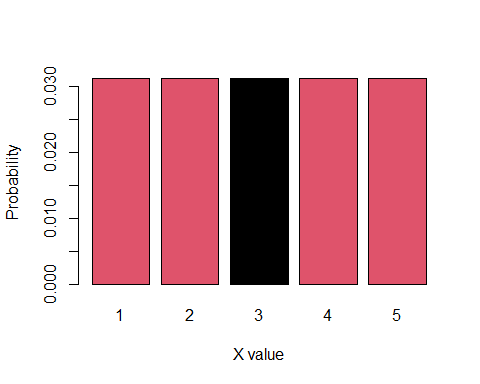
So if we want to plot a uniform kernel smoothing with bandwidth 2, just for point . it means that we want to that the probability in all of this interval is equal with each other and its equal to

that 1/4 is the probability of a . and the 0.125 is equal to a x = 3 probability that we distributed that in our interval.

## Plotting a distributed point x = 3 with new model

now we understand that , so we plot it.

p1\_5 = rep(0.125/4 , 5)  
barplot(p1\_5 , col = c(2,2,1,2,2) , names.arg = 1:5 ,  
 xlab = "X value" , ylab = "Probability")



## Making a uniform Kernel Smoothing Model for Example 4.7 Data with bandwidth 2 for Point x = 7

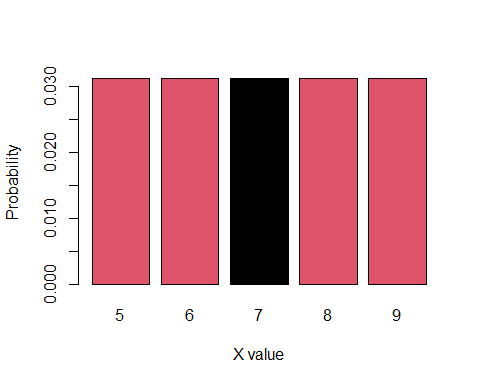
So if we want to plot a uniform kernel smoothing with bandwidth 2, just for point . it means that we want to that the probability in all of this interval is equal with each other and its equal to

that 1/4 is the probability of a . and the 0.25 is equal to a x = 7 probability that we distributed that in our interval.

## Plotting a distributed point x = 7 with new model

now we understand that , so we plot it.

p5\_9 = rep(0.25/4 , 5)  
barplot(p1\_5 , col = c(2,2,1,2,2) , names.arg = 5:9 ,  
 xlab = "X value" , ylab = "Probability")

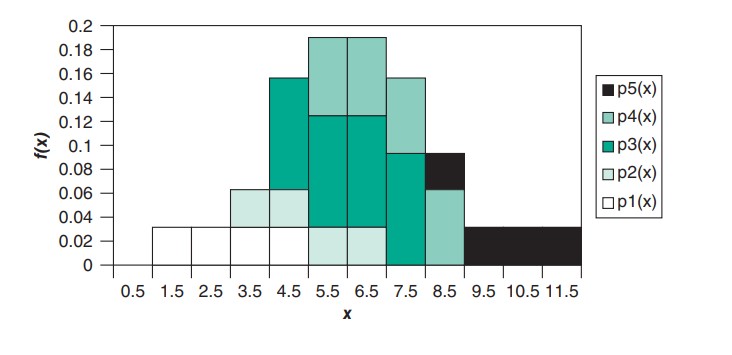


## Uniform Kernel smoothing with bandwidth 2 formula

The probability density function is: and we know that the uniform kernel smoothing with bandwidth 2, is :

where the sum is taken over the five points where the original model has positive probability.

## Plot of Example 4.7



## Note

Note that both the kernel smoothing model and the empirical distribution can also be written as . The reason why these models are classified separately is that .

## End