Kernel Smoothing Model

Mehrab Atighi

2022-11-17

Whats the emprical model?

Definition

- The empirical model is a discrete distribution based on a sample of size n that assigns probability $1/n$ 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

$$p(x) = \begin{cases} 0.125, & x = 3 \\ 0.125, & x = 5 \\ 0.375, & x = 6 \\ 0.25, & x = 7 \\ 0.125, & x = 10 \end{cases}$$

```r
c(1/8)
```

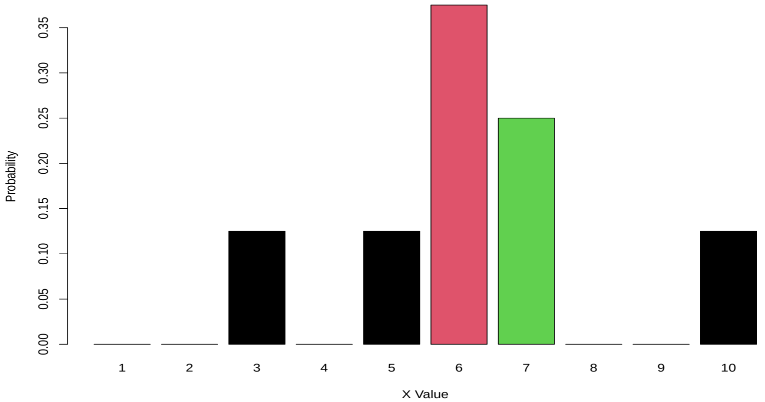
```
## [1] 0.125
```

The empirical distribution is a data-dependent distribution. Each data point contributes probability $1/n$ to the probability function, so the n parameters are the n 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 $1/n$ at each data point, a continuous density function with area $1/n$ replaces the 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 uniform kernel with bandwidth equal 2.

So we need to start with a good standpoint to understand the uniform Kernel smoothing models.
We know that a uniform Distribution have equal probability in all of that Domain. and we know that when we are telling a bandwidth is equal to 2, it means that if we are in point x, we are speaking about this interval :

$$\forall x \in [x-2, x+2]$$

.

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 $x = 3$.
it means that we want to distribute 0.125 to the $[1, 5]$ interval. that the probability in all of this interval is equal with each other and its equal to

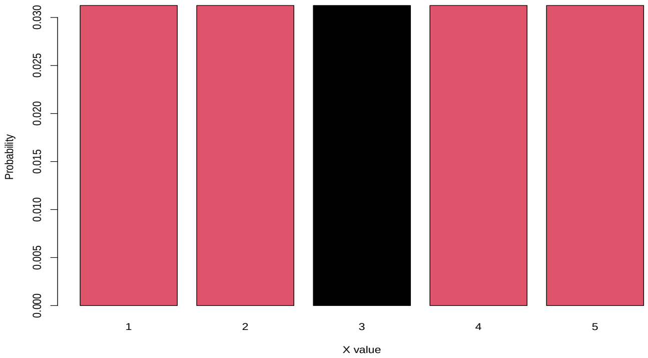
$$(1/4) * 0.125 = 0.03125$$

that $1/4$ is the probability of a uniform distribution in an interval with length equal 4. 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 $Pr(1 \leq x \leq 5) = 0.03125$, 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 $x = 7$.
it means that we want to distribute 0.25 to the $[5, 9]$ interval. that the probability in all of this interval is equal with each other and its equal to

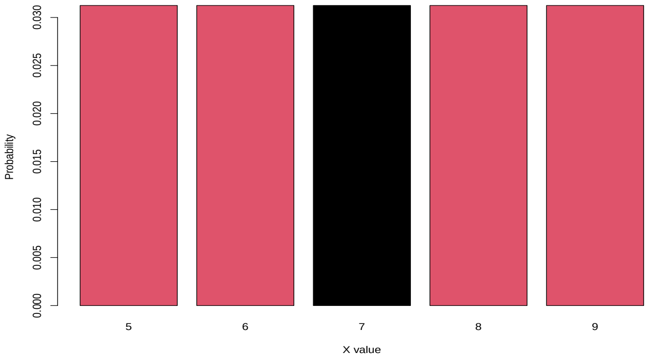
$$(1/4) * 0.25 = 0.0625$$

that $1/4$ is the probability of a uniform distribution in an interval with length equal 4. 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 $Pr(5 \leq x \leq 9) = 0.0625$, 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:

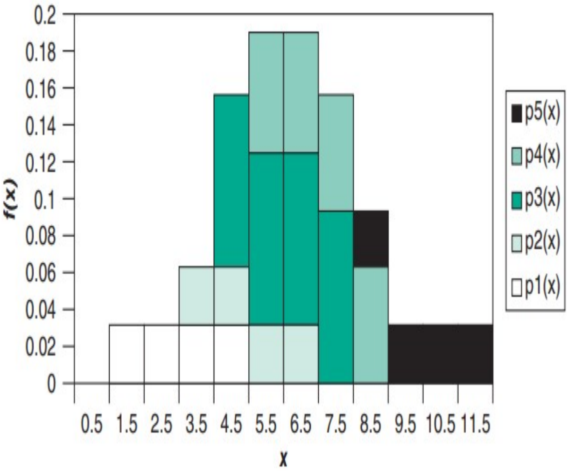
$$f(x) = \sum_{j=1}^{5} p(x_j) K_j(x)$$

and we know that the uniform kernel smoothing with bandwidth 2, is :

$$K_j(x) = \begin{cases} 0, & |x - x_j| > 2, \\ 0.25, & |x - x_j| \leq 2 \end{cases}$$

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 mixture distributions. The reason why these models are classified separately is that the number of components relates to the sample size rather than to the phenomenon and its random variable.

- Thanks for your attention