Colm Mooney - 20325583

#3. The effective channel length (in microns) is measured for 1225 field effect transistors. The data is

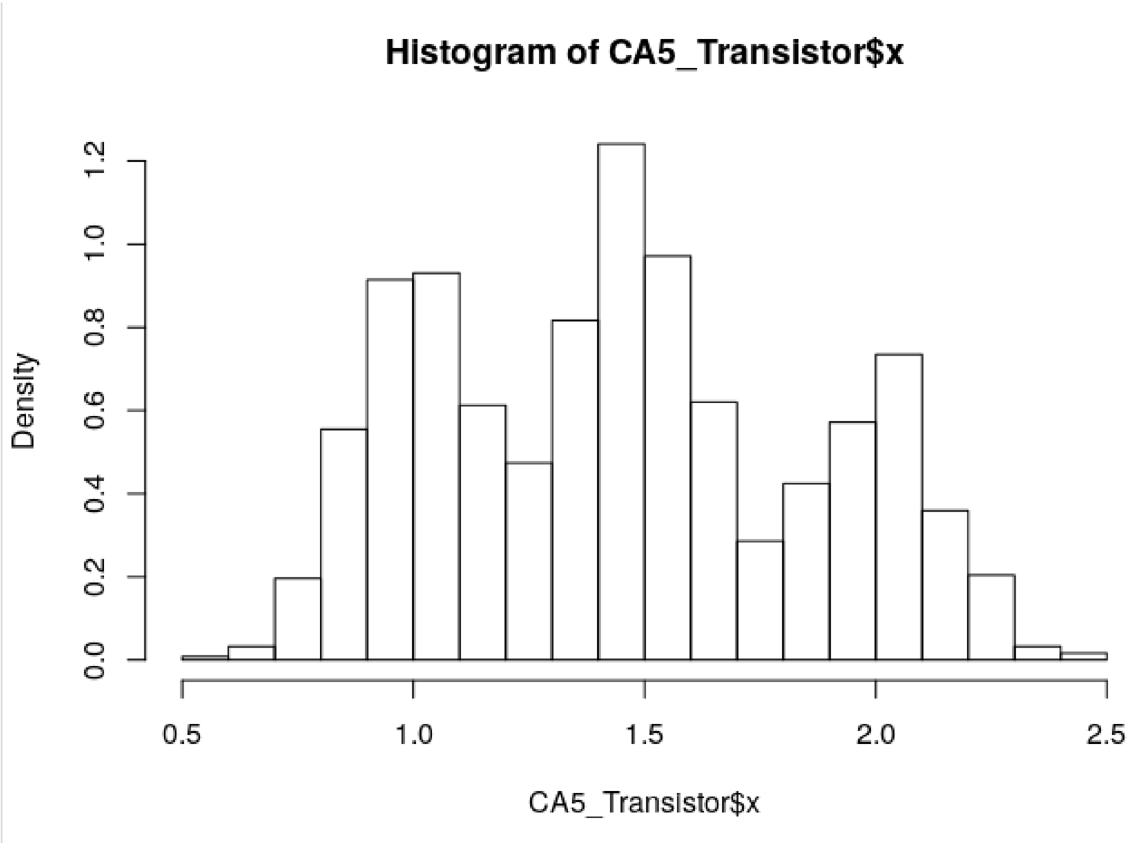
#recorded in CA5 Transistor.csv. Use kernel density estimation to find a suitable density estimate for this data.

#(a) First, generate a density histogram for the length data. Use freq=FALSE and breaks=16.

library(readr)

CA5_Transistor <- read_csv("SharedFiles/ST204/Data/CA5_Transistor.csv")

hist(CA5_Transistor$x, freq = FALSE, breaks = 16)



#(b) Then, examine various bandwidths to identify a suitable value (using the Gaussian kernel).Justify your choice in writing.

par(mfrow=c(3,2))

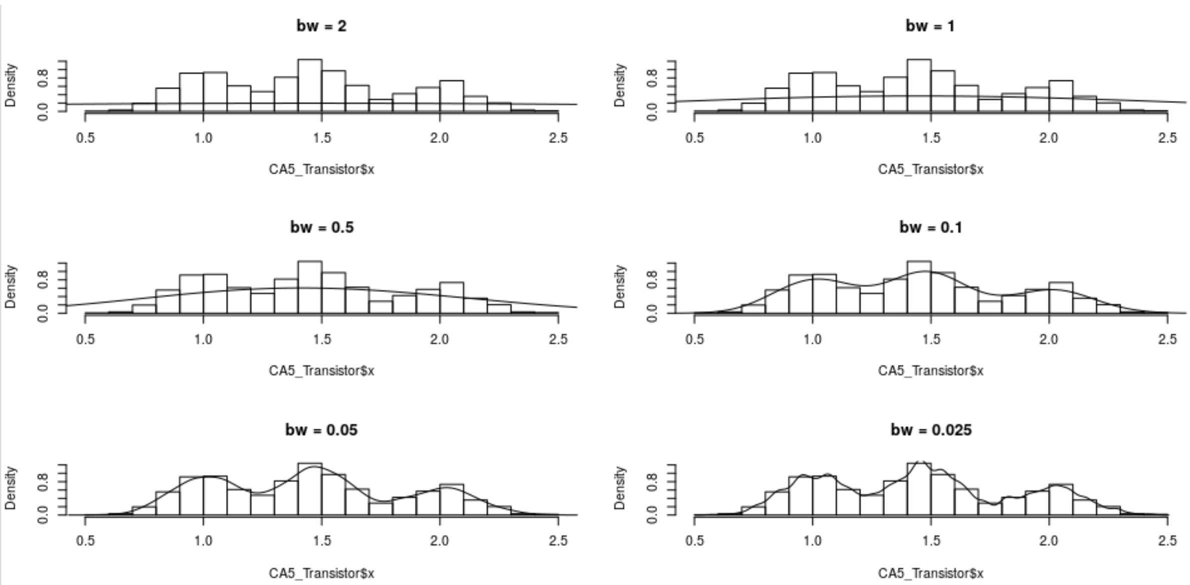
hist(CA5_Transistor$x, freq = FALSE, breaks = 16, main = "bw = 2")

lines(density(CA5_Transistor$x, bw = 2, kernel = "gaussian"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16,main = "bw = 1")

lines(density(CA5_Transistor$x, bw = 1, kernel = "gaussian"))

```
hist(CA5_Transistor$x, freq = FALSE, breaks = 16,main = "bw = 0.5")

lines(density(CA5_Transistor$x, bw = 0.5, kernel = "gaussian"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16,main = "bw = 0.1")

lines(density(CA5_Transistor$x, bw = 0.1, kernel = "gaussian"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16,main = "bw = 0.05")

lines(density(CA5_Transistor$x, bw = 0.05, kernel = "gaussian"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16,main = "bw = 0.025")

lines(density(CA5_Transistor$x, bw = 0.025, kernel = "gaussian"))

par(mfrow=c(1,1))

#The choice I choose is bw = 0.05, I think it is the BW most suitable value as it's the line that keeps most in form with the graph without losing any of

#the smoothness that values like 0.025 has. it’s likely the True X.
```

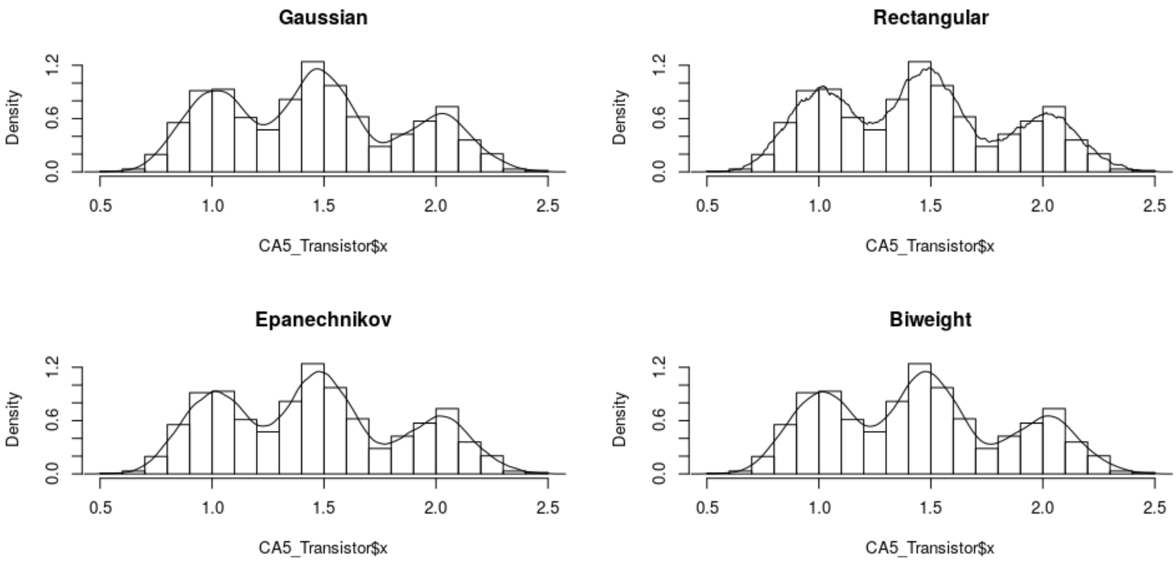



```
#(c) Examine various kernels to identify a suitable one. Ensure that you also try some of the kernel options that were not discussed in class. Justify your choice in writing.
```

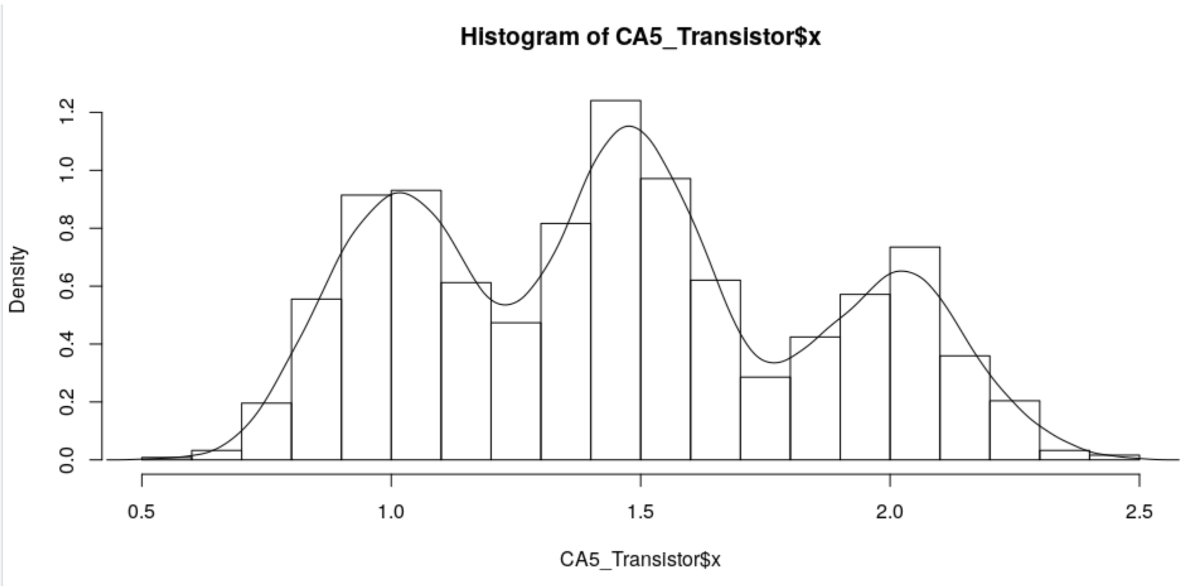

#(c) Examine various kernels to identify a suitable one. Ensure that you also try some of the kernel options that were not discussed in class. Justify your choice in writing.


```
par(mfrow=c(2, 2))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16, main = "Gaussian")

lines(density(CA5_Transistor$x, bw = 0.05, kernel = "gaussian"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16, main = "Rectangular")

lines(density(CA5_Transistor$x, bw = 0.05, kernel = "rectangular"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16, main = "Epanechnikov")
```

```
lines(density(CA5_Transistor$x, bw = 0.05, kernel = "epanechnikov"))

hist(CA5_Transistor$x, freq = FALSE, breaks = 16, main = "Biweight")

lines(density(CA5_Transistor$x, bw = 0.05, kernel = "biweight"))

par(mfrow=c(1, 1))

#I think biweight is the most suitable of the choices as it's line is the one that is most accurate with the histogram.
```




```
#(d) Superimpose the chosen density function on the histogram generated in (a).

hist(CA5_Transistor$x, freq = FALSE, breaks = 16)

lines(density(CA5_Transistor$x, bw = 0.05, kernel = "biweight"))
```



#Q4

$x = 9.5$ bw $= 2.5$

$n = 14$ Gaussian: $w(t) = \frac{1}{\sqrt{2\pi}} e^{(-t^2/2)}$

data $= 4, 7, 14, 12, 13, 8, 14, 7, 5, 6, 9, 21, 11, 16$

(4) $\hat{f}(x) = \frac{1}{\Delta n} \sum_{i=1}^{n} w\left(\frac{x - x_i}{\Delta}\right)$, $\hat{f}(9.5) = \frac{1}{14(2.5)}\left[w\left(\frac{9.5-4}{2.5}\right) + w\left(\frac{9.5-7}{2.5}\right) + \ldots + w\left(\frac{9.5-16}{2.5}\right)\right]$

$\frac{1}{35}\left[w(2.2) + w(1) + w(-1.8) + w(-1) + w(-1.4) + w(0.6) + w(-3.8) + w(1) + w(1.8) + w(1.4) + w(0.2) + w(-4.6) + w(-0.6) + w(-2.6)\right]$

$\frac{1}{35\sqrt{2\pi}}\left[e^{(-2.2^2/2)} + e^{(1^2/2)} + e^{(-(-1.8)^2/2)} + e^{(-(-1)^2/2)} + e^{(-(-1.4)^2/2)} + e^{(-(0.6)^2/2)} + e^{(-(-3.8)^2/2)} + e^{(-(1)^2/2)} + e^{(-(1.8)^2/2)} + e^{(-(1.4)^2/2)} + e^{(-(0.2)^2/2)} + e^{(-(-4.6)^2/2)} + e^{(-(-0.6)^2/2)} + e^{(-(-2.6)^2/2)}\right]$

$\frac{1}{35\sqrt{2\pi}}\left[0.0889216 + 0.6065307 + 0.1978987 + 0.6065307 + 0.3753111 + 0.8352702 + e^{(-(-3.8)^2/2)} + 0.6065307 + 0.1978987 + 0.3753111 + 0.9801987 + e^{(-(-4.6)^2/2)} + 0.8352702 + 0.0340475\right]$

$\frac{1}{35\sqrt{2\pi}}\left[5.7397199 + e^{(-(-3.8)^2/2)} + e^{(-(-4.6)^2/2)}\right] = \frac{1}{35\sqrt{2\pi}}\left[5.740477122\right] = 0.06543197239$

$= 0.06543197239$

#Q5

(5) Use ~~Epa~~ Epanechnikov kernal: $w(t)=\frac{3}{4}(1-t^2)$, $|t|\le 1$, $n=14$ $bw=4.5$

$$f(9.5)=\frac{1}{14\times 4.5}\sum_{i=1}^{14} w\left(\frac{9.5-x_i}{4.5}\right)$$

~~data = (9, 10, 11, 13, 15, 14, 18, 17, 14, 13, ...)~~

data = (4, 7, 14, 12, 13, 8, 19, 7, 5, 6, 9, 21, 11, 16)

Answer should be: $\frac{38}{567}$

$$= \frac{1}{63}\left[w\left(\frac{9.5-4}{4.5}\right)+w\left(\frac{9.5-7}{4.5}\right)+\ldots+w\left(\frac{9.5-16}{4.5}\right)\right]$$

$$= \frac{1}{63}\Big[w(11/9) \quad w(5/9)+w(-1)+w(-5/9)+w(-\tfrac{7}{9})+w(\tfrac{1}{3})+w(-\tfrac{19}{9})+w(\tfrac{5}{9})+w(1)+w(\tfrac{7}{9})$$

$$+w(\tfrac{1}{9})+w(-\tfrac{23}{9})+w(-\tfrac{1}{3})+w(-\tfrac{13}{9})\Big]$$

$$\frac{1}{63}\Big[\tfrac{3}{4}(1-(\tfrac{11}{9})^2)+\tfrac{3}{4}(1-(\tfrac{5}{9})^2)+\tfrac{3}{4}(1-(-1)^2)+\tfrac{3}{4}(1-(-\tfrac{5}{9})^2)+\tfrac{3}{4}(1-(-\tfrac{7}{9})^2)+\tfrac{3}{4}(1-(\tfrac{1}{3})^2)$$

$$+\tfrac{3}{4}(1-(-\tfrac{19}{9})^2)+\tfrac{3}{4}(1-(\tfrac{5}{9})^2)+\tfrac{3}{4}(1-(1)^2)+\tfrac{3}{4}(1-(\tfrac{7}{9})^2)+\tfrac{3}{4}(1-(\tfrac{1}{9})^2)$$

$$+\tfrac{3}{4}(1-(-\tfrac{23}{9})^2)+\tfrac{3}{4}(1-(-\tfrac{1}{3})^2)+\tfrac{3}{4}(1-(-\tfrac{13}{9})^2)\Big]$$

$$\frac{3}{252}=\frac{1}{84}\Big[(-\tfrac{40}{81})+(\tfrac{56}{81})+0+(\tfrac{56}{81})+(\tfrac{32}{81})+(\tfrac{8}{9})+(-\tfrac{280}{81})+(\tfrac{56}{81})+0+(\tfrac{32}{81})+(\tfrac{80}{81})+(\tfrac{8}{9})+(-\tfrac{448}{81})+(-\tfrac{88}{81})\Big]$$

$$\cancel{\frac{1}{84}\left[(-\tfrac{16}{27})+(-\tfrac{352}{81})\right]=\frac{1}{84}\left[-(\tfrac{400}{81})\right]=-\frac{100}{1701}}$$

$$w(\tfrac{11}{9})+w(5/9)+w(-1)+w(-5/9)+w(-\tfrac{7}{9})+w(\tfrac{1}{3})+w(-\tfrac{19}{9})+w(\tfrac{5}{9})+w(1)$$

$$+w(\tfrac{7}{9})+w(\tfrac{1}{9})+w(-\tfrac{23}{9})+w(-\tfrac{1}{3})+w(-\tfrac{13}{9})$$

$|t|\le 1$

$$w(5/9)+w(-1)+w(-5/9)+w(-\tfrac{7}{9})+w(\tfrac{1}{3})+w(\tfrac{5}{9})+w(1)+w(\tfrac{7}{9})+w(\tfrac{1}{9})+w(-\tfrac{1}{3})$$

$$=\frac{1}{84}\left[\frac{56}{81}+0+\frac{56}{81}+\frac{32}{81}+\frac{8}{9}+\frac{56}{81}+0+\frac{32}{81}+\frac{80}{81}+\frac{8}{9}\right]$$

$$\frac{1}{84}\left[5.6296293\right]=\frac{38}{567}$$