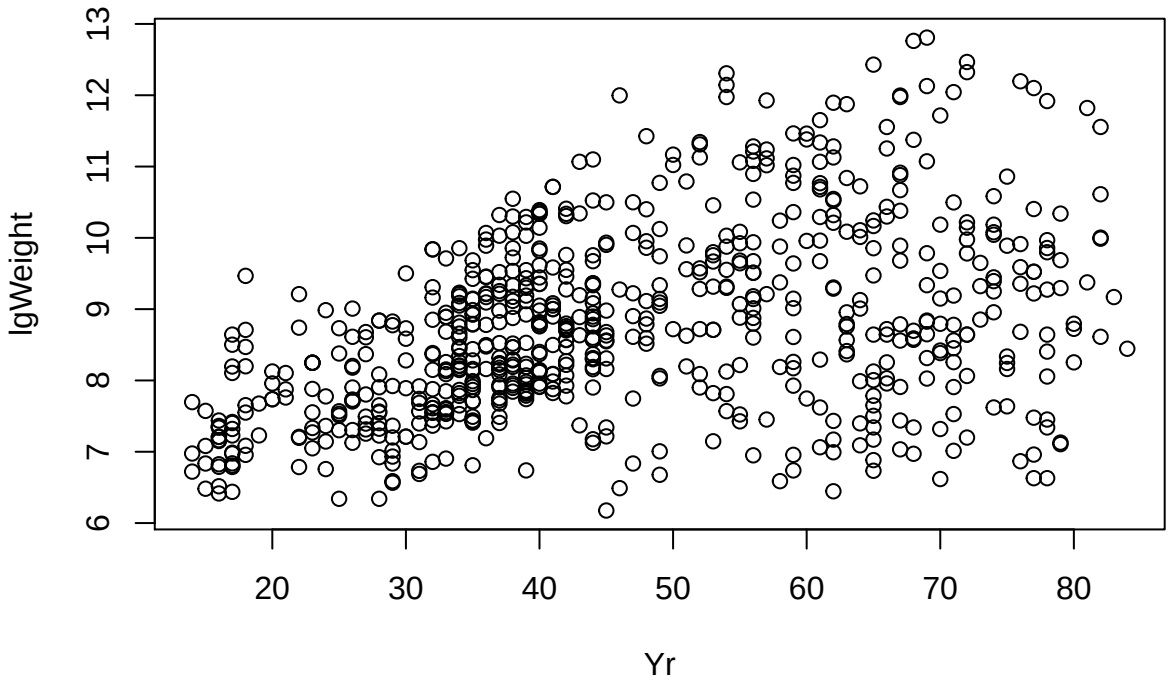
Non-parametric models - Estimating conditional variance

Haoran Mo, Alexandra Yamaui

In this excercise we want to estimate the function that represents the conditional variance σ^2 of the variable lgWeigth from the aircraft dataset given that the explanatory variable (Yr) is equal to a value x, and we will use nonparametric methods to do that.

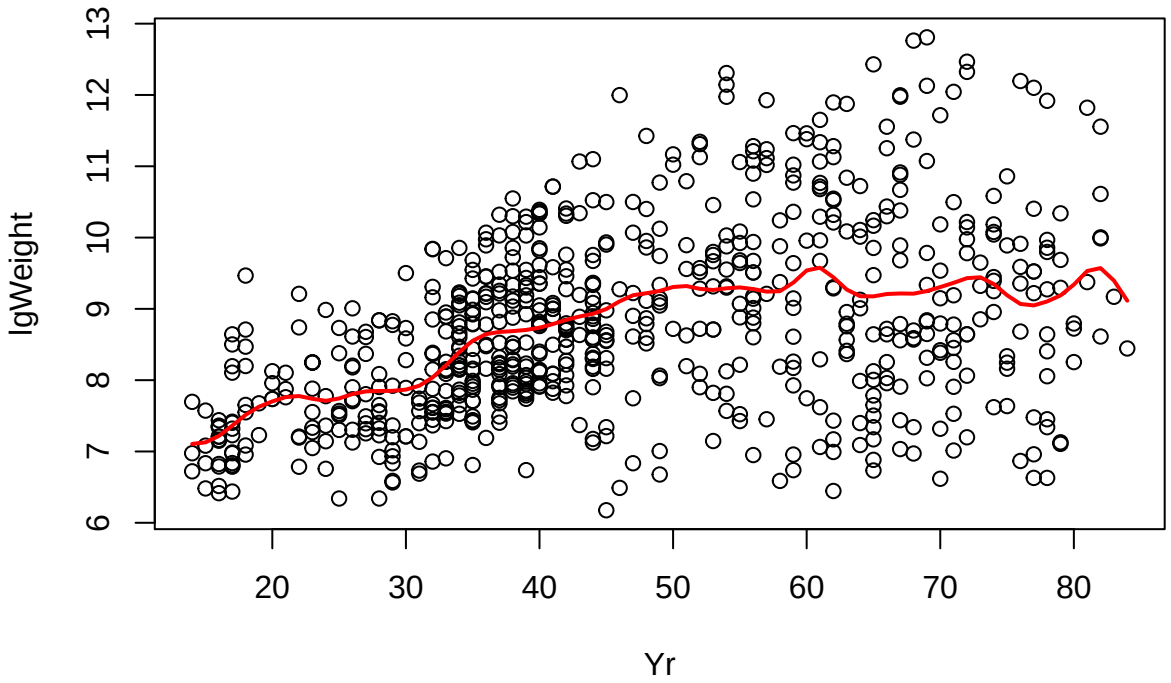
```
data(aircraft)
attach(aircraft)
lgWeight <-log(Weight)
plot(Yr,lgWeight)
```



Initially, we are going to fit a local linear regression model to obtain an estimation $\hat{m}(x)$ of every point x (Yr vs lgWeight). The general idea is to build a grid of intervals (t_i) centered around each point x and estimate a local linear regression in each interval. We are going to try multiple sizes of the intervals and several values for the smoothing parameter h, which controls weight concentration around the points x.

To make the regression function smooth weights are assigned to each pair (t_i, y_1) using a kernel function, which, in this case, is the Normal density function centered at 0, with h as standard deviation.

```
# step 1 Fit a nonparametric regression to data (xi,yi) and save the estimated values m^ (xi).
op <- par(mfrow=c(1,1))
llr <- loc.lin.reg(x=Yr,y=lgWeight, tg=Yr)
plot(Yr,lgWeight)
lines(Yr, llr$mt, col=2, lwd=2)
```



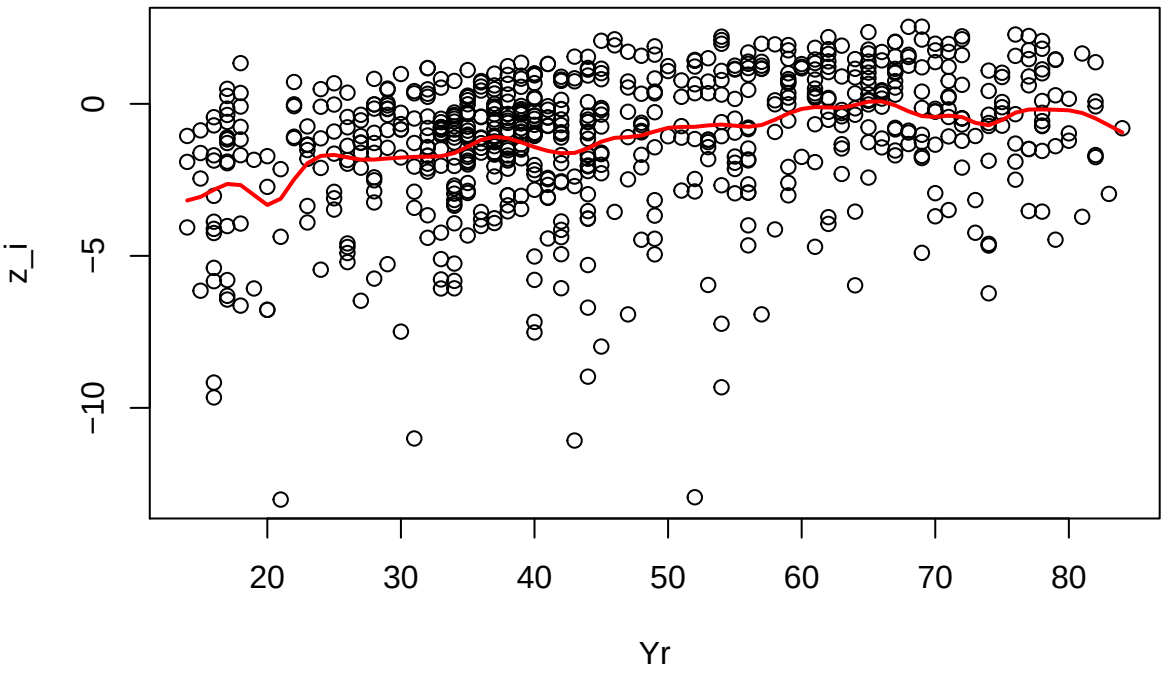

```
# TODO is this regression good enough?
```

Now we are going to apply logarithm over the square residuals $(y_i - \hat{m}(x_i))^2$, which is represents to the variance (the square deviation) of the model.

```
# step 2 Transform the estimated residuals hat.e? = y_i - llr$mt
z_i = log((lgWeight - llr$mt)^2)
```

Now we are to perform a regression over (x_i, z_i) to estimate the (logarithm of the) variance.

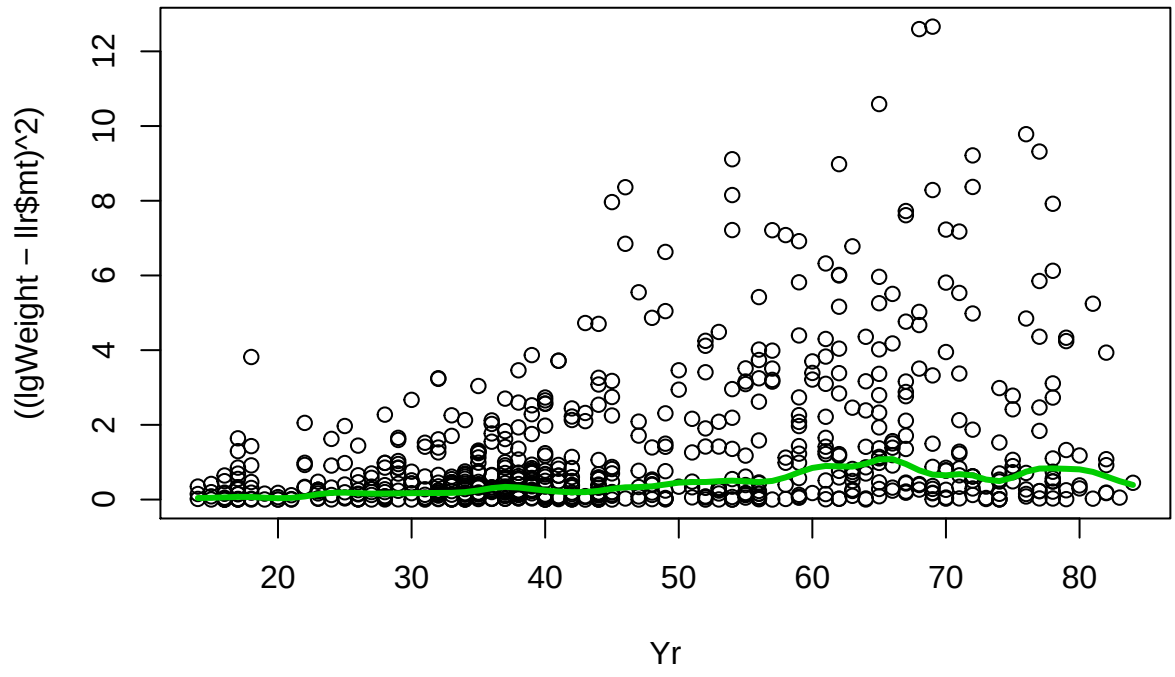
```
# step 3 Fit a nonparametric regression to data (Yr,z_i) and call the estimated function q^(x).
plot(Yr,z_i)
llr2 <-loc.lin.reg(x=Yr,y=z_i, tg=Yr)
lines(Yr, llr2$mt, col=2, lwd=2)
```



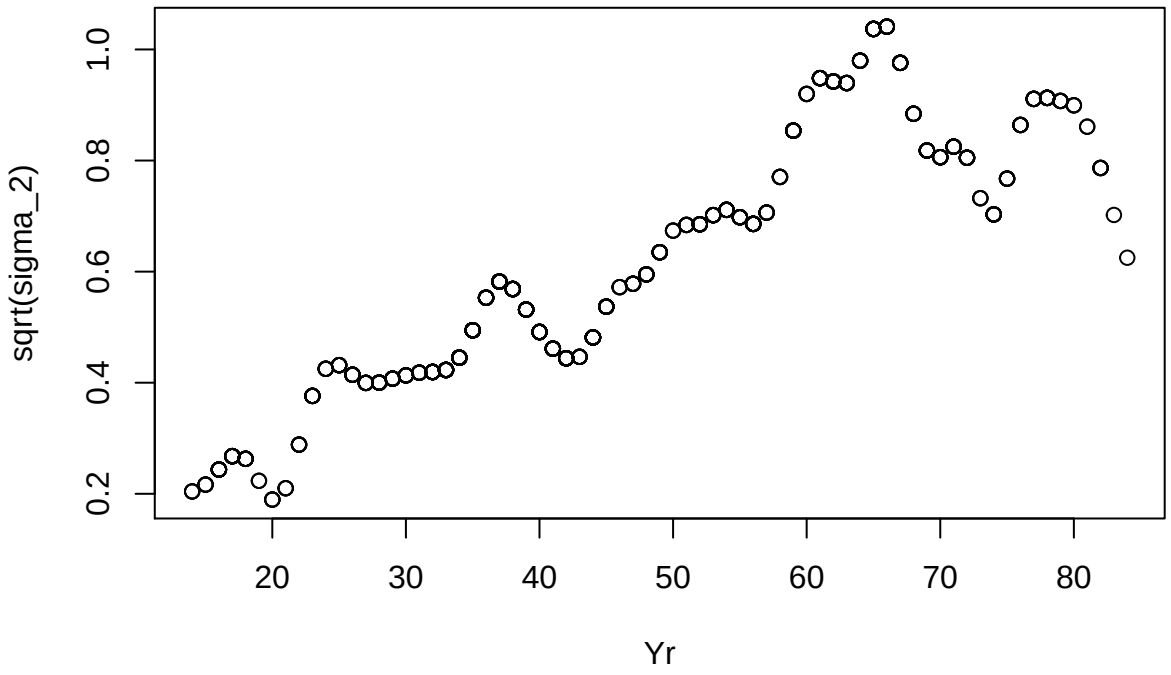
Finally, we can obtain the conditional variance applying exponential

```
# step 4 Estimate sigma_2(x)
sigma_2 =  exp(llr2$mt)
```

```
plot(Yr,((lgWeight - llr$mt)^2))
lines(Yr, sigma_2, col=3, lwd=3)
```




```
plot(Yr,sqrt(sigma_2))
```


sqrt(sigma_2)
0.2
0.4
0.6
0.8
1.0
20
30
40
50
60
70
80
Yr