Homework 10

due Apr 13, 2021

(For questions **1, 2, 3**) Aerial surveys sometimes rely on visual methods to estimate the number of animals in an area. For example, to study snow geese in their summer range areas west of Hudson Bay in Canada, a small aircraft was used to fly over the range, and when a flock of geese was spotted, an experienced person estimated the number of geese in the flock. To investigate the reliability of this method of counting, an experiment was conducted in which an airplane carrying two observers flew over $n = 45$ flocks, and each observer made an independent estimate of the number of birds in each flock. Also, a photograph of the flock was taken so that a more or less exact count of the number of birds in the flock could be made. The resulting data are given in the attached data file `snowgeese.tsv`, which can be downloaded from **HERE**. The three variables in the data sets are `photo` = photo count; `obs1` = aerial count by observer 1; and `obs2` = aerial count by observer 2.

1. Read in the data set and draw the scatter plots of `photo` (response) vs `obs1` (explanatory) and `photo` (response) vs `obs2` (explanatory).

2. Compute the regression of `photo` on `obs1` using `lm()` (That is, find the best fitting line). Overlay the best fitting line to the scatter plot.

3. Suppose that observer 2 made a count of 55 in an additional aerial survey performed later. What is your best guess for the actual count?

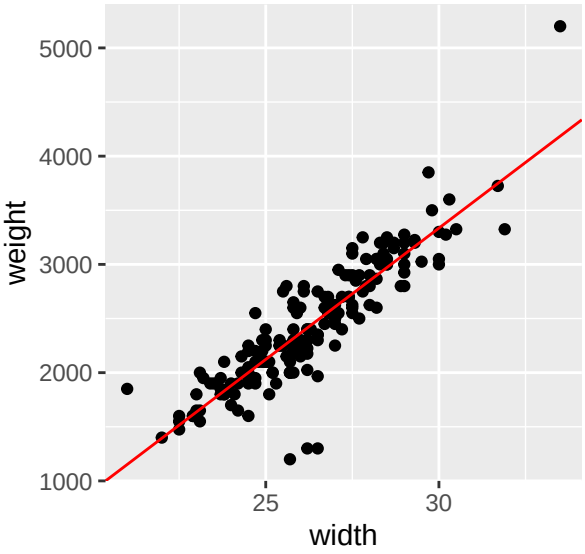
(For questions **4, 5**) Consider the `crabs` data set that we used in class:

```
crabs <- read.table(file = "crabs.tsv", header = T, sep = "\t")
dim(crabs); names(crabs)
```

```
## [1] 173   5
```

```
## [1] "color"  "spine"  "width"  "satell" "weight"
```

4. Obtain the following scatter plot with the best fitting line:



5.. Compute R^2, the coefficient of determination.