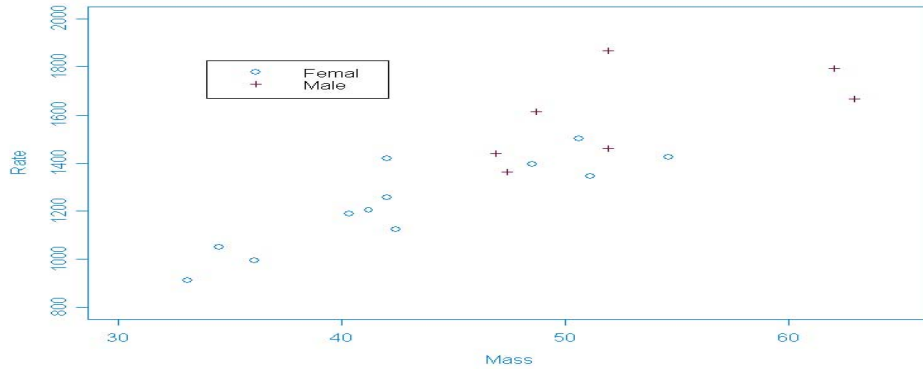


Question 1 (P131, 2.20)

a). The plot is as follow:



Comments: correlation for women should be higher than for men because the values for women are closer to a straight line.

b). For women, $r=0.9$. For men, $r=0.6$.

c). The mean for women is 43.03 and the mean for men is 53.1. This does not affect correlation. Correlation is unaffected by changes in location.

d). Correlation is not affected by the units in the problem.

Question 2 (P214, 1.119)

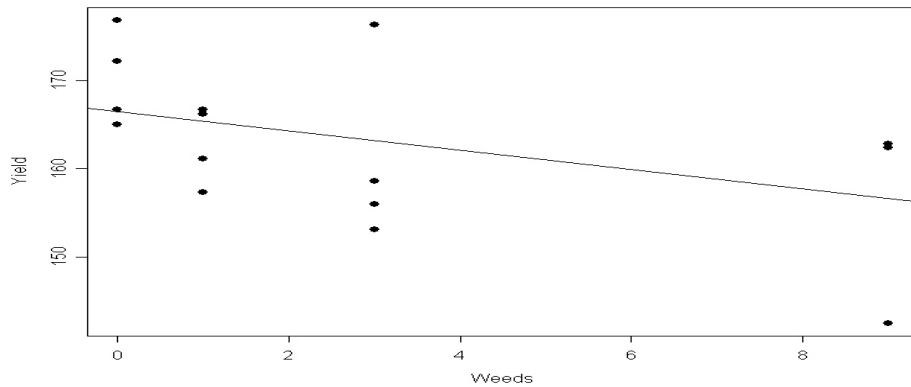
a). The explanatory variable is Weed. The response variable is yield.

b). The side-by-side stemplots of the yields are:

0 Weeds	1 Weeds	3 Weeds	9 Weeds
14	14	14	14 2
14	14	14	14
15	15	15 3	15
15	15 7	15 69	15
16	16 1	16	16 233
16 57	16 67	16	16
17 2	17	17	17
17 7	17	17 6	17

From the above plot, we can see that as weeds increase, yields seem to decrease.

c). The scatterplot is:

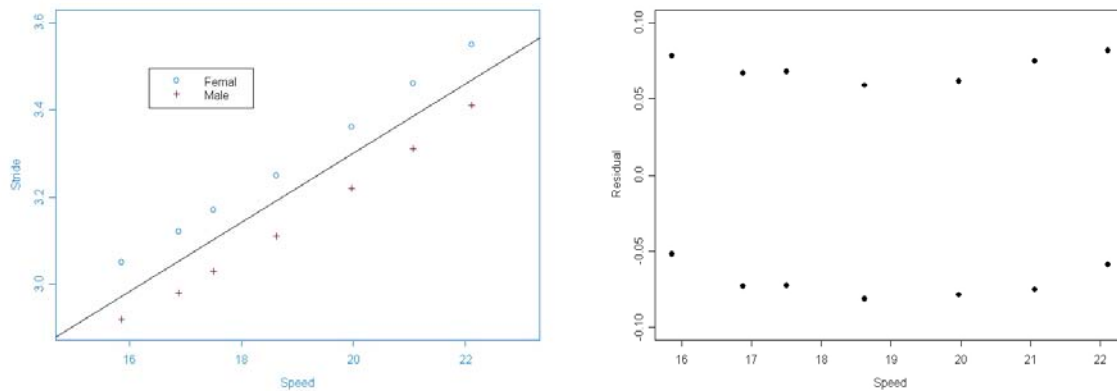


The regression line is: $Y=166.483 - 1.0987X$.
 Because the slope is negative, we see that Yield decreases as Weeds increase.

d). When $x=6$, $y=166.483 - 1.0987*6=159.89$.

Question 3 (P215 2.120)

a). The plot is as below:



b). The least-squares line for the whole data set is:

$$Y = 1.71166 + 0.0794574 * X$$

c). The residuals are:

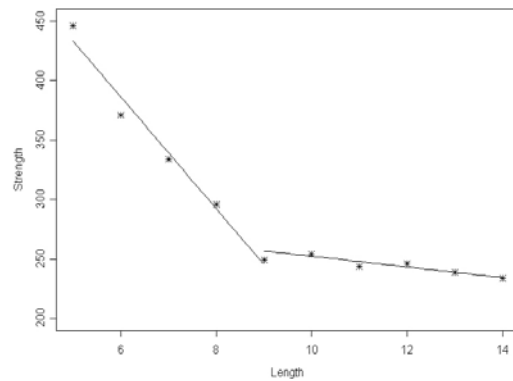
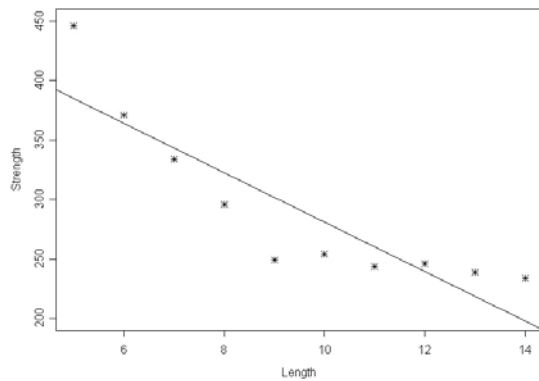
0.07815, 0.06710, 0.06784, 0.05884, 0.06158, 0.07497, 0.08154, -0.05185,
 -0.07290, -0.07216, -0.08116, -0.07842, -0.07503, -0.05846.

and the residual plot is shown above.

From the above plot, we can see that there are two distinct groups of residuals. The positive residuals are for the women and the negative residuals are for the men.

Question 4 (P216, 2.121)

a). The scatterplot is as follow (on the left):



b). The scatterplot suggests two different straight lines – one up to 8 or 9 inches and another from 9 through 14 inches. No outliers.

c). The regression line is:

$$Y = 488.382 - 20.746 * X$$

The line does not adequately describe the data. The points are above the line on the ends and below the lines in the middle.

d). The two lines are shown above (on the right). They fit the data quite well. I would like to ask the experts what differences there are between sample < 9 and > 9 in length.