STA 108 HW #1

Ch 1 : # 2, 8, 12, 27 (on Computer), 30

2) Members of health spa pay annual membership of \$300 plus a charge of \$2 for each visit to the spa.

Y denotes the dollar cost

X the number of visits

Given Annual membership : \$300 flat fee

Cost per visit : \$2

Mathematical Expression: $Y = 300 + 2X$

This is a functional relation because for every value of x, there is exactly one corresponding value y.

8) In figure 1.6, suppose another Y observation is made.

observation obtained at $X = 45$

Would $E[Y]$ for this new observation still be 104?

$E[Y] = 9.5 + 2.1x$

$X = 45$

$= 9.5 + 2.1(45)$

$E[Y] = 104$ ∴ $E[Y]$ is going to remain the same regardless

Would $E[Y] = 108$?

It can be, however there would need to be an error term of ± 4.

12) Study on Senior citizens

relationship between physical activity and frequency of colds

weekly time spent over 5 years

- negative statistical relation exists

Investigator concluded that exercise can reduce colds

a) The data obtained was observational.

b) I believe there is some validity but also some logical fallacies related to his conclusion. Although there is a relationship between the two, there are too many external factors that could affect this observational data.

c) 1. Overall Health / Immune system strength – Healthier seniors are stronger and their immune systems strengthen as a result, decreasing the likelihood of a cold.

2. Diet – A good diet provides energy to exercise and reduce cold frequency.

d. This experiment could be made valid if the experiment was completed in a controlled environment. We could randomly assign senior citizens to different groups with varying health levels of exercise. We should also maintain control over the age groups, diet, and health status.

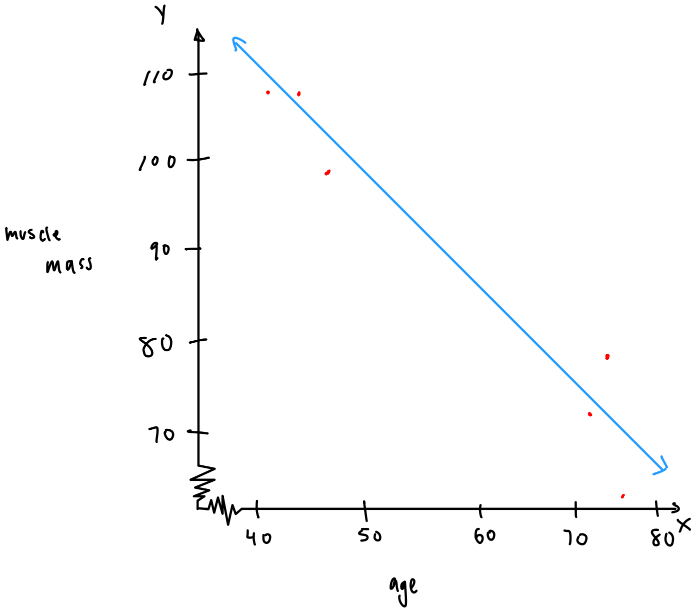
27) a) Muscle Mass is expected to decrease as age increases

$n = 15$ woman from each 10 year age group

$[40, 79]$

X - age

Y - measure of muscle mass



The linear regression line fits the plot and supports the conclusion. There is a negative relation supporting that muscle mass decreases as age increases.

Estimate linear regression

$\bar{y} = b_0 + b_1 X$

$$b_1 = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\Sigma (x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Datapoints: (43,106), (41,106), (47,97), (76,56), (72, 70) (76, 74)

$\Sigma x = 43 + 41 + 47 + 76 + 72 + 76 = 355$

$\Sigma y = 106 + 106 + 97 + 56 + 70 + 74 = 509$

$\Sigma x^2 = 43^2 + 41^2 + 47^2 + 72^2 = 22,475$

$\Sigma xy = 43 \times 106 + 41 \times 106 + 47 \times 97 + 76 \times 56 + 72 \times 70 + 76 \times 74 = 28,383$

$n = 6$ in the table given to us

$\bar{x} = \Sigma x / n = 355/6 = 59.17$

$\bar{y} = \Sigma y / n = 509/6 = 84.83$

$S(xx) = \Sigma x^2 - (\Sigma x)^2/n = 22475 - (355)^2/6 = 1470.83$

$S(xy) = \Sigma xy - (\Sigma x)(\Sigma y)/n = 28,383 - (355 \times 509)/6 = -1758.17$

$b_1 = \frac{S_{xy}}{S_{xx}} = \frac{-1758.17}{1470.83} = -1.195 \approx -1.20$

$b_0 = \bar{y} - b_1 \bar{x} = 84.83 - (-1.20)(59.17) \approx 155.8$

$$\boxed{\text{Estimated linear formula: } \bar{y} = 155.8 - 1.20x}$$

The negative slope suggests a support for the existance of muscle mass decreasing as we get older.

b) 1) R-code = Point Estimate : -1.19

2) R-code = Point Estimate of mean muscle mass: 84.95

3) R-code = Residual value 8 : 4.443

4) point estimate of σ^2 = MSE

$$MSE = \frac{\text{Square of Sum Residuals}}{n-2}$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\bar{x} = \frac{1}{60} \sum x_i \approx 59.03$$

$$\bar{y} = \frac{1}{60} \sum y_i \approx 86.35$$

① Find regression line:

$$b_1 = \frac{\sum (x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sum (x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

*Used calculator (desmos)

$b_1 = -0.726$

$b_0 = 129.22$

$\hat{y} = 129.22 - 0.726x$

② residuals

$$SSE = \sum_{i=1}^{60} (y_i - \bar{y}_i)^2 \approx 3874.45$$

$$\hat{\sigma}^2 = \frac{SSE}{n-2} = \frac{3874.45}{60-2} = \frac{3874.45}{58} = \boxed{66.8}$$

30) Refer to regression model (1.1). What is the implication for the regression function is $\beta_1 = 0$ so that the model is $Y_i = \beta_0 + \varepsilon_i$? How would regression function plot on a graph?

model 1.1 = $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

if $\beta_1 = 0$

$Y_i = \beta_0 + \varepsilon_i$

When $\beta_1 = 0$, it implies that there is no dependent relationship between X and Y. The regression function in this case would be a horizontal line at height β_0.