Section 3.7 → p. 123 Eq. 6

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

→ For regression line, we have the assumption:

$\varepsilon \sim N(\mu = 0, \sigma^2)$ that is constant variance and a 0 mean.

$$E(Y) = E(\beta_0 + \beta_1 x + \varepsilon_0)$$

$$\bar{y} = \beta_0 + \beta_1 \bar{x} + E(\varepsilon)$$

now using the assumption $\varepsilon_0 = 0$ =>

$$\Rightarrow \bar{y} = \beta_0 + \beta_1 \bar{x}$$

=> Least square line passes through the point $(\bar{x}, \bar{y})$

Section 3.7 → p. 120 (ex 1, 3, 4a)

=> (1) Describe Null hypothesis

- By using the Null hypothesis we can decide if the predictors are useful in predicting the response.

$H_0 = \beta_p = 0$

$H_1 = \beta_1$ is not zero

$$F = \frac{(TSS - RSS)/p}{RSS/(n-p-1)}$$

If there is no relationship between response and predictors F-statistic is close to 1 (H_0 is true). Else, H_a is true, F-statistic >1

p-value is associated with F-statistic; when it's very small it ~~sho~~ gives a good sign that the predictor is associated with increased sales;

=> here we see that TV and radio are closely associated with increased sales, while newspaper is not (p-value of newspaper is almost 1);

---

(3) $\hat{y} = 50 + 20x_1 + 0.07x_2 + 35x_3 + 0.01x_4 - 10x_5$

$\hat{y}_m = 50 + 20x_1 + 0.07x_2 + 35(0) + 0.01x_4 - 10(0)$

$\hat{y}_f = 50 + 20x_1 + 0.07x_2 + 35(1) + 0.01x_4 - 10x_5$

=> iii) For a fixed value of IQ and GPA, males earn more on average than females provided that the GPA is high enough.

b) female, IQ = 110, GPA = 4.0

$\hat{y}_f = 85 + 20 \times 4 + 0.07 \times 110 + 0.01(4 \times 110) - 10(4 \times 1) = \boxed{137.1}$

c) It's false because we have to take into account the standard error of beta estimates;

---

(4) a) It's hard to tell without more information. However, as the true relationship between x and y is linear, we can expect the least squares line to be close to true regression line => RSS may be lower than for cubic regression;