Chapter 12

Simple Linear Regression

12.2:

Fitting a straight line to a set of data yields the prediction line

$$\hat{Y}_i = 7 + 2X_i$$

The values of X used to find the prediction line range from 1 to 25

a) Should this model be used to predict the mean value of Y when X = 3?

b) Should this model be used to predict the mean value of Y when X=-3?

c) Should this model be used to predict the mean value of Y when X=0?

d) Should this model be used to predict the mean value of Y when X=24?

Solution:

a) Yes

b) No

c) No

d) Yes

12.5:

Zimmer's posts restaurant ratings for various locations in the United States. A sample of 100 restaurants in New York city was selected.

a) Develop a regression model to predict the cost per person?

$b_0 = -46.7718$ $b_1 = 1.4963$

$$\hat{Y} = b_0 + b_1 X_i$$

$$\hat{Y} = -46.7718 + 1.4963 \ X_i$$

b) Predict the mean cost per person for a restaurant when $X_i = 50$

$$\hat{Y} = -46.7718 + 1.4963(50) = 28.0432$$

12.17

If SSR=9740.062 and SST=17844.75 from a sample of 100

a) Compute the coefficient of determination, r^2,and interpret its meaning.

Solution:

$$r^2 = \frac{SSR}{SST} = \frac{9740.062}{17844.75} = 0.5458$$

54.58% of the variation in the cost of a restaurant meal can be explained by the variation in the rating.

b) Determine the standard error of the estimate

$$S_{XY} = \sqrt{\frac{SSE}{n-2}}$$

$$SSE = SST - SSR$$

$$SSE = 17844.75 - 9740.062 = 8104.6871$$

*The standard error of the estimate

$$S_{XY} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{8104.6871}{100-2}} = 9.0940$$

c)How useful do you think this regression model is for predicting the cost of a restaurant meal

Based on (a) the model is only moderately useful for predicting the cost of a restaurant meal.

12.43: based on 12.5

$b_1 = 1.4963$ and $S_{b_1} = 0.1379$

a) At the 0.05 level of significance. is there evidence of a linear relationship between rating of a restaurant and the cost of a meal?

Solution:

Step 1:

$H_0: \beta_1 = 0$ (There is no linear relationship between X and Y, the slope is zero)

$H_1: \beta_1 \neq 0$ (There is linear relationship between X and Y, the slope is not zero)

Step 2 :

$$\boxed{t_{stat} = \frac{b_1 - \beta_1}{S_{b_1}}} = \frac{b_1 - 0}{S_{b_1}} = \frac{b_1}{S_{b_1}} = \frac{1.4963}{0.1379} = 10.8506$$

Step 3 :

At the $\alpha = 0.05$ level of significance

$$t_{\frac{\alpha}{2},n-2} = t_{\frac{0.05}{2},100-2} = t_{0.025,98} = \pm 1.9845$$

Step 4 : decision

Since the $t_{stat} = 10.8506$ is greater than the upper critical value $t_{\frac{\alpha}{2},n-2} = 1.9845$, reject the null hypothesis. There is evidence of a linear relationship between the cost of a meal and the rating.

b) Construct a 95% confidence interval estimate of the population slope, β1.

$$\beta 1 = b_1 \pm t_{\frac{\alpha}{2}, n-2} S_{b_1}$$

$$= 1.4963 \pm 1.9845(0.1379)$$

$$= 1.4963 \pm 0.2737$$

1.2206 < β1 < 1.7700

12.51:

The table below contains the calories and fat, in grams, of seven different types of coffee drinks

Coffee	Calories(X)	Fat(Y)
1	238	7.9
2	259	3.4
3	346	22.2
4	347	19.8
5	419	16.3
6	505	21.5
7	527	18.5

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.697973
R Square	0.487166
Adjusted R	0.384599
Standard E	5.666141
Observatic	7

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	152.491373	152.491373	4.749747443
Residual	5	160.5257699	32.10515397	
Total	6	313.0171429		

	Coefficient	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-1.2703	8.056900953	-0.157668478	0.880887605	-21.98124255	19.440604	-21.9812425	19.44060393
X Variable 1	0.04487	0.020586676	2.17939153	0.081175276	-0.008053308	0.0977862	-0.00805331	0.097786163

a) At the 0.05 level of significance, is there a significant linear relationship between calories and fat? (Using t Test)

Solution:

Step 1:

$H_0: \beta_1 = 0$ (There is no linear relationship between X and Y, the slope is zero)

$H_1: \beta_1 \neq 0$ (There is linear relationship between X and Y, the slope is not zero)

step 2 :

$$t_{stat} = \frac{b_1 - \beta_1}{S_{b_1}} = \frac{b_1 - 0}{S_{b_1}} = \frac{b_1}{S_{b_1}} = \frac{0.04487}{0.0205} = 2.1887$$

Step 3:

$$t_{\frac{\alpha}{2},n-2} = t_{\frac{0.05}{2},7-2} = t_{0.025,5} = \pm 2.5706$$

$$-t_{\frac{\alpha}{2},n-2} < t_{stat} < +t_{\frac{\alpha}{2},n-2}$$

Step 4:

Decision: Do not reject H_0. There is no significant linear relationship between calories and fat

b) At the 0.05 level of significance, is there a significant linear relationship between calories and fat? (Using F Test)

$$F_{\text{stat}} = \frac{MSR}{MSE}$$

$$MSR = \frac{SSR}{C-1}$$

SSR= 152.49137

SSE= 160.5257699

$$\text{MSR} = \frac{SSR}{C-1} = \frac{152.49137}{2-1} = 152.49137$$

$$\text{MSE} = \frac{SSE}{n-c} = \frac{160.5257699}{7-2} = 32.10515$$

$$F_{\text{stat}} = \frac{MSR}{MSE} = \frac{152.49137}{32.10515} = 4.7497$$

**** $F_{\alpha, (C-1),(n-c)}$

$$F_{0.05, 1, 5} = 6.61$$

$$F_{STAT} < F_{\alpha}$$

Decision: Do not reject H_0. There is no significant linear relationship between calories and fat