Winter 2019 QMS202 Linear Regression Review

Chapter 15 Simple Linear Regression Tests

Calculator Lesson 18B

Example 15.6

The sales for Sunflowers Apparel, a chain of apparel stores for women, have increased during the past 12 years as the chain has increased the number of its stores. Until now, Sunflowers senior manager selected sites based on subjective factors such as the availability of a good lease or the perception that a location seemed ideal for an apparel store. As the new director of planning, you need to develop a systematic approach to selecting new sites that will allow Sunflowers to make better-informed decisions for opening additional stores. This plan must be able to forecast annual sales for all potential stores under consideration. You believe that the size of the store significantly contributes to the success of a store and you want to use this relationship in the decision-making process. You wish to predict annual sales based on the size of the store in square feet.

Store #	Square Feet (000)	Annual Sales (in millions of dollars)	Store #	Square Feet (000)	Annual Sales (in millions of dollars)
1	1.7	3.7	8	1.1	2.7
2	1.6	3.9	9	3.2	5.5
3	2.8	6.7	10	1.5	2.9
4	5.6	9.5	11	5.2	10.7
5	1.3	3.4	12	4.6	7.6
6	2.2	5.6	13	5.8	11.8
7	1.3	3.7	14	3.0	4.1

The dependent variable, y, is the annual sales (in millions of dollars), and the independent variable, x, is the size of the store (in square feet).

1. Compute the regression coefficient – b_0 (intercept) and b_1 (slope)
   To solve this on the calculator, Under Main menu, go to **STAT**, Input the data of store sizes (x values) in List 1 and the annual sales (y values) in List 2. Then go to **Calc** (F2) and **Set** (F6). Go to the 3rd row and enter 2 Var XList: List 1 ; 2 Var YList: List2 ; 2 Var Freq: 1. EXE to return to the display of the data. Now press **REG** (F3) and press **X** (F1). At this point, you are offered 2 options: ax+b (F1) or **a+bx** (F2).
   Preferably choose (F2) a +bx, then the calculator gives the following results:
   LinearReg (a+bx), a = 0.96447365 $=b_0$ b = 1.66986231$=b_1$ r = 0.95088327
   r^2 = 0.904179 MSe = 0.93388968 y = a +bx

The values are put together to obtain the simple linear regression equation (the prediction line): $\hat{y} = 0.96447365 + 1.66986231x$
The coefficient of determination, r^2, is 0.904179.

2. Hypothesis tests for the slope and correlation coefficient
   At the 0.05 level of significance, is there evidence of a linear relationship between the size of the store and annual sales?
   Solution: We test for a linear relationship between the variables x and y by testing whether the population regression coefficient (slope), B, is different from zero.
   H_0: $\beta_1 = 0$ (no linear relationship) H_1: β_1=/=0 (a linear relationship exists)
   Using the same store size and annual sales values lists from before, from **MENU, STAT, TEST, t, REG,** we get:
   LinearReg tTest: β & ρ : =/= (F1) 0 XList: List1 YList: List2 Freq: 1
   Save Res: None **EXE**
   LinearReg tTest: β =/= 0 & ρ =/= 0 t = 10.6411237 p= 1.8226E-07
   df = 12 a = 0.96447365 b = 1.66986232 se = 0.96637967
   r = 0.95088327 $r^2 = 0.904179$
   Conclusion: p-value = 1.8226E-07 = 0.00000018226 < 0.05 = α
   Hence reject H_0.

3. Compute the prediction value
   What is the average annual sales if the size of the store is 1.8(000) square feet?
   Make sure the data is recorded in the calculator and the simple linear regression is performed prior to computing the prediction value.
   Go to **MENU RUN-MAT EXE** then enter 1.8 **OPTN** select **STAT** (F5) **ŷ** (F2) **EXE** the calculator gives 3.97022583, which we interpret as :
   $\hat{y} = 0.96447365 + 1.66986231 \cdot (1.8) = 3.97022583$

4. Compute the residuals
   MENU, STAT, TEST, t, REG, we get:
   LinearReg tTest: β & ρ : =/= (F1) 0 XList: List1 YList: List2 Freq: 1
   Save Res: List3 EXE
   Press **Exit** and you will see the residuals appear in List 3.

Chapter 16 Multiple Regression (for 2 or more variables)

Multiple Regression Equation with 2 Independent Variables

$$\hat{Y}_i = b_0 + b_1X_{1i} + b_2X_{2i}$$

Interpreting the Figure 16.2 on Page 708 (SPSS regression results worksheet for OmniPower sales data)

Variables Entered/Removed

Model	Variables Entered	Variables Removed	Method
1	Price = X_{1i} Promotion = X_{2i}		Enter

a. All requested variables entered.
b. Dependent Variable: Sales

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.870 $= \sqrt{R}$	0.758 $= R^2 = \frac{SSR}{SST}$	0.742 $= R^2_{adj}$ $R^2_{adj} = 1 - [\,(1 - R^2) * \frac{n-1}{n-k-1}\,]$	638.065 $= \sqrt{MSE}$

a. Predictors: (Constant), Price, Promotion

ANOVA

Model	Degrees of Freedom	Sum of Squares	Mean Squares (Variance)	*F*	Sig.
1 Regression	2 = k = # of independent variables in the regression model	39472730.773 =SSR $= \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$ = MSR * k	19736365.387 $= MSR$ $= \frac{SSR}{k}$	48.477 $= F_{STAT}$ $= \frac{MSR}{MSE}$	0.000 = p-value
Error	31 = n – k – 1	12620946.668 = SSE $= \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ = MSE * (n – k – 1)	407127.312 $= MSE$ $= \frac{SSE}{n-k-1}$		
Total	33 = n – k – 1 + k = n – 1	52093677.441 = SST = SSR + SSE			

a. Predictors: (Constant), Price, Promotion b. Dependent Variable: Sales

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	t = t_{STAT} = b_j / S_{bj}	Sig. = p-value
1 (Constant)	5837.521 = b_0	628.150 = S_{b0}		9.293 = b_0 / S_{b0}	0.000
Price	-53.217 = b_1	6.852 = S_{b1}	-0.690	-7.766 = b_1 / S_{b1}	0.000
Promotion	3.613 = b_2	0.685 = S_{b2}	0.468	5.273 = b_2 / S_{b2}	0.000