Another Regression Example

Example:

Age of Batteries (No. of flashes)	Recharge Time
0	1.1
10	1.4
20	1.6
30	1.9
40	2.1
50	2.3
60	2.8
70	2.7
80	3.2
90	3.1

	Age	Time	Age^2	Time^2	Age X Time
	0	1.1	0	1.21	0
	10	1.4	100	1.96	14
	20	1.6	400	2.56	32
	30	1.9	900	3.61	57
	40	2.1	1600	4.41	84
	50	2.3	2500	5.29	115
	60	2.8	3600	7.84	168
	70	2.7	4900	7.29	189
	80	3.2	6400	10.24	256
	90	3.1	8100	9.61	279
Total	**450**	**22.2**	**28500**	**54.02**	**1194**

Basic computations:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{450}{10} = 45$$

$$\bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{22.2}{10} = 2.22$$

$$S_{xx} = \sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 = 28500 - \frac{1}{10}450^2 = 8250$$

$$S_{yy} = \sum_{i=1}^{n}\left(y_i - \bar{Y}\right)^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2 = 54.02 - \frac{1}{10}22.2^2 = 4.736$$

$$S_{xy} = \sum_{i=1}^{n}\left(x_i - \bar{X}\right)\left(y_i - \bar{Y}\right) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) = 1194 - \frac{1}{10}450 \times 22.2 = 195$$

Regression coefficients:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{195}{8250} = 0.023636364$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X} = 2.22 - 0.023636364 \times 45 = 1.156363636$$

***R*²** **computation:**

$$r^2 = \frac{SSR}{SST} = \frac{\hat{\beta}_1^2 S_{xx}}{S_{yy}} = \frac{S_{xy}^2}{S_{xx} S_{yy}} = \frac{195^2}{8250 \times 4.736} = 0.973203317$$

$$r = \sqrt{r^2} = \sqrt{0.973203317} = 0.9865107$$

ANOVA comutations:

$$SST = S_{yy} = 4.736$$

$$SSR = \hat{\beta}_1^2 S_{xx} = 0.023636364^2 \times 8250 = 4.609090909$$

$$SSE = SST - SSR = 4.736 - 4.609090909 = 0.126909091$$

$$MSR = \frac{SSR}{1} = \frac{4.609090909}{1} = 4.609090909$$

$$MSE = \frac{SSE}{n-2} = \frac{0.12609091}{8} = 0.015863636$$

$$F = \frac{MSR}{MSE} = \frac{4.609090909}{0.015863636} = 290.5444126$$

Significance F = FDIST(290.544126,1,8)=1.42525E-07

Standard error of regression:

$$s = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{0.12609091}{8}} = 0.125950928$$

Significance of regression coefficients:

$$SE\left(\hat{\beta}_0\right) = s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n S_{xx}}} = 0.1259509\sqrt{\frac{28500}{10 \times 8250}} = 0.074028$$

$$t = \frac{\hat{\beta}_0}{SE\left(\hat{\beta}_0\right)} = \frac{1.156363636}{0.074028} = 15.61923147 = 15.6206$$ for β_0

P-value for β_0 = TDIST(15.61923,8,2)=2.81306E-07

$$SE\left(\hat{\beta}_1\right) = \frac{s}{\sqrt{S_{xx}}} = \frac{0.125950928}{\sqrt{8250}} = 0.001386674$$

$$t = \frac{\hat{\beta}_1}{SE\left(\hat{\beta}_1\right)} = \frac{0.0236363636}{0.001386674} = 17.04536$$ *for β_1*

P-value for β_1 = TDIST(17.04536,8,2)=1.42525E-07

95% CI for β_0 =

$$\hat{\beta}_0 - t_{n-2,0.025} SE\left(\hat{\beta}_0\right) \le \beta_0 \le \hat{\beta}_0 + t_{n-2,0.025} SE\left(\hat{\beta}_0\right)$$
$$1.1563636 - 2.306 \times 0.074028 \le \beta_0 \le 1.1563636 + 2.306 \times 0.074028$$
$$0.98565 \le \beta_0 \le 1.32707$$

95% CI for β_1 =

$$\hat{\beta}_1 - t_{n-2,0.025} SE\left(\hat{\beta}_1\right) \le \beta_1 \le \hat{\beta}_1 + t_{n-2,0.025} SE\left(\hat{\beta}_1\right)$$
$$0.02363636 - 2.306 \times 0.001386674 \le \beta_0 \le 0.02363636 + 2.306 \times 0.001386674$$
$$0.02044 \le \beta_0 \le 0.023834$$

Example: 95% <u>Confidence</u> interval for μ* at AGE=65

$$\hat{\mu}^* = \hat{\beta}_0 + \hat{\beta}_1 x^* = 1.15636364 + 0.023636364 \times 65 = 2.692727$$

$$\hat{\mu}^* - t_{n-2,\alpha/2}\, s\sqrt{\frac{1}{n} + \frac{\left(x^* - \overline{X}\right)^2}{S_{xx}}} \le \mu^* \le \hat{\mu}^* + t_{n-2,\alpha/2}\, s\sqrt{\frac{1}{n} + \frac{\left(x^* - \overline{X}\right)^2}{S_{xx}}}$$

$$2.69273 - 2.306 \times 0.12595\sqrt{\frac{1}{10} + \frac{(65-45)^2}{8250}} \le \mu^* \le 2.69273 - 2.306 \times 0.12595\sqrt{\frac{1}{10} + \frac{(65-45)^2}{8250}}$$

$$2.58081 \le \mu^* \le 2.80465$$

Example: 95% <u>Prediction</u> interval for Y* at AGE=65

$$\hat{Y}^* = \hat{\beta}_0 + \hat{\beta}_1 x^* = 1.15636364 + 0.023636364 \times 65 = 2.692727$$

$$\hat{Y}^* - t_{n-2,\alpha/2}\, s\sqrt{\frac{1}{n} + \frac{\left(x^* - \overline{X}\right)^2}{S_{xx}}} \le Y^* \le \hat{Y}^* + t_{n-2,\alpha/2}\, s\sqrt{\frac{1}{n} + \frac{\left(x^* - \overline{X}\right)^2}{S_{xx}}}$$

$$2.69273 - 2.306 \times 0.12595\sqrt{1 + \frac{1}{10} + \frac{(65-45)^2}{8250}} \le \mu^* \le 2.69273 - 2.306 \times 0.12595\sqrt{1 + \frac{1}{10} + \frac{(65-45)^2}{8250}}$$

$$2.38147 \le \mu^* \le 3.00399$$

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.986511
R Square	0.973203
Adjusted R Square	0.969854
Standard Error	0.125951
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	4.609091	4.609091	290.5444	1.42525E-07
Residual	8	0.126909	0.015864		
Total	9	4.736			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	1.156364	0.074028	15.6206	2.81E-07	0.985654332	1.327073
Age	0.023636	0.001387	17.04536	1.43E-07	0.020438685	0.026834