

## Math 10 MPS Homework 11 – Regression

1. We list for 11 printers, their cost (**C**, in hundreds of dollars) and their rating (**R**, on a scale from 0-100). The results are

C	4	5	2.5	2	4	2.85	2.1	2.5	2	1.8	1.3
R	90	90	85	83	78	78	75	70	68	63	61

Minitab output is attached. Discuss the scatter plot of the data where C is the dependent variable and R is the independent variable.

- Interpret the correlation coefficient ( $r$ ).
- Report and Interpret the least square line (regression line) which will be used to predict the Cost of the printer from the Rating.
- Test the hypothesis that Cost and Rating are correlated. Explain why this is the same as testing whether the slope of the regression line is not equal to zero.
- Report and interpret the standard error of C|R (the standard deviation “with respect to” the regression line.)
- Report and interpret  $r^2$ , the coefficient of determination.
- Analyze the residuals and note any unusual observations.
- Make Cost predictions for the following printers. Are any of these predictions inappropriate?
  - Printer X – Rating =87
  - Printer Y – Rating =75
  - Printer Z – Rating =21
- Find a 95% confidence interval for the expected Cost of Printers X and Y.
- Find a 95% prediction interval for the actual Cost of Printers X and Y.

## Minitab Output for Question 1

### Regression Analysis: Cost versus Rating

The regression equation is  
 $Cost = -3.74 + 0.0846 \text{ Rating}$

Predictor	Coef	SE Coef	T	P
Constant	-3.738	1.892	-1.98	0.080
Rating	0.08462	0.02455	3.45	0.007

S = 0.785173    R-Sq = 56.9%    R-Sq(adj) = 52.1%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	7.3229	7.3229	11.88	0.007
Residual Error	9	5.5485	0.6165		
Total	10	12.8714			

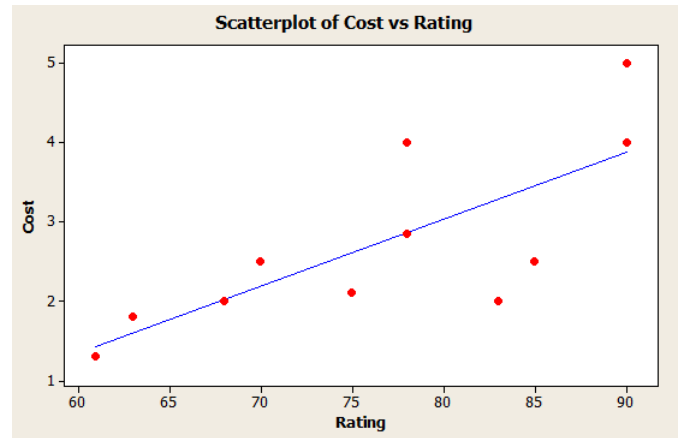
Obs	Rating	Cost	Fit	SE Fit	Residual	St Resid
1	90.0	4.000	3.878	0.408	0.122	0.18
2	90.0	5.000	3.878	0.408	1.122	1.67
3	85.0	2.500	3.455	0.316	-0.955	-1.33
4	83.0	2.000	3.286	0.286	-1.286	-1.76
5	78.0	4.000	2.863	0.240	1.137	1.52
6	78.0	2.850	2.863	0.240	-0.013	-0.02
7	75.0	2.100	2.609	0.239	-0.509	-0.68
8	70.0	2.500	2.186	0.285	0.314	0.43
9	68.0	2.000	2.016	0.315	-0.016	-0.02
10	63.0	1.800	1.593	0.406	0.207	0.31
11	61.0	1.300	1.424	0.447	-0.124	-0.19

#### Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	3.624	0.351	( 2.831, 4.418)	( 1.679, 5.570)
2	2.609	0.239	( 2.067, 3.150)	( 0.752, 4.466)
3	-1.961	1.382	(-5.087, 1.166)	(-5.556, 1.635)

#### Values of Predictors for New Observations

New Obs	Rating
1	87.0
2	75.0
3	21.0



2. The following regression analysis was used to test Poverty (percentage living below the poverty line) as a predictor for Dropout (High School Dropout Percentage). Five items have been blanked out below and can be calculated based on other information in the output.

$r^2$	<input type="text"/>	n	50
r	<input type="text"/>	k	1
Std. Error	<input type="text"/>	Dep. Var.	<b>HSDropouts</b>

ANOVA table

Source	SS	df	MS	F	p-value
Regression	67.45	1	67.45	<input type="text"/>	
Residual	216.18	48	4.50		
Total	283.62	49			

Regression output

variables	coefficients	std. error
Intercept	6.212	1.086
Poverty	0.291	0.075

Predicted values for: HSDropouts

Poverty	Predicted	95% Confidence Intervals		95% Prediction Intervals	
		lower	upper	lower	upper
10	9.117	8.273	9.961	4.767	13.466
15	<input type="text"/>	9.944	11.195	6.257	14.882

- Fill in the missing information from the output
  - $r^2$
  - r
  - Std. Error
  - F Test Statistic
  - Predicted Value for Poverty = 15
- Write out the regression equation.
- Conduct the Hypothesis Test that Poverty and HSDropout are correlated with  $\alpha = .01$  (Critical Value for F is 7.19 ( $\alpha = .01$ , DFnum=1, DF=48)).
- What percentage of the variability of High School Dropout Rates can be explained by Poverty?
- North Dakota has a Poverty Rate of 11.9 percent and a HS Dropout Rate of 4.6 percent.
  - Calculate the predicted HS Dropout Rate for North Dakota from the regression equation.
  - The Standard Error (from part a-iii) is the standard deviation with respect to the regression line. Calculate the Z-score for the actual North Dakota HS Dropout Rate of 4.6 (Subtract the predicted value and divide by the Standard Error). Do you think that the North Dakota HS Dropout Rate is unusual? Explain.