University of Guelph Numeracy Project

Multiple Regression: Examples

TABLE OF CONTENTS

Multiple Regression: Examples

Population Regression Equation

Population Regression Equation

► The following example demonstrates an application of multiple regression to a real life situation.

Ricardo has concerns over his coming final statistics exam. In response, his professor outlines how Ricardo can estimate his grade on the final exam through consideration of the grades he received throughout the semester:

Midterm 1: x1
Midterm 2: x2
Assignments: x3

As such, Ricardo can predict his final exam grade (response variable y) using the three scores identified above (n = 3 explanatory variables)

Multiple Linear Regression Model

Multiple Linear Regression Model

► Refer back to the example involving Ricardo. We can now use the prediction equation to estimate his final exam grade.

In a past statistics class, a regression of final exam grades for Test 1, Test 2 and Assignment grades resulted in the following equation:

$\hat{y}_{final} = -5.70_{b_0} + 0.38\text{Test1}_{b_1 x_1} + 0.42\text{Test2}_{b_2 x_2} + 0.61\text{Assign}_{b_3 x_3}$

Under the assumption that Ricardo scored 70% on Test 1, 60% on test 2 and 80% on the assignments, his predicted final exam grade would have been:

$\hat{y} = -5.70 + 0.38(70) + 0.42(60) + 0.16(80)$
$\hat{y} = 58.9$
As such, Ricardo might expect to receive a 58.9 on his final exam

Multiple Regression Example Question

Example Question

► A researcher wants to relate the taste of cheese to its concentrations of 3 chemicals: Acetic acid, Hydrogen Sulfide and Lactic acid. The acetic acid and hydrogen sulfide (H_2S) measurements are actually natural logs of their concentrations (i.e. they used a transformation of these 2 variables).

The following SPSS output summarizes the results:

Source	DF	Sum of Squares	Mean Square	F	Sig.
Regression	3	4994.476	1664.825	16.221	0.000(a)
Residual	26	2668.411	102.631		
Total	29	7662.887			

In this example, we want to relate the taste of cheese to 3 ingredients:

y = taste
x_1 = acetic acid
x_2 = hydrogen sulfide
x_3 = lactic acid

Our hypothesis is as follows:

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$
H_a: at least one β_j is not = 0

Given our small p-value (Sig. = 0.000), we can reject H_0 and be confident that not all the slopes are equal to 0. However, we are also confident that at least one of the explanatory variables (x_1, x_2, x_3) is helping predict taste (y).

We can now investigate which explanatory variables are important predictors of taste.

Model	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	-28.877	19.735	-1.463	0.155
Acetic acid o	0.328	4.46	0.73	0.942
H_2S	3.912	1.248	3.133	0.004
Lactic acid	19.671	8.629	2.28	0.031

The above ANOVA results indicate that both H_2S and Lactic acid are important predictors of taste, as indicated by their significance. Let us see what will happen when we consider Acetic acid as the ONLY explanatory variable.

Model	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	-61.499	24.846	-2.475	0.02
Acetic acid	15.648	4.496	3.481	0.002

When considering Acetic acid in the model on its own, it is a good predictor of taste, which is demonstrated by its small p-value (i.e. significance level). However, as we saw on the previous page, Acetic acid is not a good predictor of taste when paired with lactic acid and H_2S.

Output after removal of Acetic acid:

Source	DF	Sum of Squares	Mean Square	F	Sig.
Regression	2	4993.321	2496.961	25.26	0.000(a)
Residual	27	2668.965	98.851		
Total	29	7662.887			

You can see that there is still a small p-value (significance = 0.000), even after the removal of Acetic acid.

Model	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	-27.592	8.982	-3.072	0.005
H_2S	3.946	1.136	3.475	0.002
Lactic acid	19.887	7.959	2.499	0.019

So, both H_2S and Lactic acid are explanatory variables used to predict taste (y). This gives a final prediction model of:

Taste = -27.592 + 3.946 H_2S + 19.887 Lactic acid