

Instructions for carrying out statistical procedures and tests using Minitab

These instructions are closely linked to the author's book:

Essential Statistics for the Pharmaceutical Sciences
John Wiley & Sons Ltd <http://eu.wiley.com>
2007
ISBN: 978-0-470-03468-2

For all references to chapters or tables, see the above book.

**Using Minitab to perform a
multiple regression analysis**

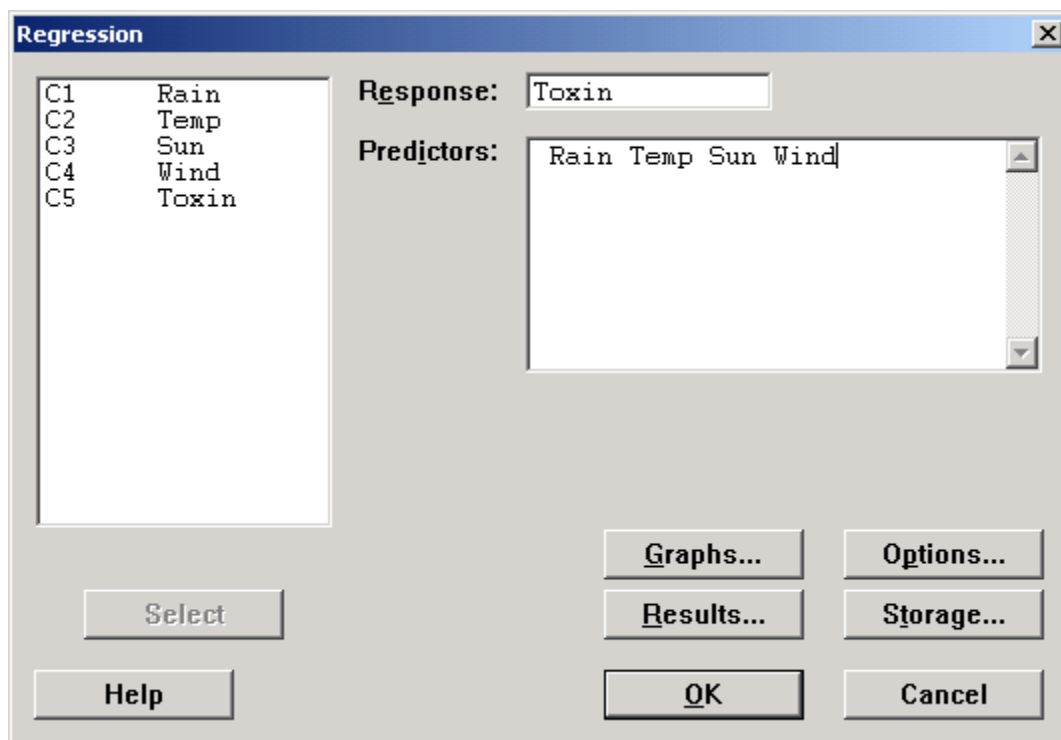
Using Minitab to perform a multiple regression analysis

Example: Table 14.6 Rainfall, temperature, sunshine and wind speed at growing sites and concentration of fungal toxin in nuts.

Enter the data into 5 appropriately labelled columns and then follow the menus:

Stat / Regression / Regression...

Toxin is entered into the Response box, and all 4 meteorological factors into the Predictors box. The Regression box should then be as below:



The output is shown on the next page:

Regression Analysis: Toxin versus Rain, Temp, Sun, Wind

The regression equation is

$$\text{Toxin} = 31.6 + 7.07 \text{ Rain} - 0.420 \text{ Temp} - 0.237 \text{ Sun} - 0.794 \text{ Wind}$$

Predictor	Coef	SE Coef	T	P
Constant	31.608	7.105	4.45	0.007
Rain	7.068	1.003	7.05	0.001
Temp	-0.4201	0.2413	-1.74	0.142
Sun	-0.2375	0.5086	-0.47	0.660
Wind	-0.7936	0.2977	-2.67	0.045

S = 1.57397 R-Sq = 91.9% R-Sq(adj) = 85.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	139.782	34.946	14.11	0.006
Residual Error	5	12.387	2.477		
Total	9	152.169			

Source	DF	Seq SS
Rain	1	114.842
Temp	1	4.712
Sun	1	2.625
Wind	1	17.603

Under 'Analysis of Variance', near the end of the output, the overall regression is shown to be statistically significant ($P = 0.006$). However, further up the output (Just below the regression equation), we find P values for each individual predictor. Rain and Wind are shown with P values of 0.001 and 0.045, so both are significant. The other two (Temperature and Sunshine) are apparently not significant ($P = 0.142$ and 0.660 respectively).

We must now remove one predictor. Since Temperature had a P value of 0.142 it has a stronger claim to be retained than Sunshine ($P = 0.660$). (See Chapter 14). So, repeat the regression, using Rainfall, Temperature and Wind, but omitting Sunshine. The new output is shown on the next page:

Regression Analysis: Toxin versus Rain, Temp, Wind

The regression equation is

$$\text{Toxin} = 31.6 + 7.01 \text{ Rain} - 0.479 \text{ Temp} - 0.822 \text{ Wind}$$

Predictor	Coef	SE Coef	T	P
Constant	31.565	6.625	4.76	0.003
Rain	7.0108	0.9285	7.55	0.000
Temp	-0.4790	0.1919	-2.50	0.047
Wind	-0.8218	0.2718	-3.02	0.023

S = 1.46782 R-Sq = 91.5% R-Sq(adj) = 87.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	139.242	46.414	21.54	0.001
Residual Error	6	12.927	2.155		
Total	9	152.169			

Source	DF	Seq SS
Rain	1	114.842
Temp	1	4.712
Wind	1	19.688

The equation as a whole is significant (Under Analysis of Variance, near end of output, P = 0.001), and all the individual predictors (Rain, Temp and Wind) have significant P values (<0.001, 0.047 and 0.023 respectively).

The regression equation (Near the top of the output) is:

$$\text{Toxin} = 31.6 + 7.01 \text{ Rain} - 0.479 \text{ Temp} - 0.822 \text{ Wind}$$

Which means:

$$\text{Toxin} = 31.6 + 7.01 \times \text{Rain} - 0.479 \times \text{Temp} - 0.822 \times \text{Wind}$$

For this final model, R-squared and R-squared (adjusted) are provided about half way down the output, as 91.5% and 87.3% respectively. (See Section 14.2.5 for explanation of R-squared.)