Subject: Multiple linear regression (MLR): significance of the calculated terms (2 sigma interval)
Posted by kisCA on Wed, 03 Apr 2013 19:27:54 GMT
View Forum Message <> Reply to Message

Hello,

I have been reading MLR results for a while now in published papers. They always talk about the significance of the calculated terms of the regression by saying "it's significant at a 2 sigma level". Is there any body here that would explain to me what it really means and how it is done ? Is it simply when 2sigma is largely superior to the calculated term ?

Thank you for your help

Subject: Re: Multiple linear regression (MLR): significance of the calculated terms (2 sigma interval)
Posted by Craig Markwardt on Thu, 04 Apr 2013 05:48:01 GMT
View Forum Message <> Reply to Message

On Wednesday, April 3, 2013 3:27:54 PM UTC-4, kisCA wrote:
>
> I have been reading MLR results for a while now in published papers. They always talk about the significance of the calculated terms of the regression by saying "it's significant at a 2 sigma level". Is there any body here that would explain to me what it really means and how it is done ? Is it simply when 2sigma is largely superior to the calculated term ?

A quick and dirty way to do it is calculate the T-value which is
T = (fitted parameter value) / (uncertainty of fitted parameter)
and call that a gaussian sigma. Example, a regression coefficient is reported as 4.5 units and the uncertainty is 1.5 units, so the regression coefficient is significant at (4.5)/(1.5) = 3 sigma. Quick and dirty of course.

The following course indicates to use the Student T test to test for parameter significance....
http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm
IDL has a standard library routine TM_TEST which doesn't quite do what you want. If you peek at how it works, it computes significance from the T value as computed above. The formula is,
P = IBETA(0.5*DOF, 0.5, DOF/(DOF + T^2))
where DOF is number of degrees of freedom (number of data points minus number of regression coefficients).

The preference in my field of work is to use an F-test. (see my MPFTEST)

People usually use a shorthand statement. If something is 2-sigma significant, they usually mean that they calculated the significance (P-value) from the Student T test or F test, and then gave the gaussian "sigma" value that produced the same significance.

Example: your significance test produces a P-value of 0.0025 (in other words such a large

regresssion coefficient would happen by random chance 0.0025 x 100 = 0.25% of the time). For the probability of a gaussian variable x to be less than 0.25%, the random variable must be |x|>3.02, so that would be a "3.02-sigma" result. (use GAUSS_CVF(0.0025/2))

Craig

MPFTEST download
http://www.physics.wisc.edu/~craigm/idl/fitting.html#MPFIT

---

Subject: Re: Multiple linear regression (MLR): significance of the calculated terms (2 sigma interval)
Posted by kisCA on Thu, 04 Apr 2013 20:50:22 GMT
View Forum Message <> Reply to Message

Oh my, that was complete and awesomely understandable ! Thank you so much !

---