A Regression Approach to Bias between Methods

Application Note: 5

By

Allan Fraser

The approach to comparing regression data between two different datasets requires an assessment of the *translational* and *rotational* biases, where the former is the value of the y-intercept and the latter is the slope or "rotation" of the slope of the regressed line.

In this type of regression analysis, the variance of the x-variable is somewhat smaller than that of the y-variable *i.e.* the smaller variance is plotted on the x-axis and the larger on the y-axis. Should the variances be very different, then a paired t-test is more suitable (Thompson & Lowthian, 2011).

The relevant hypotheses are:

Y-intercept:

H_0: a=0

H_1: a≠0

Slope:

H_0: b=1

H_1: b≠1

The ideal is: a=0 and b=1

Therefore, a systematic difference between two data sets is identified if the estimated intercept if the estimated intercept differs significantly from zero or if the slope deviates significantly from 1.0. This is decided by using the decision rules for the *p*-value.

For illustration purposes, a set of results as shown in Table 1 are subjected to a regression analysis (outcome shown in Table 2). The results are interpreted as follows:

For intercept a: a = 0 (null hypothesis: no bias; ideal intercept is zero);
The intercept is 0.131 with a *p*-value for the intercept of 0.009 which is less than 0.05, therefore the intercept is significantly different to 0 and a bias exists. We reject the null hypothesis. In addition, the upper and lower 95% confidence intervals for the intercept do not span zero (+0.0469 and +0.2155), therefore the intercept is significantly different from zero. A *translational bias* exists between the two methods.

For the slope b: b = 1 (ideal slope: H_0: b=1 and H_1: b≠1).
The slope is 1.0078±0.020 and therefore does span 1.000, *i.e.* upper 95% is 1.028 and the lower 95% is 0.988.

Therefore, in this example we only see only translational bias between the two methods and rotational bias is not evident.

Table 1. Data for $\%SiO_2$ between two different lab sites on paired samples.

$\%SiO_2$ Site 1	**$\%SiO_2$ Site 2**
0.1235	0.001
1.124	0.999
2.123	2.001
3.234	2.998
4.112	4.002
5.19	4.998
6.231	5.997
7.139	7.01

Table 2. Regression results for the data between site 1 and 2.

	Coefficients	***p***	***Lower 95%***	***Upper 95%***
Intercept	0.132	0.009	+0.0470	+0.216
Slope	1.0078	<0.001	0.988	1.028
R^2	0.9996			

References

Thompson, M., Lowthian, P. (2011). *Notes on statistics and data quality for analytical chemists.* Imperial College Press: 15-115.

How to cite this document:

Fraser. A.W. (2017) . A Regression Approach to Bias between Methods. Application note 5. www.allanfraserandassociates.com (retrieved on ….date).