Contents

Chapter 1

Error In Variable Models

1.1 Regression Methods

Conventional regression models are estimated using the ordinary least squares (OLS) technique, and are referred to as 'Model I regression'(Cornbleet and Cochrane, 1979; Ludbrook, 1997). A key feature of Model I models is that the independent variable is assumed to be measured without error. As often pointed out in several papers (Altman and Bland, 1983; Ludbrook, 1997), this assumption invalidates simple linear regression for use in method comparison studies, as both methods must be assumed to be measured with error.

The use of regression models that assumes the presence of error in both variables X and Y have been proposed for use instead. (Cornbleet and Cochrane, 1979; Ludbrook, 1997), These methodologies are collectively known as 'Model II regression'. They differ in the method used to estimate the parameters of the regression.

Regression estimates depend on formulation of the model. A formulation with one method considered as the X variable will yield different estimates for a formulation where it is the Y variable. With Model I regression,the models fitted in both cases will entirely different and inconsistent. However with Model II regression, they will be consistent and complementary.

1.2 Conclusions about Existing Methodologies

Scatterplots are recommended by Altman and Bland (1983) for an initial examination of the data, facilitating an initial judgement and helping to identify potential outliers. They are not useful for a thorough examination of the data. O'Brien et al. (1990) notes that data points will tend to cluster around the line of equality, obscuring interpretation.

The Bland Altman methodology is well noted for its ease of use, and can be easily implemented with most software packages. Also it doesn't require the practitioner to have more than basic statistical training. The plot is quite informative about the variability of the differences over the range of measurements. For example, an inspection of the plot will indicate the 'fan effect'. They also can be used to detect the presence of an outlier.

Ludbrook (1997, 2002)criticizes these plots on the basis that they presents no information on effect of constant bias or proportional bias. These plots are only practicable when both methods measure in the same units. Hence they are totally unsuitable for conversion problems. The limits of agreement are somewhat arbitrarily constructed. They may or may not be suitable for the data in question. It has been found that the limits given are too wide to be acceptable. There is no guidance on how to deal with outliers. Bland and Altman recognize effect they would have on the limits of agreeement, but offer no guidance on how to correct for those effects.

There is no formal testing procedure provided. Rather, it is upon the practitioner opinion to judge the outcome of the methodology.

1.3 Background

In method comparison studies, it is of importance to assure that the presence of a difference of medical importance is detected. For a given difference, the necessary number of samples depends on the range of values and the analytical standard deviations of the methods involved. For typical examples, the present study evaluates the statistical power of least-squares and Deming regression analyses applied to the method comparison data.

1.4 Constant and Proportional Bias

Linear Regression is a commonly used technique for comparing paired assays. The Intercept and Slope can provide estimates for the constant bias and proportional bias occurring between both methods. If the basic assumptions underlying linear regression are not met, the regression equation, and consequently the estimations of bias are undermined. Outliers are a source of error in regression estimates.

Constant or proportional bias in method comparison studies using linear regression can be detected by an individual test on the intercept or the slope of the line regressed from the results of the two methods to be compared.

- Model I regression
- Model II regression

Model I regression [Criterion v Test] [Cornbleet Gochman 1979] define this analysis as the case in which the independent variable, X, is measured without error, with y as the dependent variable.

In method comparison studies, the X variable is a precisely measured reference method. In the [Cornbleet Gochman 1979] paper. It is argued that criterion may be regarded as the correct value. Other papers dispute this.

Model II regression [Test V Test] In this type of analysis,both of the measurement methods are test methods, with both expected to be subject to error. Deming regression is an approach to model II regression.

Regression approaches are useful for a making a detailed examination of the biases across the range of measurements, allowing bias to be decomposed into fixed bias and proportional bias. Fixed bias describes the case where one method gives values that are consistently different to the other across the whole range. Proportional bias describes the difference in measurements getting progressively greater, or smaller, across the range of measurements. A measurement method may have either an attendant fixed bias or proportional bias, or both. (?). Determination of these biases shall be discussed in due course.

1.5 Cornbleet - Cochran

This regression method also calculates a line of best fit for two sets of data. It differs from Model I regression in that it is derived in a way that factors in for error in the x-axis, as well as the y-axis. Cornbleet Gochman (1979) refer to it as 'Model II regression'.

- Model II regression [Test V Test] In this type of analysis,both of the measurement methods are test methods, with both expected to be subject to error. Deming regression is an approach to model II regression.
- Model I regression [Criterion v Test] [Cornbleet Gochman 1979] define this analysis as the case in which the independent variable, X, is measured without error, with y as the dependent variable.
- In method comparison studies, the X variable is a precisely measured reference method. In the [Cornbleet Gochman 1979] paper It is argued that criterion may be regarded as the correct value. Other papers dispute this.

1.6 Model I and II Regression

On account of the fact that one set of measurements are linearly related to another, one could surmise that Linear Regression is the most suitable approach to analyzing comparisons. This approach is unsuitable on two counts. Firstly one of the assumptions of Regression analysis is that the independent variable values are without error. In method comparison studies one must assume the opposite; that there is error present in the measurements. Secondly a regression of X on Y would yield and entirely different result from Y on X.

Model I regression is unsuitable for method comparison studies. Even in the case where one method is a gold standard , it is disputed as to whether it is a valid approach. Model II regression is suitable for method comparison studies, but it is more difficult to execute. Both Model I and II regression models are unduly influenced by outliers. Regression Models can not easily be used to analyze repeated measurements

Cochrane and Cornbleet argue for the use of methods that based on the assumption that both methods are imprecisely measured ,and that yield a fitting that is consistent with both 'X on Y' and 'Y on X' formulations. These methods uses alternatives to the OLS approach to determine the slope and intercept.

They describe three such alternative methods of regression; Deming, Mandel, and Bartlett regression. Collectively the authors refer to these approaches as Model II regression techniques.

1.6.1 Model I regression [Criterion v Test]

[Cornbleet Gochman 1979] define this analysis as the case in which the independent variable, X, is measured without error, with y as the dependent variable.

In method comparison studies, the X variable is a precisely measured reference method. In the [Cornbleet Gochman 1979] paper. It is argued that criterion may be regarded as the correct value. Other papers dispute this.

Simple Linear Regression is well known statistical technique, wherein estimates for slope and intercept of the line of best fit are derived according to the Ordinary Least Square (OLS) principle. This method is known to Cornbleet and Cochrane as Model I regression.

Simple linear regression is defined as such with the name 'Model I regression' by Cornbleet Gochman (1979), in contrast to 'Model II regression'.

1.6.2 Model II regression [Test V Test]

In this type of analysis, both of the measurement methods are test methods, with both expected to be subject to error. Deming regression is an approach to model II regression.

In Model I regression, the independent variable is assumed to be measured without error. For method comparison studies, both sets of measurement must be assumed to be measured with imprecision and neither case can be taken to be a reference method. Arbitrarily selecting either method as the reference will yield two conflicting outcomes. A fitting based on 'X on Y' will give inconsistent results with a fitting based on 'Y on X'. Consequently model I regression is inappropriate for such cases.

Simple linear regression is defined as such with the name 'Model I regression' by Cornbleet Gochman (1979), in contrast to 'Model II regression'.

Simple linear regression calculates a line of best fit for two sets of data, n which the independent variable, X, is measured without error, with y as the dependent variable.

SLR (Model I) regression is considered by many Altman and Bland (1983); Cornbleet and Cochrane (1979); Ludbrook (1997) to be wholly unsuitable for method comparison studies, although recommended for use in calibration studies [Corncoch]. Even in the case where one method is a gold standard , it is disputed as to whether it is a valid approach. Model II regression is more suitable for method comparison studies, but it is more difficult to execute. Both Model I and II regression models are unduly influenced by outliers. Regression Models can not be used to analyze repeated measurements

Conventional regression models are estimated using the ordinary least squares (OLS) technique, and are referred to as 'Model I regression'(Cornbleet and Cochrane, 1979; Ludbrook, 1997). A key feature of Model I models is that the independent variable is assumed to be measured without error. As often pointed out in several papers (Altman and Bland, 1983; Ludbrook, 1997), this assumption invalidates simple linear regression for use in method comparison studies, as both methods must be assumed to be measured with error.

The use of regression models that assumes the presence of error in both variables X and Y have been proposed for use instead. (Cornbleet and Cochrane, 1979; Ludbrook, 1997), These methodologies are collectively known as 'Model II regression'. They differ in the method used to estimate the parameters of the regression.

Regression estimates depend on formulation of the model. A formulation with one method considered as the X variable will yield different estimates for a formulation where it is the Y variable. With Model I regression,the models fitted in both cases will entirely different and inconsistent. However with Model II regression, they will be consistent and complementary.
Conversely, Cornbleet Cochrane state that when the independent variable X is a precisely measured reference method, Model I regression may be considered suitable. They qualify this statement by referring the X as *the correct value*, tacitly implying that there must still be some measurement error present. The validity of this approach has been disputed elsewhere.

This regression method also calculates a line of best fit for two sets of data. It differs from Model I regression in that it is derived in a way that factors in for error in the x-axis, as well as the y-axis. Cornbleet Gochman (1979) refer to it as 'Model II regression'.

1.6.3 Comparison of Model II regressions

Cornbleet and Cochrane comparing the three methods, citing studies by other authors, concluding that Deming regression is the most useful of these methods. They found the Bartlett method to be flawed in determining slopes.

However the author point out that *clinical laboratory measurements usually increase in absolute imprecision when larger values are measured.*However one of the assumptions that underline Deming and Mandel regression is constancy of the measurement errors throughout the range of values.

1.7 Regression Methods

Conventional regression models are estimated using the ordinary least squares (OLS) technique, and are referred to as 'Model I regression' (Cornbleet and Cochrane, 1979; Ludbrook, 1997). A key feature of Model I models is that the independent variable is assumed to be measured without error. As often pointed out in several papers (Altman and Bland, 1983; Ludbrook, 1997), this assumption invalidates simple linear regression for use in method comparison studies, as both methods must be assumed to be measured with error.

1.8 Contention

Several papers have commented that this approach is undermined when the basic assumptions underlying linear regression are not met, the regression equation, and consequently the estimations of bias are undermined. Outliers are a source of error in regression estimates.In method comparison studies, the X variable is a precisely measured reference method. Cornbleet Gochman (1979) argued that criterion may be regarded as the correct value. Other papers dispute this.

Chapter 2

Errors-in-variables Regression

2.1 Errors-in-variables models

Errors-in-variables models or measurement errors models are regression models that account for measurement errors in the independent variables, as well as the dependent variable.

2.2 Introduction

Deming regression method also calculates a line of best fit for two sets of data. It differs from simple linear regression in that it is derived in a way that factors in for error in the x-axis, as well as the y-axis.

2.3 Origin of Deming Regression

Deming regression is a type of error-in-variable regression approach that assumes that the ratio $\lambda = \sigma^2_{epsilon}/\sigma^2_{\eta}$ is known. This approach would be appropriate when errors in y and x are both caused by measurements, and the accuracy of measuring devices or procedures are known. The case when $\lambda = 1$ is also known as the ***orthogonal regression***.

Deming regression is a regression fitting approach that assumes error in both variables. The sum of squared distances from measured sets of values to the regression line is minimized at an angles specified by the ratio λ of the residual variance of both variables. The variance of the ratio,λ, specifies the angle. When λ is one, the angle is 45 degrees. In ordinary linear regression, the distances are minimized in the vertical directions (Linnet, 1999).

As stated previously, the fundamental flaw of simple linear regression is that it allows for measurement error in one variable only. This causes a downward biased slope estimate.

The Deming regression method also calculates a line of best fit for two sets of data. Deming Regression differs from simple linear regression in that it is derived in a way that factors in for error in the x-axis, as well as the y-axis.

2.4 Bivariate Least Squares Regression

Since there are errors in both methods, a regression technique that takes into account the individual errors in both axes (bivariate least-squares, BLS) should be used. In this paper, we demonstrate that the errors made in estimating the regression coefficients by the BLS method are fewer than with the ordinary least-squares (OLS) or weighted least-squares (WLS) regression techniques and that the coefficient can be considered normally distributed.

2.5 Deming's Regression

The most commonly known Model II methodology is known as Deming's Regression, (also known an Ordinary Least Product regression). Deming regression is recommended by Cornbleet and Cochrane (1979) as the preferred Model II regression for use in method comparison studies.

As previously noted, the Bland Altman Plot is uninformative about the comparative influence of proportional bias and fixed bias. Deming's regression provides independent tests for both types of bias.

Deming regression method also calculates a line of best fit for two sets of data. It differs from simple linear regression in that it is derived in a way that factors in for error in the x-axis, as well as the y-axis.

2.6 Deming Regression

The Intercept and Slope are calculated according to Combleet & Gochman, 1979. The standard errors and confidence intervals are estimated using the jackknife method (Armitage et al., 2002). The 95% confidence interval for the Intercept can be used to test the hypothesis that A=0. This hypothesis is accepted if the confidence interval for A contains the value 0. If the hypothesis is rejected, then it is concluded that A is significantly different from 0 and both methods differ at least by a constant amount. The 95% confidence interval for the Slope can be used to test the hypothesis that B=1. This hypothesis is accepted if the confidence interval for B contains the value 1. If the hypothesis is rejected, then it is concluded that B is significantly different from 1 and there is at least a proportional difference between the two methods.

2.7 Deming Regression

- Informative analysis for the purposes of method comparison, Deming Regression is a regression technique taking into account uncertainty in both the independent and dependent variables.
- Demings method always results in one regression fit, regardless of which variable takes the place of the predictor variables.
- The measurement error (lambda or λ) is specified with measurement error vari-

ance related as

$$\lambda = \sigma_y^2/\sigma_x^2$$

(where σ_x^2 and σ_y^2 is the measurement error variance of the x and y variables, respectively).

- In the case where λ is equal to one, (i.e. equal error variances), the methodology is equivalent to ***orthogonal regression***.

- Deming approaches the matter by simultaneously minimizing the sum of the square of the residuals of both variables. This derivation results in the best fit to minimize the sum of the squares of the perpendicular distances from the data points.

- To compute the slope by Demings formula, normally distributed error of both variables is assumed, as well as a constant level of imprecision throughout the range of measurements.

2.8 Deming Regression

Deming regression is a regression fitting approach that assumes error in both variables. Deming regression is recommended by Cornbleet and Cochrane (1979) as the preferred Model II regression for use in method comparison studies. The sum of squared distances from measured sets of values to the regression line is minimized at an angles specified by the ratio λ of the residual variance of both variables.

The sum of squared distances from measured sets of values to the regression line is minimized at an angles specified by the ratio λ of the residual variance of both variables. The variance of the ratio,λ, specifies the angle. When λ is one, the angle is 45 degrees. In ordinary linear regression, the distances are minimized in the vertical directions (Linnet, 1999).

As stated previously, the fundamental flaw of simple linear regression is that it allows for measurement error in one variable only. This causes a downward biased slope estimate.

When λ is one, the angle is 45 degrees. In ordinary linear regression, the distances are minimized in the vertical directions (Linnet, 1999). In cases involving only single measurements by each method, λ may be unknown and is therefore assumes a value of one. While this will bias the estimates, it is less biased than ordinary linear regression.

2.9 Deming Regression

The Deming regression line is estimated by minimizing the sums of squared deviations in both the x and y directions at an angle determined by the ratio of the analytical standard deviations for the two methods.

This ratio can be estimated if multiple measurements were taken with each method, but if only one measurement was taken with each method, it can be assumed to be equal to one.

2.10 Deming Regression for MCS

As noted before, Deming regression is an important and informative methodology in method comparison studies. For single measurement method comparisons, Deming regression offers a useful complement to LME models.

2.11 Drawbacks of Deming Regression

Deming's Regression suffers from some crucial drawback. Firstly it is computationally complex, and it requires specific software packages to perform calculations. Secondly it is uninformative about the comparative precision of two methods of measurement. Most importantly Carroll and Ruppert (1996) states that

Deming's regression is acceptable only when the precision ratio (λ, in their paper as η) is correctly specified, but in practice this is often not the case, with the λ being underestimated. This underestimation leads to an overcorrection for attenuation.

2.12 Diagnostics for Deming Regression

Model selection and diagnostic technique are well developed for classical linear regression methods. Typically an implementation of a linear model fit will be accompanied by additional information, such as the coefficient of determination and likelihood and information criterions, and a regression ANOVA table. Such additional information has not, as yet, been implemented for Deming regression.

2.13 Single measurements

In cases involving only single measurements by each method, λ may be unknown and is therefore assumes a value of one. While this will bias the estimates, it is less biased than ordinary linear regression.

2.14 Performance in the presence of outliers

All least square estimation methods are sensitive to outliers. In common with all regression methods, Deming regression is vulnerable to outliers.

Bland Altman's 1986 paper contains a data set, measurement of mean velocity of circumferential fibre shortening (VCF) by the long axis and short axis in M-mode echocardiography. Evident in this data set are outliers. Choosing the most noticeable, we shall use the deming regression method on this data set, both with and without this outlier, to assess its influence.

- In the presence of the outlier, the intercept and slope are estimated to be -0.0297027 and 1.0172959 respectively.

- Without the outlier the intercept and slope are estimated to be -0.11482220 and 1.09263112 respectively.

- We therefore conclude that Deming Regression is adversely affected by outliers , in the same way model I regression is.

2.15 Deming Regression : Parameters and Estimation

2.15.1 RMSE

The root mean squared error is an estimate of the total error of the slope and includes the random error and the systemic error.

The root mean square error, RMSE, is given by

$$\text{RMSE} = \sqrt{\sum (b-1)^2/nruns} = \sqrt{(\text{Bias})^2 + (\text{SE})^2}$$

where 'nruns' is the number of runs.

2.15.2 Estimating the Variance ratio

$$x_i = \mu + \beta_0 + \epsilon_{xi}$$
$$y_i = \mu + \beta_1 + \epsilon_{yi}$$

The inter-method bias is the difference of these biases. In order to determine an estimate for the residual variances, one of the method biases must be assumed to be zero, i.e. $\beta_0 = 0$. The inter-method bias is now represented by β_1.

$$
\begin{aligned}
x_i &= \mu + \epsilon_{xi} \\
y_i &= \mu + \beta_1 + \epsilon_{yi}
\end{aligned}
$$

The residuals can be expressed as

$$
\begin{aligned}
\epsilon_{xi} &= x_i - \mu \\
\epsilon_{yi} &= y_i - (\mu + \beta_1)
\end{aligned}
$$

The variance of the residuals are equivalent to the variance of the corresponding observations, $\sigma^2_{\epsilon x} = \sigma^2_x$ and $\sigma^2_{\epsilon y} = \sigma^2_y$.

$$
\lambda = \frac{\sigma^2_{yx}}{\sigma^2_y}. \tag{2.1}
$$

Assuming constant standard deviations, and given duplicate measurements, the analytical standard deviations are given by

$$
SD^2_{ax} = \frac{1}{2n}\sum(x_{2i} - x_{1i})^2
$$
$$
SD^2_{ay} = \frac{1}{2n}\sum(y_{2i} - y_{1i})^2
$$

2.16 Estimating the variance ratio (Linnet)

Using duplicate measurements, one can estimate the analytical standard deviations and compute their ratio. This ratio is then used for computing the slope by the Deming method.[Linnet]

2.17 Inference Procedures

A 95% confidence interval for the intercept estimate can be used to test the intercept, and hence fixed bias, is equal to zero. This hypothesis is accepted if the confidence interval for the estimate contains the value 0 in its range. Should this be, it can be concluded that fixed bias is not present. Conversely, if the hypothesis is rejected, then it is concluded that the intercept is non zero, and that fixed bias is present.

Testing for proportional bias is a very similar procedure. The 95% confidence interval for the slope estimate can be used to test the hypothesis that the slope is equal to 1. This hypothesis is accepted if the confidence interval for the estimate contains the value 1 in its range. If the hypothesis is rejected, then it is concluded that the slope is significant different from 1 and that a proportional bias exists.

- The 95% confidence interval for the Intercept can be used to test the hypothesis that A=0. This hypothesis is accepted if the confidence interval for A contains the value 0.
- If the hypothesis is rejected, then it is concluded that A is significant different from 0 and both methods differ at least by a constant amount.
- The 95% confidence interval for the Slope can be used to test the hypothesis that B=1. This hypothesis is accepted if the confidence interval for B contains the value 1.
- If the hypothesis is rejected, then it is concluded that B is significant different from 1 and there is at least a proportional difference between the two methods. Cochrane Cornbleet
- The authors make the distinction between model I and model II regression types.
- Model II regression is the appropriate type when the predictor variable x is measured with imprecision. Cornbleet and Cochrane remark that clinical laboratory

measurements usually increase in absolute imprecision when larger values are measured.

A 95% confidence interval for the intercept estimate can be used to test the intercept, and hence fixed bias, is equal to zero. This hypothesis is accepted if the confidence interval for the estimate contains the value 0 in its range. Should this be, it can be concluded that fixed bias is not present. Conversely, if the hypothesis is rejected, then it is concluded that the intercept is non zero, and that fixed bias is present.

Testing for proportional bias is a very similar procedure. The 95% confidence interval for the slope estimate can be used to test the hypothesis that the slope is equal to 1. This hypothesis is accepted if the confidence interval for the estimate contains the value 1 in its range. If the hypothesis is rejected, then it is concluded that the slope is significant different from 1 and that a proportional bias exists.

- Model II regression is the appropriate type when the predictor variable x is measured with imprecision. Cornbleet and Cochrane remark that clinical laboratory measurements usually increase in absolute imprecision when larger values are measured.

Guidelines

Always plot the data. Suspected outliers may be identified from the scatter plot. S_{ex} represents the precision of a single x measurement near the mean value of X

$$\lambda = fracS^2_{ex}S^2_{ey}$$

2.18 Using LME models to estimate the ratio (BXC)

$$y_{mi} = \mu + \beta_m + b_i + \epsilon_{mi}$$

with β_m is a fixed effect for the method m and b_i is a random effect associated with patient i, and ϵ_{mi} as the measurement error. This is a simple single level LME model. Pinheiro and Bates (1994) provides for the implementation of fitting a model. The variance ratio of the residual variances is immediately determinable from the output. This variance ratio can be use to fit a Deming regression, as described in chapter 1.

2.19 Zhange Example

For convenience, a new data set shall be introduced to demonstrate Deming regression. Measurements of transmitral volumetric flow (MF) by doppler echocardiography, and left ventricular stroke volume (SV) by cross sectional echocardiography in 21 patients with aortic valve disease are tabulated in Zhang et al. (1986). This data set features in the discussion of method comparison studies in Altman (1991, p.398) .

Patient	MF (cm^3)	SV (cm^3)	Patient	MF (cm^3)	SV (cm^3)	Patient	MF (cm^3)	SV (cm^3)
1	47	43	8	75	72	15	90	82
2	66	70	9	79	92	16	100	100
3	68	72	10	81	76	17	104	94
4	69	81	11	85	85	18	105	98
5	70	60	12	87	82	19	112	108
6	70	67	13	87	90	20	120	131
7	73	72	14	87	96	21	132	131

Table 2.19.1: Transmitral volumetric flow(MF) and left ventricular stroke volume (SV) in 21 patients. (Zhang et al 1986)

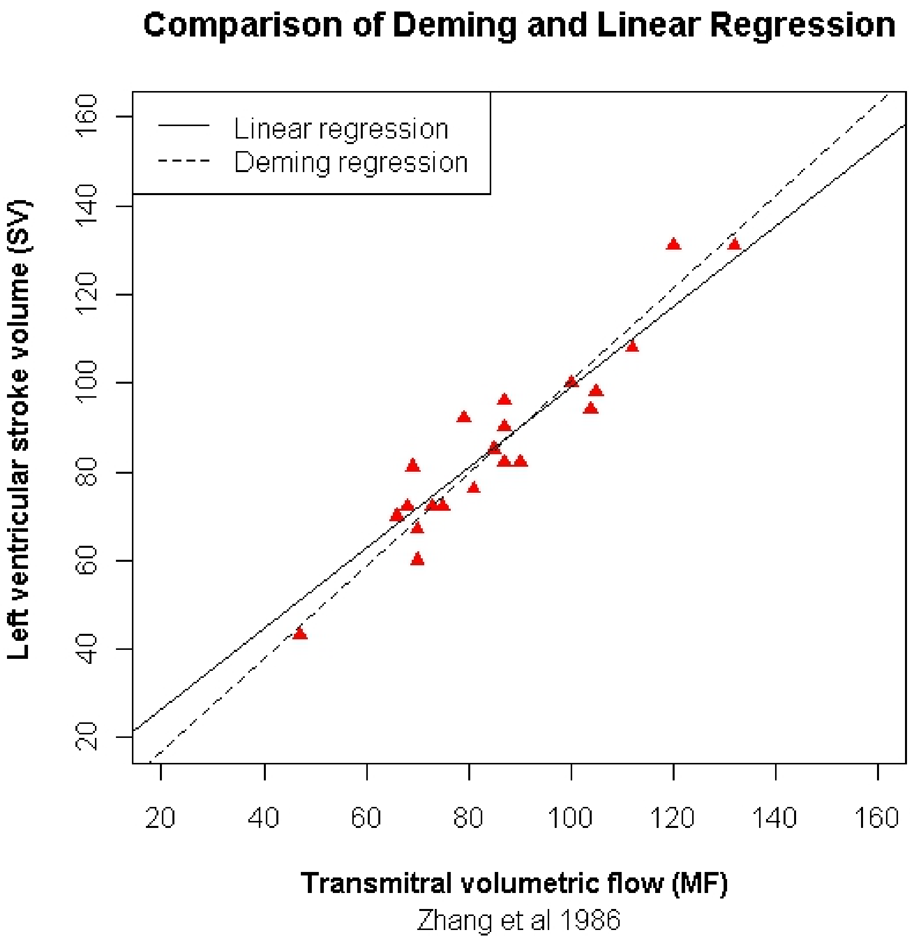

Figure 2.19.1: Deming Regression For Zhang's Data

2.20 Kummel's Estimates

For a given λ, Kummel (1879) derived the following estimate that would later be used for the Deming regression slope parameter. The intercept estimate α is simply estimated in the same way as in conventional linear regression, by using the identity $\bar{Y} - \hat{\beta}\bar{X}$;

$$\hat{\beta} = \frac{S_{yy} - \lambda S_{xx} + [(S_{yy} - \lambda S_{xx})^2 + 4\lambda S_{xy}^2]^{1/2}}{2S_{xy}} \tag{2.2}$$

, with λ as the variance ratio. As stated previously λ is often unknown, and therefore must be assumed to equal one.

For a given λ, Kummel (1879) derived the following estimate that would later be used for the Deming regression slope parameter. The intercept estimate α is simply estimated in the same way as in conventional linear regression, by using the identity $\bar{Y} - \hat{\beta}\bar{X}$;

$$\hat{\beta} = \frac{S_{yy} - \lambda S_{xx} + [(S_{yy} - \lambda S_{xx})^2 + 4\lambda S_{xy}^2]^{1/2}}{2S_{xy}} \tag{2.3}$$

, with λ as the variance ratio. As stated previously λ is often unknown, and therefore must be assumed to equal one.

Carroll and Ruppert (1996) states that Deming regression is acceptable only when the precision ratio (λ,in their paper as η) is correctly specified, but in practice this is often not the case, with the λ being underestimated. Several candidate models, with varying variance ratios may be fitted, and estimates of the slope and intercept are produced. However no model selection information is available to determine the best fitting model.

As with conventional regression methodologies, Deming regression calculates an estimate for both the slope and intercept for the fitted line, and standard errors thereof. Therefore there is sufficient information to carry out hypothesis tests on both estimates, that are informative about presence of fixed and proportional bias.

A 95% confidence interval for the intercept estimate can be used to test the intercept, and hence fixed bias, is equal to zero. This hypothesis is accepted if the confidence interval for the estimate contains the value 0 in its range. Should this be, it can be concluded that fixed bias is not present. Conversely, if the hypothesis is rejected, then it is concluded that the intercept is non zero, and that fixed bias is present.

Testing for proportional bias is a very similar procedure. The 95% confidence interval for the slope estimate can be used to test the hypothesis that the slope is equal to 1. This hypothesis is accepted if the confidence interval for the estimate contains the value 1 in its range. If the hypothesis is rejected, then it is concluded that the slope is significant different from 1 and that a proportional bias exists.

`Use of deming regression in method comparison studies.`

`Henk Konings`

Accuracy is closeness to the true value, or alternatively, having a low measurement error.

The determination of a true value for a biological specimen is difficult and sometimes impossible.

Precision is expressed in terms of standard deviation, coefficient of variance or variance.

In Deming regression, the errors between methods are assigned to both methods in proportion to the variances of the methods.

2.21 Implementations

Thus far, one of the few `R` implementations of Deming regression is contained in the 'MethComp' package. (Carstensen et al., 2008).

Unless specified otherwise, the variance ratio λ has a default value of one. A means of computing likelihood functions would potentially allow for an algorithm for estimating the true variance ratio.

2.22 Performance in the presence of outliers

All least square estimation methods are sensitive to outliers. In common with all regression methods, Deming regression is vulnerable to outliers.

Bland Altman's 1986 paper contains a data set, measurement of mean velocity of circumferential fibre shortening (VCF) by the long axis and short axis in M-mode echocardiography. Evident in this data set are outliers. Choosing the most noticeable, we shall use the deming regression method on this data set, both with and without this outlier, to assess its influence.

- In the presence of the outlier, the intercept and slope are estimated to be -0.0297027 and 1.0172959 respectively.

- Without the outlier the intercept and slope are estimated to be -0.11482220 and 1.09263112 respectively.

- We therefore conclude that Deming Regression is adversely affected by outliers , in the same way model I regression is.

2.23 Using LMEs to estimate the ratio

$$y_{mi} = \mu + \beta_m + b_i + \epsilon_{mi}$$

with β_m is a fixed effect for the method m and b_i is a random effect associated with patient i, and ϵ_{mi} as the measurement error.

This is a simple single level LME model. Pinheiro and Bates (1994) provides for the implementation of fitting a model.

The variance ratio of the residual variances is immediately determinable from the output. This variance ratio can be use to fit a Deming regression, as described in chapter 1.

2.24 Least Products Regression

Used as an alternative to Bland-Altman Analysis, this method is also known as 'Geometric Mean Regression' and 'Reduced Major Axis Regression'. This regression model minimizes the areas of the right triangles formed by the data points' vertical and horizontal deviations from the fitted line and the fitted line.

- Model II regression analysis caters for cases in which random error is attached to both dependent and independent variables. Comparing methods of measurement is just such a case.(Ludbrook)

- Least products regression is the reviewer's preferred technique for analysing the Model II case. In this, the sum of the products of the vertical and horizontal deviations of the x,y values from the line is minimized.

- Least products regression analysis is suitable for calibrating one method against another. It is also a sensitive technique for detecting and distinguishing fixed and proportional bias between methods.

Least-products regression can lead to inflated SEEs and estimates that do not tend to their true values an N approaches infinity (Draper and Smith, 1998).

2.25 Deming Regression

- Informative analysis for the purposes of method comparison, Deming Regression is a regression technique taking into account uncertainty in both the independent and dependent variables.

- Demings method always results in one regression fit, regardless of which variable takes the place of the predictor variables.

- The measurement error (lambda or λ) is specified with measurement error variance related as

$$\lambda = \sigma_y^2/\sigma_x^2$$

  (where σ_x^2 and σ_y^2 is the measurement error variance of the x and y variables, respectively).

- In the case where λ is equal to one, (i.e. equal error variances), the methodology is equivalent to ***orthogonal regression***.

- Deming approaches the matter by simultaneously minimizing the sum of the square of the residuals of both variables. This derivation results in the best fit to minimize the sum of the squares of the perpendicular distances from the data points.

- To compute the slope by Demings formula, normally distributed error of both variables is assumed, as well as a constant level of imprecision throughout the range of measurements.

2.26 Deming Regression- Line of Best Fit

Deming regression method also calculates a line of best fit for two sets of data. It differs from simple linear regression in that it is derived in a way that factors in for error in the x-axis, as well as the y-axis.

2.27 Weighted Deming Regression

Weighted Linear Regression

Weighted linear regression allows for non-constancy of the standard deviation of the y variable. However it is assumed that X is without measurement error. Weighted Deming regression takes into account the non-constant proportional measurement errors in both variables.Despite the non-constancy, it is necessary to retain the constant value of λ.

In **both forms** of Deming regression, λ is assumed to be constant through out the range of measurements. For WDR weights w_i are used to compute the sums of squares and cross products. The weights are inversely proportional to the squared analytical variance at any given value.

2.28 Using LMEs to estimate the ratio

$$y_{mi} = \mu + \beta_m + b_i + \epsilon_{mi}$$

with β_m is a fixed effect for the method m and b_i is a random effect associated with patient i, and ϵ_{mi} as the measurement error.

This is a simple single level LME model. Pinheiro and Bates (1994) provides for the implementation of fitting a model.

The variance ratio of the residual variances is immediately determinable from the output. This variance ratio can be use to fit a Deming regression, as described in chapter 1.

sectionEstimating the variance ratio

2.29 Methods

Theoretical calculations and simulations were used to consider the statistical power for detection of slope deviations from unity and intercept deviations from zero. For situations with proportional analytical standard deviations, weighted forms of regression analysis were evaluated.

2.30 Results

In general, sample sizes of 40100 samples conventionally used in method comparison studies often must be reconsidered. A main factor is the range of values, which should be as wide as possible for the given analyte. For a range ratio (maximum value divided by minimum value) of 2, 544 samples are required to detect one standardized slope deviation; the number of required samples decreases to 64 at a range ratio of 10 (proportional analytical error). For electrolytes having very narrow ranges of values, very large sample sizes usually are necessary. In case of proportional analytical error, application of a weighted approach is important to assure an efficient analysis; e.g., for a range ratio of 10, the weighted approach reduces the requirement of samples by ¿50

2.31 Conclusions

Estimation of the necessary sample size for a method comparison study assures a valid result; either no difference is found or the existence of a relevant difference is confirmed.

2.32 Rejection Rule

Rejection rule for outliers.

2.33 Weighted Deming Regression

Weighted linear regression allows for non-constancy of the standard deviation of the y variable. However it is assumed that X is without measurement error. Weighted Deming regression takes into account the non-constant proportional measurement errors in both variables.Despite the non-constancy, it is necessary to retain the constant value of λ.

In **both forms** of Deming regression, λ is assumed to be constant through out the range of measurements. For WDR weights w_i are used to compute the sums of squares and cross products. The weights are inversely proportional to the squared analytical variance at any given value.

2.34 Least Product Regression

- Least Product Regression , also known as 'Model II regression' caters for cases in which random error is attached to both dependent and independent variables. Ludbrook cites this methodology as being pertinent to Method comparison studies.

- The sum of the products of the vertical and horizontal deviations of the x,y values from the line is minimized.

- Least products regression analysis is considered suitable for calibrating one method against another.Ludbrook comments that it is also a sensitive technique for detecting and distinguishing fixed and proportional bias between methods.

- Proposed as an alternative to Bland & Altman methodology ,this method is also known as 'Geometric Mean Regression' and 'Reduced Major Axis Regression'.

2.35 Least Products Regression

Used as an alternative to Bland-Altman Analysis, this method is also known as 'Geometric Mean Regression' and 'Reduced Major Axis Regression'. This regression model minimizes the areas of the right triangles formed by the data points' vertical and horizontal deviations from the fitted line and the fitted line.

- Model II regression analysis caters for cases in which random error is attached to both dependent and independent variables. Comparing methods of measurement is just such a case.(Ludbrook)

- Least products regression is the reviewer's preferred technique for analysing the Model II case. In this, the sum of the products of the vertical and horizontal deviations of the x,y values from the line is minimized.

- Least products regression analysis is suitable for calibrating one method against another. It is also a sensitive technique for detecting and distinguishing fixed and proportional bias between methods.

Least-products regression can lead to inflated SEEs and estimates that do not tend to their true values an N approaches infinity (Draper and Smith, 1998).

Difference with Least Squares Regression

Least-products regression can lead to inflated SEEs and estimates that do not tend to their true values an N approaches infinity (Draper and Smith, 1998).

2.36 Linnet - References

The statistical procedures are described in: Linnet K. Necessary sample size for method comparison studies based on regression analysis. Clin Chem 1999; 45: 882-94. Linnet K. Performance of Deming regression analysis in case of misspecified analytical error

ratio in method comparison studies. Clin Chem 1998; 44: 1024-1031. Linnet K. Evaluation of regression procedures for methods comparison studies. Clin Chem 1993; 39. 424-432. Linnet K. Estimation of the linear relationship between measurements of two methods with proportional errors. Stat Med 1990; 9: 1463-1473.

2.37 Drawbacks

Deming's Regression suffers from some crucial drawback. Firstly it is computationally complex, and it requires specific software packages to perform calculations.Secondly it is uninformative about the comparative precision of two methods of measurement. Most importantly Carroll and Ruppert (1996) states that Deming's regression is acceptable only when the precision ratio (λ, in their paper as η) is correctly specified ,but in practice this is often not the case, with the λ being underestimated.

2.38 Bootstap Techniques

Use of Bootstap Techniques to obtain Confidence Interval estimates

2.39 References

Carpenter, J., Bithell, J. (2000) Bootstrap condence intervals: when, which, what? A practical guide for medical statisticians. Stat Med, 19 (9), 11411164.

2.40 Estimating the Variance ratio

$$x_i = \mu + \beta_0 + \epsilon_{xi}$$

$$y_i = \mu + \beta_1 + \epsilon_{yi}$$

The inter-method bias is the difference of these biases. In order to determine an estimate for the residual variances, one of the method biases must be assumed to be zero, i.e. $\beta_0 = 0$. The inter-method bias is now represented by β_1.

$$
\begin{aligned}
x_i &= \mu + \epsilon_{xi} \\
y_i &= \mu + \beta_1 + \epsilon_{yi}
\end{aligned}
$$

The residuals can be expressed as

$$
\begin{aligned}
\epsilon_{xi} &= x_i - \mu \\
\epsilon_{yi} &= y_i - (\mu + \beta_1)
\end{aligned}
$$

The variance of the residuals are equivalent to the variance of the corresponding observations, $\sigma^2_{\epsilon x} = \sigma^2_x$ and $\sigma^2_{\epsilon y} = \sigma^2_y$.

$$\lambda = \frac{\sigma^2_{yx}}{\sigma^2_y}. \tag{2.4}$$

Assuming constant standard deviations, and given duplicate measurements, the analytical standard deviations are given by

$$
\begin{aligned}
SD^2_{ax} = \frac{1}{2n}\sum(x_{2i} - x_{1i})^2 \\
SD^2_{ay} = \frac{1}{2n}\sum(y_{2i} - y_{1i})^2
\end{aligned}
$$

Using duplicate measurements, one can estimate the analytical standard deviations and compute their ratio. This ratio is then used for computing the slope by the Deming method.[Linnet]

2.41 determining lambda

assuming constant standard deviations, and given duplicate measurements, the analytical standard deviations are given by

$$SD^2_{ax} = \frac{1}{2n}\sum(x_{2i} - x_{1i})^2$$
$$SD^2_{ay} = \frac{1}{2n}\sum(y_{2i} - y_{1i})^2$$

2.42 performance in the presence of oultiers

All least square estimation methods are sensitive to outliers.

2.43 weighted least square regression

The constancy of variance is a necessary assumption for ordinary linear regression.

$$\mathrm{SD}_{ay} = ch(x_i) \tag{2.5}$$

2.44 Weighted Deming Regression

Weighted linear regression allows for non-constancy of the standard deviation of the y variable. However it is assumed that X is without measurement error. Weighted Deming regression takes into account the non-constant proportional measurement errors in both variables.Despite the non-constancy, it is necessary to retain the constant value of λ.

In **both forms** of Deming regression, λ is assumed to be constant through out the range of measurements. For WDR weights w_i are used to compute the sums of squares and cross products. The weights are inversely proportional to the squared analytical variance at any given value.

2.45 Variance Ratio

The Bland Altman Plot is uninformative about the comparative influence of proportional bias and fixed bias. Model II approaches, such as Deming regression, can provide independent tests for both types of bias.

As with conventional regression methodologies, Deming regression calculates an estimate for both the slope and intercept for the fitted line, and standard errors thereof. Therefore there is sufficient information to carry out hypothesis tests on both estimates, that are informative about presence of fixed and proportional bias.

2.46 Ordinary Least Product Regression

Ludbrook (1997) states that the grouping structure can be straightforward, but there are more complex data sets that have a hierarchical(nested) model.

Observations between groups are independent, but observations within each groups are dependent because they belong to the same subpopulation. Therefore there are two sources of variation: between-group and within-group variance. Mean correction

is a method of reducing bias.

Chapter 3

REMOVE

3.1 Simple Linear Regression

Simple linear regression is defined as such with the name ‘Model I regression’ by Cornbleet Gochman (1979), in contrast to ’Model II regression’.

On account of the fact that one set of measurements are linearly related to another, one could surmise that Linear Regression is the most suitable approach to analyzing comparisons. This approach is unsuitable on two counts. Firstly one of the assumptions of Regression analysis is that the independent variable values are without error. In method comparison studies one must assume the opposite; that there is error present in the measurements. Secondly a regression of X on Y would yield and entirely different result from Y on X.

Simple linear regression calculates a line of best fit for two sets of data, n which the independent variable, X, is measured without error, with y as the dependent variable.

SLR (Model I) regression is considered by many Altman and Bland (1983); Cornbleet and Cochrane (1979); Ludbrook (1997) to be wholly unsuitable for method comparison studies, although recommended for use in calibration studies [Corncoch]. Even in the case where one method is a gold standard , it is disputed as to whether it is a valid approach. Model II regression is more suitable for method comparison studies, but it is more difficult to execute. Both Model I and II regression models are unduly influenced by outliers. Regression Models can not be used to analyze repeated measurements

Regression Analysis

Another inappropriate approach is the regressing one set of measurements against the other. According to this methodology the measurement methods could considered equivalent if the confidence interval for the regression coefficient included 1. Analysts sometimes use least squares (referred to by Ludbrook as Model I) regression analysis to calibrate one method of measurement against another. In this technique, the sum of the squares of the vertical deviations of y values from the line is minimized. This approach is invalid, because both y and x values are attended by random error.

The Identity Plot

This is a simple graphical approach, advocated by Bland and Altman (1986), that yields a cursory examination of how well the measurement methods agree. In the case of good agreement, the co-variates of the plot accord closely with the $X = Y$ line.

Advantages of Regression Approaches for MCS

- These methods can be employed in conversion problems.
- Bland and Altman have stated that regression analysis offers insights into MCS problems.

Disadvantages

- Regression methods are uninformative about the variability of the differences.
- Regression methods can determine the presence of bias, and the levels of constant bias and proportional bias thereof Ludbrook (1997, 2002).

3.2 Constant and Proportional Bias

Linear Regression is a commonly used technique for comparing paired assays. The Intercept and Slope can provide estimates for the constant bias and proportional bias occurring between both methods. If the basic assumptions underlying linear regression are not met, the regression equation, and consequently the estimations of bias are undermined. Outliers are a source of error in regression estimates.

Constant or proportional bias in method comparison studies using linear regression can be detected by an individual test on the intercept or the slope of the line regressed from the results of the two methods to be compared.

Bartko's Discussion of BB

Let $y = X_1 - X_2$ and $x = (X_1 - X_2)/2$. The Bradley-Blackwood procedure fits y on x, such that

$$y = \beta_0 + \beta_1 x$$

The slope and intercepte are given bu

$$beta_1 = \frac{(\sigma_1^2 = \sigma_2^2)}{2\sigma_x^2}$$

Pitman's Test on Correlated variances

$H_0 \, : \, \sigma_1^2 = \sigma_2^2$

$H_0 \, : \, \sigma_1^2 = \sigma_2^2$

Pitman's test is identical to the slope equal to zero in the regression of y on x.

3.3 Bradley-Blackwood Test (Kevin Hayes Talk)

This work considers the problem of testing $\mu_1 = \mu_2$ and $\sigma_1^2 = \sigma_2^2$ using a random sample from a bivariate normal distribution with parameters $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

The new contribution is a decomposition of the Bradley-Blackwood test statistic (*Bradley and Blackwood, 1989*)for the simultaneous test of $\mu_1 = \mu_2$; $\sigma_1^2 = \sigma_2^2$ as a sum of two statistics.

One is equivalent to the Pitman-Morgan (*Pitman, 1939; Morgan, 1939*) test statistic for $\sigma_1^2 = \sigma_2^2$ and the other one is a new alternative to the standard paired-t test of $\mu_D = \mu_1 = \mu_2 = 0$.

Surprisingly, the classic Student paired-t test makes no assumptions about the equality (or otherwise) of the variance parameters.

The power functions for these tests are quite easy to derive, and show that when $\sigma_1^2 = \sigma_2^2$, the paired t-test has a slight advantage over the new alternative in terms of power, but when $\sigma_1^2 \neq \sigma_2^2$, the new test has substantially higher power than the paired-t test.

While Bradley and Blackwood provide a test on the joint hypothesis of equal means and equal variances their regression based approach does not separate these two issues.

The rejection of the joint hypothesis may be due to two groups with unequal means and unequal variances; unequal means and equal variances, or equal means and unequal variances.

We propose an approach for resolving this (model selection) problem in a manner controlling the magnitudes of the relevant type I error probabilities.

3.4 Conclusions about Existing Methodologies

The Bland Altman methodology is well noted for its ease of use, and can be easily implemented with most software packages. Also it doesn't require the practitioner to have more than basic statistical training. The plot is quite informative about the variability of the differences over the range of measurements. For example, an inspection of the plot will indicate the 'fan effect'. They also can be used to detect the presence of an outlier.

Ludbrook (1997, 2002) criticizes these plots on the basis that they presents no information on effect of constant bias or proportional bias. These plots are only practicable when both methods measure in the same units. Hence they are totally unsuitable for conversion problems. The limits of agreement are somewhat arbitrarily constructed. They may or may not be suitable for the data in question. It has been found that the limits given are too wide to be acceptable. There is no guidance on how to deal with outliers. Bland and Altman recognize effect they would have on the limits of agreeement, but offer no guidance on how to correct for those effects.

There is no formal testing procedure provided. Rather, it is upon the practitioner opinion to judge the outcome of the methodology.

Introduction Quotes from Bartko's Paper

"Can two methods of measurement be used interchangeably?"

Bartko's Ellipse

$$\frac{x-\bar{x}}{\sigma_x^2} - \frac{2\rho(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} + \frac{y-\bar{y}}{\sigma_y^2} = \chi^2(2df_(1-\rho^2)$$

section*Remarks

- Pearson's Correlation of (x,y) is the same as Pitman's correlation of sums and differences.
- Techniques for plotting an ellipse can be found in Douglas Altman's book.

3.5 A regression based approach based on Bland Altman Analysis

Bland and Altman have stated that regression analysis offers insights into method comparison studies. Regression methods can determine the presence of bias, and the levels of constant bias and proportional bias thereof Ludbrook (1997, 2002). While they are informative about inter-method bias, Regression methods offer the analyst no insights into the relative precision of both methods. These methods can be employed in conversion problems, however errors are attended. ***Lu et al*** used such a technique in their comparison of DXA scanners. They also used the Blackwood Bradley test. However it was shown that, for particular comparisons, agreement between methods was indicated according to one test, but lack of agreement was indicated by the other.

Remarks

- Pearson's Correlation of (x,y) is the same as Pitman's correlation of sums and differences.
- Techniques for plotting an ellipse can be found in Douglas Altman's book.

3.6 The MCR R pacakge - Regression Techniques for MCS

The ***mcr*** packages provides a set of regression techniques to quantify the relation between two measurement methods.

In particular, it address regression problems with errors in both variables, but without repeated measurements. The ***mcr*** package follows the CLSI EP09-A3 recommendations for analytical method comparison and estimation of bias using patient samples.

*Methods featured in the **mcr** package*

- Deming Regression
- Weighted Deming Regression
- Passing-Bablock Regression

The *creatinine* gives the blood and serum preoperative creatinine measurements in 110 heart surgery patients.

```
library("mcr")
data("creatinine", package="mcr")
tail(creatinine)



fit.lr <- mcreg(as.matrix(creatinine), method.reg="LinReg", na.rm=TRUE)
fit.wlr <- mcreg(as.matrix(creatinine), method.reg="WLinReg", na.rm=TRUE)
compareFit( fit.lr, fit.wlr )
```

3.7 Implementation of Deming Regression with Rs

Thus far, one of the few R implementations of Deming regression is contained in the 'MethComp' package. (Carstensen et al., 2008).

Unless specified otherwise, the variance ratio λ has a default value of one. A means of computing likelihood functions would potentially allow for an algorithm for estimating the true variance ratio.

3.8 Linnet - References

The statistical procedures are described in: Linnet K. Necessary sample size for method comparison studies based on regression analysis. Clin Chem 1999; 45: 882-94. Linnet K. Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. Clin Chem 1998; 44: 1024-1031. Linnet K. Evaluation of regression procedures for methods comparison studies. Clin Chem 1993; 39. 424-432. Linnet K. Estimation of the linear relationship between measurements of two methods with proportional errors. Stat Med 1990; 9: 1463-1473.

Bibliography

Altman, D. and J. Bland (1983). Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician) 32*(3), 307–317.

Altman, D. G. (1991). *Practical Statistics for Medical Research*. Chapman and Hall.

Bland, J. and D. Altman (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet i*, 307–310.

Carroll, R. and D. Ruppert (1996). The use and misuse of orthogonal regression in linear errors-in-variables models. *The American Statistician 50*(1), 1–6.

Carstensen, B., J. Simpson, and L. C. Gurrin (2008). Statistical models for assessing agreement in method comparison studies with replicate measurements. *The International Journal of Biostatistics 4*(1).

Cornbleet, P. J. and D. Cochrane (1979). Regression methods for assessing agreement between two methods of clinical measurement. *Journal of Clinical Chemistry 24*(2), 342–345.

Kummel, C. (1879). Reduction of observation equations which contain more than one observed quantity. *The Analyst 6*, 97–105.

Linnet, K. (1999). Necessary sample size for method comparison studies based on regression analysis. *Clinical Chemistry 45*(6), 882–894.

Ludbrook, J. (1997). Comparing methods of measurement. *Clinical and Experimental Pharmacology and Physiology 24*, 193–203.

Ludbrook, J. (2002). Statistical techniques for comparing measurers and methods of measurement: a critcal review. *Clinical and Experimental Pharmacology and Physiology 29*, 527–536.

O'Brien, E., J. Petrie, W. Littler, M. de Swiet, P. L. Padfield, D. Altman, M. Bland, A. Coats, and N. Atkins (1990). The British Hypertension Society protocol for the evaluation of blood pressure measuring devices. *Journal of Hypertension 8*, 607–619.

Pinheiro, J. and D. Bates (1994). *Mixed Effects Models in S and S plus* (2nd ed.). Reading, Massachusetts: Springer.

Zhang, Y., S. Nitter-Hauge, H. Ihlen, K. Rootwelt, and E. Myhre (1986). Measurement of aortic regurgitation by doppler echocardiography. *British Heart Journal 55*, 32–38.