

0.1 Carstensen's Limits of agreement

? presents a methodology to compute the limits of agreement based on LME models. The method of computation is the same as Roy's model, but with the covariance estimates set to zero.

In cases where there is negligible covariance between methods, the limits of agreement computed using Roy's model accord with those computed using Carstensen's model. In cases where some degree of covariance is present between the two methods, the limits of agreement computed using models will differ. In the presented example, it is shown that Roy's LoAs are lower than those of Carstensen, when covariance is present.

Importantly, estimates required to calculate the limits of agreement are not extractable, and therefore the calculation must be done by hand.

1 Computing LoAs from LME models

One important feature of replicate observations is that they should be independent of each other. In essence, this is achieved by ensuring that the observer makes each measurement independent of knowledge of the previous value(s). This may be difficult to achieve in practice. (Check who said this)

2 New

Computing limits of agreement features prominently in many method comparison studies, further to ?? . ? addresses the issue of computing LoAs in the presence of replicate measurements, suggesting several computationally simple approaches. When repeated measures data are available, it is desirable to use all the data to compare the two methods. However, the original Bland-Altman method was developed for two sets of measurements done on one occasion (i.e. independent data), and so this approach is not suitable for replicate measures data. However, as a naive analysis, it may be used to explore the data because of the simplicity of the method. ? computes the limits of agreement to the case with replicate measurements by using LME models.

? formulates a very powerful method of assessing whether two methods of measurement, with replicate measurements, also using LME models. Roy's approach is based on the construction of variance-covariance matrices. Importantly, Roy's approach does not address the issue of limits of agreement (though another related analysis , the coefficient of repeatability, is mentioned).

This paper seeks to use Roy's approach to estimate the limits of agreement. These estimates will be compared to estimates computed under Carstensen's formulation.

In computing limits of agreement, it is first necessary to have an estimate for the standard deviations of the differences. When the agreement of two methods is analyzed using LME models, a clear method of how to compute the standard deviation is required. As the estimate for inter-method bias and the quantile would be the same

for both methodologies, the focus is solely on the standard deviation.

2.1 Roy's LOAs

The limits of agreement computed by Roy's method are derived from the variance covariance matrix for overall variability. This matrix is the sum of the between subject VC matrix and the within-subject VC matrix.

The standard deviation of the differences of methods x and y is computed using values from the overall VC matrix.

$$\text{var}(x - y) = \text{var}(x) + \text{var}(y) - 2\text{cov}(x, y)$$

The respective estimates computed by both methods are tabulated as follows. Evidently there is close correspondence between both sets of estimates.

? uses the "Blood" data set, which featured in ?.

3 Sampling Strategy

? describes the sampling method when discussing of a motivating example

Diabetes patients attending an outpatient clinic in Denmark have their HbA_{1c} levels routinely measured at every visit. Venous and Capillary blood samples were obtained from all patients appearing at the clinic over two days. Samples were measured on four consecutive days on each machines, hence there are five analysis days.

? notes that every machine was calibrated every day to the manufacturers guidelines. Measurements are classified by method, individual and replicate. In this case the replicates are clearly not exchangeable, neither within patients nor simulataneously for all patients.

3.1 Limits of Agreement in LME models

? uses LME models to determine the limits of agreement. Between-subject variation for method m is given by d_m^2 and within-subject variation is given by λ_m^2 . ? remarks that for two methods A and B , separate values of d_A^2 and d_B^2 cannot be estimated, only their average. Hence the assumption that $d_x = d_y = d$ is necessary. The between-subject variability \mathbf{D} and within-subject variability $\mathbf{\Lambda}$ can be presented in matrix form,

$$\mathbf{D} = \begin{pmatrix} d_A^2 & 0 \\ 0 & d_B^2 \end{pmatrix} = \begin{pmatrix} d^2 & 0 \\ 0 & d^2 \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_A^2 & 0 \\ 0 & \lambda_B^2 \end{pmatrix}.$$

The variance for method m is $d_m^2 + \lambda_m^2$. Limits of agreement are determined using the standard deviation of the case-wise differences between the sets of measurements by two methods A and B , given by

$$\text{var}(y_A - y_B) = 2d^2 + \lambda_A^2 + \lambda_B^2. \quad (1)$$

Importantly the covariance terms in both variability matrices are zero, and no covariance component is present.

? has demonstrated a methodology whereby d_A^2 and d_B^2 can be estimated separately. Also covariance terms are present in both \mathbf{D} and $\mathbf{\Lambda}$. Using Roy's methodology, the

variance of the differences is

$$\text{var}(y_{iA} - y_{iB}) = d_A^2 + \lambda_B^2 + d_A^2 + \lambda_B^2 - 2(d_{AB} + \lambda_{AB}) \quad (2)$$

All of these terms are given or determinable in computer output. The limits of agreement can therefore be evaluated using

$$\bar{y}_A - \bar{y}_B \pm 1.96 \times \sqrt{\sigma_A^2 + \sigma_B^2 - 2(\sigma_{AB})}. \quad (3)$$

For Carstensen's 'fat' data, the limits of agreement computed using Roy's method are consistent with the estimates given by ?; $0.044884 \pm 1.96 \times 0.1373979 = (-0.224, 0.314)$.

In computing limits of agreement, it is first necessary to have an estimate for the standard deviations of the differences. When the agreement of two methods is analyzed using LME models, a clear method of how to compute the standard deviation is required. As the estimate for inter-method bias and the quantile would be the same for both methodologies, the focus is solely on the standard deviation.

3.2 Carstensen's LOAs

Carstensen presents a model where the variation between items for method m is captured by σ_m and the within item variation by τ_m .

Further to his model, Carstensen computes the limits of agreement as

$$\hat{\alpha}_1 - \hat{\alpha}_2 \pm \sqrt{2\hat{\tau}^2 + \hat{\sigma}_1^2 + \hat{\sigma}_2^2}$$

3.3 Roy's LOAs

The limits of agreement computed by Roy's method are derived from the variance covariance matrix for overall variability. This matrix is the sum of the between subject VC matrix and the within-subject VC matrix.

The standard deviation of the differences of methods x and y is computed using values from the overall VC matrix.

$$\text{var}(x - y) = \text{var}(x) + \text{var}(y) - 2\text{cov}(x, y)$$

The respective estimates computed by both methods are tabulated as follows. Evidently there is close correspondence between both sets of estimates.

? formulates an LME model, both in the absence and the presence of an interaction term.? uses both to demonstrate the importance of using an interaction term. Failure to take the replication structure into account results in over-estimation of the limits of agreement. For the Carstensen estimates below, an interaction term was included when computed.

? uses the “Blood” data set, which featured in ?.

? describes the sampling method when discussing of a motivating example

Diabetes patients attending an outpatient clinic in Denmark have their HbA_{1c} levels routinely measured at every visit. Venous and Capillary blood samples were obtained from all patients appearing at the clinic over two days. Samples were measured on four consecutive days on each machines, hence there are five analysis days.

? notes that every machine was calibrated every day to the manufacturers guidelines. Measurements are classified by method, individual and replicate. In this case the replicates are clearly not exchangeable, neither within patients nor simultaneously for all patients.

3.4 Limits of Agreement in LME models

? uses LME models to determine the limits of agreement. Between-subject variation for method m is given by d_m^2 and within-subject variation is given by λ_m^2 . ? remarks that for two methods A and B , separate values of d_A^2 and d_B^2 cannot be estimated, only their average. Hence the assumption that $d_x = d_y = d$ is necessary. The between-subject variability \mathbf{D} and within-subject variability $\mathbf{\Lambda}$ can be presented in matrix form,

$$\mathbf{D} = \begin{pmatrix} d_A^2 & 0 \\ 0 & d_B^2 \end{pmatrix} = \begin{pmatrix} d^2 & 0 \\ 0 & d^2 \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_A^2 & 0 \\ 0 & \lambda_B^2 \end{pmatrix}.$$

The variance for method m is $d_m^2 + \lambda_m^2$. Limits of agreement are determined using the standard deviation of the case-wise differences between the sets of measurements by two methods A and B , given by

$$\text{var}(y_A - y_B) = 2d^2 + \lambda_A^2 + \lambda_B^2. \quad (4)$$

Importantly the covariance terms in both variability matrices are zero, and no covariance component is present.

? has demonstrated a methodology whereby d_A^2 and d_B^2 can be estimated separately. Also covariance terms are present in both \mathbf{D} and $\mathbf{\Lambda}$. Using Roy's methodology, the variance of the differences is

$$\text{var}(y_{iA} - y_{iB}) = d_A^2 + \lambda_B^2 + d_B^2 + \lambda_A^2 - 2(d_{AB} + \lambda_{AB}) \quad (5)$$

All of these terms are given or determinable in computer output. The limits of agreement can therefore be evaluated using

$$\bar{y}_A - \bar{y}_B \pm 1.96 \times \sqrt{\sigma_A^2 + \sigma_B^2 - 2(\sigma_{AB})}. \quad (6)$$

For Carstensen's 'fat' data, the limits of agreement computed using Roy's method are consistent with the estimates given by ?; $0.044884 \pm 1.96 \times 0.1373979 = (-0.224, 0.314)$.

3.5 Limits of agreement for Carstensen’s data

? describes the calculation of the limits of agreement (with the inter-method bias implicit) for both data sets, based on his formulation;

$$\hat{\alpha}_1 - \hat{\alpha}_2 \pm 2\sqrt{2\hat{\tau}^2 + \hat{\sigma}_1^2 + \hat{\sigma}_2^2}.$$

For the ‘Fat’ data set, the inter-method bias is shown to be 0.045. The limits of agreement are $(-0.23, 0.32)$

Carstensen demonstrates the use of the interaction term when computing the limits of agreement for the ‘Oximetry’ data set. When the interaction term is omitted, the limits of agreement are $(-9.97, 14.81)$. Carstensen advises the inclusion of the interaction term for linked replicates, and hence the limits of agreement are recomputed as $(-12.18, 17.12)$.

4 Computing LoAs from LME models

One important feature of replicate observations is that they should be independent of each other. In essence, this is achieved by ensuring that the observer makes each measurement independent of knowledge of the previous value(s). This may be difficult to achieve in practice.

4.1 Carstensen's Model

MORE

4.2 Carstensen's LOAs

Carstensen presents a model where the variation between items for method m is captured by σ_m and the within item variation by τ_m .

Further to his model, Carstensen computes the limits of agreement as

$$\hat{\alpha}_1 - \hat{\alpha}_2 \pm \sqrt{2\hat{\tau}^2 + \hat{\sigma}_1^2 + \hat{\sigma}_2^2}$$

4.3 Roy's LOAs

The limits of agreement computed by Roy's method are derived from the variance covariance matrix for overall variability. This matrix is the sum of the between subject VC matrix and the within-subject VC matrix.

The standard deviation of the differences of methods x and y is computed using values from the overall VC matrix.

$$\text{var}(x - y) = \text{var}(x) + \text{var}(y) - 2\text{cov}(x, y)$$

The respective estimates computed by both methods are tabulated as follows. Evidently there is close correspondence between both sets of estimates.

? formulates an LME model, both in the absence and the presence of an interaction term.? uses both to demonstrate the importance of using an interaction term. Failure to take the replication structure into account results in over-estimation of the limits of agreement. For the Carstensen estimates below, an interaction term was included when computed.

? uses the “Blood” data set, which featured in ?.

? describes the sampling method when discussing of a motivating example

Diabetes patients attending an outpatient clinic in Denmark have their HbA_{1c} levels routinely measured at every visit. Venous and Capillary blood samples were obtained from all patients appearing at the clinic over two days. Samples were measured on four consecutive days on each machines, hence there are five analysis days.

? notes that every machine was calibrated every day to the manufacturers guidelines. Measurements are classified by method, individual and replicate. In this case the replicates are clearly not exchangeable, neither within patients nor simultaneously for all patients.

5 Hamlett

Hamlett re-analyses the data of ? to generalize their model to cover other settings not covered by the Lam method.

In many cases, repeated observation are collected from each subject in sequence and/or longitudinally.

$$y_i = \alpha + \mu_i + \epsilon$$

5.1 T

he classical model is based on measurements y_{mi} by method $m = 1, 2$ on item $i = 1, 2, \dots$

$$y_{mi} = \alpha_m + \mu_i + e_{mi}$$

$$e_{mi} \sim \mathcal{N}(0, \sigma_m^2)$$

Even though the separate variances can not be identified, their sum can be estimated by the empirical variance of the differences.

Like wise the separate α can not be estimated, only their difference can be estimated as \bar{D}