

Contents

| | | |
|-----|---|----|
| 0.1 | Roy's methodology | 2 |
| 1 | Roy's LME approach | 3 |
| 2 | Roy's LME methodology for assessing agreement | 5 |
| 3 | Replicates | 6 |
| 4 | Roy's methodology with single measurements | 7 |
| 5 | Roy's examples | 7 |
| 6 | LME | 8 |
| 6.1 | Difference Variance further to Carstensen | 10 |
| 6.2 | Relevance of Roy's Methodology | 10 |
| 7 | LME models in method comparison studies | 12 |
| 8 | Roy's LME methodology for assessing agreement | 13 |

0.1 Roy's methodology

For the purposes of comparing two methods of measurement, Roy (2009a) presents a methodology utilizing linear mixed effects model. This methodology provides for the formal testing of inter-method bias, between-subject variability and within-subject variability of two methods. The formulation contains a Kronecker product covariance structure in a doubly multivariate setup. By doubly multivariate set up, Roy means that the information on each patient or item is multivariate at two levels, the number of methods and number of replicated measurements. Further to Lam et al. (1999), it is assumed that the replicates are linked over time. However it is easy to modify to the unlinked case.

Roy (2009a) sets out three criteria for two methods to be considered in agreement. Firstly that there be no significant bias. Second that there is no difference in the between-subject variabilities, and lastly that there is no significant difference in the within-subject variabilities. Roy further proposes examination of the the overall variability by considering the second and third criteria be examined jointly. Should both the second and third criteria be fulfilled, then the overall variabilities of both methods would be equal.

A formal test for inter-method bias can be implemented by examining the fixed effects of the model. This is common to well known classical linear model methodologies. The null hypotheses, that both methods have the same mean, which is tested against the alternative hypothesis, that both methods have different means. The inter-method bias and necessary t -value and p -value are presented in computer output. A decision on whether the first of Roy's criteria is fulfilled can be based on these values.

Importantly Roy (2009a) further proposes a series of three tests on the variance components of an LME model, which allow decisions on the second and third of Roy's criteria. For these tests, four candidate LME models are constructed. The differences in the models are specifically in how the D and Λ matrices are constructed, using either an unstructured form or a compound symmetry form. To illustrate these differences,

consider a generic matrix A ,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

A symmetric matrix allows the diagonal terms a_{11} and a_{22} to differ. The compound symmetry structure requires that both of these terms be equal, i.e $a_{11} = a_{22}$.

The first model acts as an alternative hypothesis to be compared against each of three other models, acting as null hypothesis models, successively. The models are compared using the likelihood ratio test. Likelihood ratio tests are a class of tests based on the comparison of the values of the likelihood functions of two candidate models. LRTs can be used to test hypotheses about covariance parameters or fixed effects parameters in the context of LMEs. The test statistic for the likelihood ratio test is the difference of the log-likelihood functions, multiplied by -2 . The probability distribution of the test statistic is approximated by the χ^2 distribution with $(\nu_1 - \nu_2)$ degrees of freedom, where ν_1 and ν_2 are the degrees of freedom of models 1 and 2 respectively. Each of these three test shall be examined in more detail shortly.

1 Roy's LME approach

The methodology uses a linear mixed effects regression fit using compound symmetry (CS) correlation structure on \mathbf{V} .

$$\Lambda = \frac{\max_{H_0} L}{\max_{H_1} L}$$

Roy (2009b) considers the problem of assessing the agreement between two methods with replicate observations in a doubly multivariate set-up using linear mixed effects models.

Roy (2009b) uses examples from Bland and Altman (1986) to be able to compare both types of analysis.

For the the RV-IC comparison, \hat{D} is given by

$$\hat{D} = \begin{bmatrix} 1.6323 & 1.1427 \\ 1.1427 & 1.4498 \end{bmatrix} \quad (1)$$

The estimate for the within-subject variance covariance matrix is given by

$$\hat{\Sigma} = \begin{bmatrix} 0.1072 & 0.0372 \\ 0.0372 & 0.1379 \end{bmatrix} \quad (2)$$

The estimated overall variance covariance matrix for the the 'RV vs IC' comparison is given by

$$Block\Omega_i = \begin{bmatrix} 1.7396 & 1.1799 \\ 1.1799 & 1.5877 \end{bmatrix}. \quad (3)$$

The power of the likelihood ratio test may depends on specific sample size and the specific number of replications, and the author proposes simulation studies to examine this further.

2 Roy's LME methodology for assessing agreement

Barnhart et al. (2007) describes the sources of disagreement as differing population means, different between-subject variances, different within-subject variances between two methods and poor correlation between measurements of two methods.

Roy (2009b)proposes the use of LME models to perform a test on two methods of agreement to determine whether they can be used interchangeably.

Bivariate correlation coefficients have been shown to be of limited use in method comparison studies (Bland and Altman, 1986). However, recently correlation analysis

has been developed to cope with repeated measurements, enhancing their potential usefulness. Roy incorporates the use of correlation into his methodology.

Roy (2009b) considers the problem of assessing the agreement between two methods with replicate observations in a doubly multivariate set-up using linear mixed effects models.

Roy (2009b) uses examples from Bland and Altman (1986) to be able to compare both types of analysis.

Roy (2009b) proposes a LME based approach with Kronecker product covariance structure with doubly multivariate setup to assess the agreement between two methods. This method is designed such that the data may be unbalanced and with unequal numbers of replications for each subject.

Roy (2009b) considers four independent hypothesis tests.

- Testing of hypotheses of differences between the means of two methods
- Testing of hypotheses in between subject variabilities in two methods,
- Testing of hypotheses of differences in within-subject variability of the two methods,
- Testing of hypotheses in differences in overall variability of the two methods.

3 Replicates

Measurements taken in quick succession by the same observer using the same instrument on the same subject can be considered true replicates. ? notes that some measurements may not be ‘true’ replicates.

Roy’s methodology assumes the use of ‘true replicates’. However data may not be collected in this way. In such cases, the correlation matrix on the replicates may require a different structure, such as the autoregressive order one $AR(1)$ structure. However determining MLEs with such a structure would be computational intense, if possible at all.

4 Roy's methodology with single measurements

5 Roy's examples

Roy provides three case studies, using data sets well known in method comparison studies, to demonstrate how the methodology should be used.

The first case study is the Systolic blood pressure data, taken from Bland and Altman (1999).

To complete the study, the relevant values are provided for the $RvsS$ comparison also.

The second data set, a comparison of two peak expiratory flow rate measurements, is referenced by Bland and Altman (1986).

The last case study is also based on a data set from Bland and Altman (1999). It contains the measurements of left ventricular cardiac eject fraction, measured by impedance cartography and radionuclide ventriculography, on twelve patients. The number of replicated differs for each patient.

The bias is shown to be 0.7040, with a p-value of 0.0204. The MLEa of the between-method and within-method variance-covariance matrices of methods RV and IC are given by

$$\hat{D} = \begin{pmatrix} 1.6323 & 1.1427 \\ 1.1427 & 1.4498 \end{pmatrix}, \quad (4)$$

$$\hat{\Sigma} = \begin{pmatrix} 1.6323 & 1.1427 \\ 1.1427 & 1.4498 \end{pmatrix}. \quad (5)$$

Roy (2009a) notes that these are the same estimate for variance as given by Bland and Altman (1999).

The repeatability coefficients are determined to be 0.9080 for the RV method and 1.0293 for the IC method.

From the estimated $\mathbf{\Omega}_i$ correlation matrix, the overall correlation coefficient is 0.7100. The overall correlation coefficients between two methods RV and IC are 0.9384 and 0.9131 respectively.

Roy (2009a) concludes that is appropriate to switch between the two methods if needed.

?

Roy (2009a) recommends to not switch between the two method.

6 LME

Fitting model according to Roy

Linear mixed-effects model fit by REML

Data: BA99

| | AIC | BIC | logLik |
|--|----------|----------|-----------|
| | 4319.707 | 4336.629 | -2155.853 |

Random effects:

Formula: ~1 | subj

(Intercept) Residual

StdDev: 29.39085 12.44454

Fixed effects: ob.js ~ method

| | Value | Std.Error | DF | t-value | p-value |
|-------------|-----------|-----------|-----|----------|---------|
| (Intercept) | 127.40784 | 3.281757 | 424 | 38.82306 | 0 |
| methodS | 15.61961 | 1.102107 | 424 | 14.17250 | 0 |

Correlation:

(Intr)

methodS -0.168

Standardized Within-Group Residuals:

| | Min | Q1 | Med | Q3 | Max |
|--|-------------|-------------|-------------|------------|------------|
| | -3.61292639 | -0.42538402 | -0.02467651 | 0.40166235 | 4.84280044 |

Number of Observations: 510 Number of Groups: 85

6.1 Difference Variance further to Carstensen

Carstensen et al. (2008) states a model where the variation between items for method m is captured by τ_m (our notation d_m^2) and the within-item variation by σ_m .

The formulation of this model is general and refers to comparison of any number of methods however, if only two methods are compared, separate values of τ_1^2 and τ_2^2 cannot be estimated, only their average value τ , so in the case of only two methods we are forced to assume that $\tau_1 = \tau_2 = \tau$ (Carstensen et al., 2008).

Another important point is that there is no covariance terms, so further to Carstensen et al. (2008) the variance covariance matrices for between-item and within-item variability are respectively.

$$\mathbf{D} = \begin{pmatrix} d_2^1 & 0 \\ 0 & d_2^2 \end{pmatrix}$$

and $\mathbf{\Sigma}$ is constructed as follows:

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_2^1 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

Under this model the limits of agreement should be computed based on the standard deviation of the difference between a pair of measurements by the two methods on a new individual, j , say:

$$\text{var}(y_{1j} - y_{2j}) = 2d^2 + \sigma_1^2 + \sigma_2^2$$

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

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

6.2 Relevance of Roy's Methodology

The relevance of Roy's methodology is that estimates for the between-item variances for both methods \hat{d}_m^2 are computed. Also the VC matrices are constructed with covariance terms and, so the difference variance must be formulated accordingly.

$$\hat{\alpha}_1 - \hat{\alpha}_2 \pm \sqrt{\hat{d}_1^2 + \hat{d}_1^2 + \hat{\sigma}_1^2 + \hat{\sigma}_2^2 - 2\hat{d}_{12} - 2\hat{\sigma}_1 2}$$

7 LME models in method comparison studies

Linear mixed effects (LME) models can facilitate greater understanding of the potential causes of bias and differences in precision between two sets of measurement. Lai and Shiao (2005) views the uses of linear mixed effects models as an expansion on the Bland-Altman methodology, rather than as a replacement. ? remarks that modern statistical computation, such as that used for LME models, greatly improve the efficiency of calculation compared to previous ‘by-hand’ methods. In this chapter various LME approaches to method comparison studies shall be examined.

8 Roy’s LME methodology for assessing agreement

Roy (2009b) proposes the use of LME models to perform a test on two methods of agreement to determine whether they can be used interchangeably.

Bivariate correlation coefficients have been shown to be of limited use in method comparison studies (Bland and Altman, 1986). However, recently correlation analysis has been developed to cope with repeated measurements, enhancing their potential usefulness. Roy incorporates the use of correlation into his methodology.

Roy’s method considers two methods to be in agreement if three conditions are met.

- no significant bias, i.e. the difference between the two mean readings is not ”statistically significant”,
- high overall correlation coefficient,
- the agreement between the two methods by testing their repeatability coefficients.

The methodology uses a linear mixed effects regression fit using compound symmetry (CS) correlation structure on \mathbf{V} .

$$\Lambda = \frac{\max_{H_0} L}{\max_{H_1} L}$$

Roy (2009b) considers the problem of assessing the agreement between two methods with replicate observations in a doubly multivariate set-up using linear mixed effects models.

Roy (2009b) uses examples from Bland and Altman (1986) to be able to compare both types of analysis.

Roy (2009b) proposes a LME based approach with Kronecker product covariance structure with doubly multivariate setup to assess the agreement between two methods. This method is designed such that the data may be unbalanced and with unequal numbers of replications for each subject.

The maximum likelihood estimate of the between-subject variance covariance matrix of two methods is given as D . The estimate for the within-subject variance covariance matrix is $\hat{\Sigma}$. The estimated overall variance covariance matrix 'Block Ω_i ' is the addition of \hat{D} and $\hat{\Sigma}$.

$$\text{Block } \Omega_i = \hat{D} + \hat{\Sigma} \quad (6)$$

For the the RV-IC comparison, \hat{D} is given by

$$\hat{D} = \begin{bmatrix} 1.6323 & 1.1427 \\ 1.1427 & 1.4498 \end{bmatrix} \quad (7)$$

The estimate for the within-subject variance covariance matrix is given by

$$\hat{\Sigma} = \begin{bmatrix} 0.1072 & 0.0372 \\ 0.0372 & 0.1379 \end{bmatrix} \quad (8)$$

The estimated overall variance covariance matrix for the the 'RV vs IC' comparison is given by

$$\text{Block}\Omega_i = \begin{bmatrix} 1.7396 & 1.1799 \\ 1.1799 & 1.5877 \end{bmatrix}. \quad (9)$$

The power of the likelihood ratio test may depends on specific sample size and the specific number of replications, and the author proposes simulation studies to examine this further.

References

- Barnhart, H., M. Haber, and L. Lin (2007). An overview of assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics* 17, 529–569.
- Bland, J. and D. Altman (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet* i, 307–310.
- Bland, J. and D. Altman (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 8(2), 135–160.
- 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).
- Lai, D. and S.-Y. P. K. Shiao (2005). Comparing two clinical measurements: a linear mixed model approach. *Journal of Applied Statistics* 32(8), 855–860.
- Lam, M., K. Webb, and D. O’Donnell (1999). Correlation between two variables in repeated measurements. *American Statistical Association, Proceedings of the Biometric Session*, 213–218.
- Roy, A. (2009a). An application of linear mixed effects model to assess the agreement between two methods with replicated observations. *Journal of Biopharmaceutical Statistics* 19, 150–173.
- Roy, A. (2009b). An application of the linear mixed effects model to ass the agreement between two methods with replicated observations. *Journal of Biopharmaceutical Statistics* 19, 150–173.