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.

- Computing limits of agreement features prominently in many method comparison studies, further to BA86,BA99. BA99 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.

Limits of Agreement

- bxc2008 computes the limits of agreement to the case with replicate measurements by using LME models.
- Roy 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).

Limits of Agreement

- 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.

Roy's method

- Roy proposes a novel method using the LME model with Kronecker product covariance structure in a doubly multivariate set-up to assess the agreement between a new method and an established method with unbalanced data and with unequal replications for different subjects (see Roy).
- Using Roy's method, four candidate models are constructed, each differing by constraints applied to the variance covariance matrices. In addition to computing the inter-method bias, three significance tests are carried out on the respective formulations to make a judgement on whether or not two methods are in agreement.

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

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

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.

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

Carstensen's LOAs

- The respective estimates computed by both methods are tabulated as follows. Evidently there is close correspondence between both sets of estimates.
- bxc2008 formulates an LME model, both in the absence and the presence of an interaction term.bxc 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.

Roy2006 uses the “Blood" data set, which featured in BA99.

Likelihood Ratio Tests

- The relationship between the respective models presented by roy is known as “nesting". A model A to be nested in the reference model, model B, if Model A is a special case of Model B, or with some specific constraint applied.
- A general method for comparing models with a nesting relationship is the **likelihood ratio test (LRTs)**.
- LRTs are a family of tests used to compare the value of likelihood functions for two models, whose respective formulations define a hypothesis to be tested (i.e. the nested and reference model).
- The significance of the likelihood ratio test can be found by comparing the likelihood ratio to the χ^2 distribution, with the appropriate degrees of freedom.

- When testing hypotheses around covariance parameters in an LME model, REML estimation for both models is recommended by West et al. REML estimation can be shown to reduce the bias inherent in ML estimates of covariance parameters west.
- Conversely, pb advises that testing hypotheses on fixed-effect parameters should be based on ML estimation, and that using REML would not be appropriate in this context.

Variance Covariance Matrices

- Under Roy's model, random effects are defined using a bivariate normal distribution. Consequently, the variance-covariance structures can be described using 2×2 matrices.
- A discussion of the various structures a variance-covariance matrix can be specified under is required before progressing.
- The following structures are relevant:
  1. the identity structure,
  2. the compound symmetric structure
  3. the symmetric structure.

Variance Covariance Matrices

- The **identity** structure is simply an abstraction of the identity matrix.
- The **compound symmetric** structure and **symmetric** structure can be described with reference to the following matrix (here in the context of the overall covariance Block-$\boldsymbol{\Omega}_i$, but equally applicable to the component variabilities $\boldsymbol{G}$ and $\boldsymbol{\Sigma}$);

$$\begin{pmatrix} \omega_1^2 & \omega_{12} \\ \omega_{12} & \omega_2^2 \end{pmatrix}$$

Symmetric structure requires the equality of all the diagonal terms, hence $\omega_1^2 = \omega_2^2$. Conversely compound symmetry make no such constraint on the diagonal elements. Under the identity structure, $\omega_{12} = 0$. A comparison of a model fitted using symmetric structure with that of a model fitted using the compound symmetric structure is equivalent to a test of the equality of variance.
In the presented example, it is shown that Roy's LOAs are lower than those of (BXC-model), when covariance between methods is present.

Remarks on the Multivariate Normal Distribution

- Diligence is required when considering the models. Carstensen specifies his models in terms of the univariate normal distribution.
- Roy's model is specified using the bivariate normal distribution.
- This gives rises to a key difference between the two model, in that a bivariate model accounts for covariance between the variables of interest.

- Roys uses and LME model approach to provide a set of formal tests for method comparison studies.
- Four candidates models are fitted to the data.
- These models are similar to one another, but for the imposition of equality constraints.
- These tests are the pairwise comparison of candidate models, one formulated without constraints, the other with a constraint. Roy's model uses fixed effects $\beta_0 + \beta_1$ and $\beta_0 + \beta_1$ to specify the mean of all observationsby
  methods 1 and 2 respectively.

This model includes a method by item interaction term.
Carstensen presents two models. One for the case where the replicates, and a second for when they are linked.
Carstensen's model does not take into account either between-item or within-item covariance between methods.
In the presented example, it is shown that Roy's LoAs are lower than those of Carstensen.

$$\begin{pmatrix} \omega_2^1 & 0 \\ 0 & \omega_2^2 \end{pmatrix} = \begin{pmatrix} \tau^2 & 0 \\ 0 & \tau^2 \end{pmatrix} + \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

Remarks on the Multivariate Normal Distribution

The multivariate normal distribution of a k-dimensional random vector $X = [X_1, X_2, \ldots, X_k]$ can be written in the following notation:

$$X \ \sim \ \mathcal{N}(\mu, \, \Sigma),$$

or to make it explicitly known that X is k-dimensional,

$$X \ \sim \ \mathcal{N}_k(\mu, \, \Sigma).$$

with k-dimensional mean vector

$$\mu = [\mathrm{E}[X_1], \mathrm{E}[X_2], \ldots, \mathrm{E}[X_k]]$$

and $k \times k$ covariance matrix

$$\Sigma = [\mathrm{Cov}[X_i, X_j]], \ i = 1, 2, \ldots, k; \ j = 1, 2, \ldots, k$$

1. Univariate Normal Distribution

$$X \sim \mathcal{N}(\mu, \sigma^2),$$

2. Bivariate Normal Distribution

(a)

$$X \sim \mathcal{N}_2(\mu, \Sigma),$$

(b)

$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}.$$

Note 1: Coefficient of Repeatability

- The coefficient of repeatability is a measure of how well a measurement method agrees with itself over replicate measurements BA99.
- Once the within-item variability is known, the computation of the coefficients of repeatability for both methods is straightforward.

Carstensen model in the single measurement case

- BXC2004 presents a model to describe the relationship between a value of measurement and its real value.
- The non-replicate case is considered first, as it is the context of the Bland-Altman plots.
- This model assumes that inter-method bias is the only difference between the two methods.

Carstensen model in the single measurement case

$$y_{mi} = \alpha_m + \mu_i + e_{mi} \qquad e_{mi} \sim \mathcal{N}(0, \sigma^2_m) \tag{1}$$

The differences are expressed as $d_i = y_{1i} - y_{2i}$.
For the replicate case, an interaction term c is added to the model, with an associated variance component.

Note 3: Model terms

It is important to note the following characteristics of this model.

- Let the number of replicate measurements on each item i for both methods be n_i, hence $2 \times n_i$ responses. However, it is assumed that there may be a different number of replicates made for different items. Let the maximum number of replicates be p. An item will have up to $2p$ measurements, i.e. $\max(n_i) = 2p$.
- Later on $\boldsymbol{X}_i$ will be reduced to a 2×1 matrix, to allow estimation of terms. This is due to a shortage of rank. The fixed effects vector can be modified accordingly.
- $\boldsymbol{Z}_i$ is the $2n_i \times 2$ model matrix for the random effects for measurement methods on item i.
- $\boldsymbol{b}_i$ is the 2×1 vector of random-effect coefficients on item i, one for each method.

Note 3: Model terms (Continued)

- ϵ is the $2n_i \times 1$ vector of residuals for measurements on item i.
- $\boldsymbol{G}$ is the 2×2 covariance matrix for the random effects.
- $\boldsymbol{R}_i$ is the $2n_i \times 2n_i$ covariance matrix for the residuals on item i.
- The expected value is given as $\mathrm{E}(\boldsymbol{y}_i) = \boldsymbol{X}_i\beta$. hamlett
- The variance of the response vector is given by $\mathrm{Var}(\boldsymbol{y}_i) = \boldsymbol{Z}_i\boldsymbol{G}\boldsymbol{Z}_i' + \boldsymbol{R}_i$ ...hamlett.