Common Slope Tests for Bivariate Errors-in-Variables Models

DAVID I. WARTON
Macquarie University
Australia

NEVILLE C. WEBER
University of Sydney
Australia

Summary

Likelihood ratio tests are derived for bivariate normal structural relationships in the presence of group structure. These tests may also be applied to less restrictive models where only errors are assumed to be normally distributed. Tests for a common slope amongst those from several datasets are derived for three different cases – when the assumed ratio of error variances is the same across datasets and either known or unknown, and when the standardised major axis model is used. Estimation of the slope in the case where the ratio of error variances is unknown could be considered as a maximum likelihood grouping method. The derivations are accompanied by some small sample simulations, and the tests are applied to data arising from work on seed allometry.

Key words: Common principal component analysis; Group structure; Grouping methods; Standardised or reduced major axis; Structural and functional relationships.

1. Introduction

Applications for functional and structural relationship models (KENDALL and STUART, 1973, Chapter 29) quite commonly arise in the applied sciences (e.g. McARDLE, 1988; WEBSTER, 1997). Despite this, few inferential procedures have been developed for such models. The development of one-sample tests for such models is usually attributed to researchers as far back as CREASY (1956) and PITMAN (1939), however tests for functional and structural relationships in the presence of group structure have a much shorter history (CLARKE, 1980; BANSAL, 1990; ISO-GAWA, 1992; DOUGHERTY, 1997). Tests for commonality are of particular interest as all research inevitably requires some consideration of the extent of its generality, and this often takes a statistical form. This paper develops likelihood ratio tests for a common slope amongst the bivariate structural relationships of several datasets, a procedure for which there was previously no guidance for practitioners.
The model of interest considers data in \( g \) groups where the \( i \)th group consists of \( n^{(i)} \) independent bivariate normal observations \((x_j^{(i)}, y_j^{(i)})\), \( i = 1, \ldots, g; \) \( j = 1, \ldots, n^{(i)} \). It is postulated that the observations are representations of unobserved variables \((X_j^{(i)}, Y_j^{(i)})\) which are linearly related for each \((i, j)\):

\[
Y_j^{(i)} = \alpha^{(i)} + \beta^{(i)}X_j^{(i)}
\]

where independently for each \((i, j)\)

\[
X_j^{(i)} \sim N(\mu^{(i)}, \sigma^{(i)^2}),
\]

\[
x_j^{(i)} \sim N(x_j^{(i)}, \sigma_{x}^{(i)^2})\] and \( y_j^{(i)} \sim N(Y_j^{(i)}, \sigma_{y}^{(i)^2}).\)

This model will be referred to as the bivariate structural relationship, and is the subject of Sections 2-4. In Section 5 the \( X_j^{(i)} \) are treated as fixed instead of (1), a model referred to as the bivariate functional relationship.

Moran (1971) and others have noted that without some additional knowledge, the \( \beta^{(i)} \) are unidentifiable. Given that \((x_j^{(i)}, y_j^{(i)})\) is bivariate normal, the distribution can be described using 5\( g \) parameters. However, under the reparameterisation described above for the structural relationship model, there are 6\( g \) parameters – \((\mu^{(i)}, \alpha^{(i)}, \beta^{(i)}, \sigma^{(i)}_x, \sigma^{(i)}_y, \sigma^{(i)}_{xy})\) for each \( i = 1, \ldots, g \). It is common in many areas to treat the error variance ratio \( V^{(i)} = \sigma^{(i)}_y^2 / \sigma^{(i)}_x^2 \) as a known quantity, hence reducing the number of model parameters to 5\( g \). This essentially reduces the two sources of error variation from the straight line to just one, as choosing \( V^{(i)} \) defines a direction in which departures from the line are measured during estimation of the line. For example, the major axis model \((V^{(i)} = 1)\) is commonly used, which measures departures perpendicular to the fitted line. The major axis slope is equivalent to the slope of the first principal component vector from the variance/covariance matrix. Another common alternative is to use the standardised major axis model \((V^{(i)} = \beta^{(i)} \), where error variances \((\sigma^{(i)}_x^2, \sigma^{(i)}_y^2)\) are assumed proportional to variances of the observed variables \( x \) and \( y \), and the gradient of the fitted line is reflected around the vertical to define the direction in which departures from the line are measured. Some considerations with the use of these approaches are reviewed by Sprent and Dolby (1980), and some alternative methods of estimation can be found in Fuller (1987).

The test of interest in this paper is \( H_{\text{com}} : \beta^{(i)} = \beta \), for some unknown \( \beta \), \( i = 1, \ldots, g \), against \( H_{\text{dif}} : \beta^{(i)} \neq \beta \) for at least one of \( i = 1, \ldots, g \). Notice that if each \( \sigma^{(i)}_x = \sigma^{(i)}_y \), then a test of \( H_{\text{com}} \) against \( H_{\text{dif}} \) is equivalent to the problem considered by Flury (1984) for a common set of principal components from \( g \) variance/covariance matrices. If instead all \( \sigma^{(i)}_x = 0 \), or if all \( \sigma^{(i)}_y = 0 \), the problem reduces to a test for common linear regression slopes (Cochran, 1957).

Clarke (1980) developed an approximate \( t \)-test for \( H_{\text{com}} \) against \( H_{\text{dif}} \) under the standardised major axis model, although this test only considered \( g = 2 \). Bansal (1990) derived the likelihood ratio statistic and its approximate distribution for \( H_{\text{com}} \) against \( H_{\text{dif}} \) under the major axis functional relationship model with \( g = 2 \).
A likelihood ratio test for the major axis functional relationship model was later developed (DOUGHERTY, 1997), assuming an arbitrary number of groups, although estimation would require replicated observations at each \((X^{(i)}_j, Y^{(i)}_j)\). Apart from work in the principal component literature, there are no techniques available in the structural relationship literature for testing \(H_{com} \) against \(H_{dif} \) for arbitrary \(g\).

We derive likelihood ratio tests of \(H_{com} \) against \(H_{dif} \) for the model outlined above, but further assuming:

\[
\begin{align*}
(1.1) \text{each } V^{(i)} &= V, \text{ a known constant;} \\
(1.2) \text{each } V^{(i)} &= V, \text{ an unknown constant;} \\
(1.3) \text{each } V^{(i)} &= \beta^{(i)^2}.
\end{align*}
\]

The test statistic given below at (7) for the test of common slope applied to the model under assumption (1.1) is a straightforward extension of FLURY (1984), enhanced by the inclusion of a Bartlett correction to improve the chi-squared approximation to the null distribution of the statistic. This form suggests statistics for the other two cases and extensions to the functional relationship model, where the \(X^{(i)}_j \) are treated as fixed.

For the bivariate normal models (1.1–1.3), generic maximisation routines could be used, starting with the 5\(g\) equations expressing moment estimates in terms of the model parameters. We describe the method in greater detail, however, to elucidate its geometric interpretation, to report simple computational methods, and identify an appropriate Bartlett correction for the test statistic. Without obtaining a closed form for the test statistic under assumption (1.1) the extension to the functional relationship case, which is not a likelihood ratio statistic, would not be apparent.

2. A Common Slopes Test with \(V\) Known

Assume that the error variance ratios of all groups are known to be \(V\), as in (1.1). Transformation to an error variance ratio of 1 would allow testing as in FLURY (1984), although see KRZANOWSKI (1984) for a simpler slope estimate.

For a bivariate (functional or) structural relationship model with slope \(\beta\), observations are modelled to err from the line at a slope of \(-V/\beta\) (e.g. SPRENT, 1969). Lines with the latter slope will be referred to as residual axes. The transformation

\[
\begin{pmatrix}
x \\
y
\end{pmatrix} = (V + \beta^2)^{-\frac{1}{2}} \begin{pmatrix}
1 & -\beta \\
\beta & V
\end{pmatrix} \begin{pmatrix}
u \\
v
\end{pmatrix}
\]

(2)
is useful for estimation of the model and ensuing inference. For the model without group structure, for example, a suitable estimate for \(\beta\) (such as the maximum likelihood estimate) is such that \(u\) and \(v\) are independent variables, so that position along the line is independent of distance from the line along residual axes. To perform a one-sample test of the slope, one could test for independence of variables transformed as in (2), with \(\beta\) taking its null value (See Appendix 1 for de-
tails). For the present model, an equivalent hypothesis to $H_{\text{com}}$ is that there is a transformation of the form of (2) such that for some $\beta$, each variance/covariance matrix $\Sigma^{(i)} i = 1, \ldots, g$ satisfies

$$\Sigma^{(i)} = \frac{1}{V + \beta^2} \begin{pmatrix} 1 & -\beta \\ \beta & V \end{pmatrix} \begin{pmatrix} \lambda_1^{(i)} & 0 \\ 0 & \lambda_2^{(i)} \end{pmatrix} \begin{pmatrix} 1 & \beta \\ -\beta & V \end{pmatrix}. \quad (3)$$

The $\lambda_j^{(i)}$ are defined as the variances of transformed variables $u$ and $v$. To find maximum likelihood estimates sequentially for the $\lambda_j^{(i)}$ and then $\beta$, the log likelihood expression for $g$ groups of bivariate normal data is written as

$$\log l_{\text{com}}(x, y) = K - \sum_{i=1}^g n^{(i)} \log |\Sigma^{(i)}| - \sum_{i=1}^g \frac{n^{(i)}}{2} \operatorname{tr}(\Sigma^{(i)-1} S^{(i)}),$$

for some constant $K$, where $S^{(i)}$ is the sample variance/covariance matrix for the $i$th group, $|A|$ and $\operatorname{tr}(A)$ are respectively the determinant and trace of $A$. Using (3), $|\Sigma^{(i)}| = \lambda_1^{(i)} \lambda_2^{(i)}$ and

$$\operatorname{tr}(\Sigma^{(i)-1} S^{(i)}) = \operatorname{tr} \left\{ \frac{1}{V + \beta^2} \begin{pmatrix} V & -\beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/\lambda_2^{(i)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/\lambda_2^{(i)} \end{pmatrix} \begin{pmatrix} V & -\beta \\ -\beta & 1 \end{pmatrix} \right\} \begin{pmatrix} S^{(i)}_{xx} & S^{(i)}_{xy} \\ S^{(i)}_{xy} & S^{(i)}_{yy} \end{pmatrix} \begin{pmatrix} V & -\beta \\ -\beta & 1 \end{pmatrix},$$

so it can be seen that estimates for the $\lambda_j^{(i)}$ satisfy

$$\begin{pmatrix} \hat{\lambda}_1^{(i)} \\ C_V^{(i)}(\beta) \hat{\lambda}_2^{(i)} \end{pmatrix} = \frac{1}{V + \beta^2} \begin{pmatrix} V & \beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} S^{(i)}_{xx} & S^{(i)}_{xy} \\ S^{(i)}_{xy} & S^{(i)}_{yy} \end{pmatrix} \begin{pmatrix} V & -\beta \\ \beta & 1 \end{pmatrix},$$

where $C_V^{(i)}(\beta)$ is a term defined in the above expression. This result could also be seen by replacing the variance/covariance matrix in (3) by its sample estimates.
Notice also that for the $i$th group, the root of $C_V^{(i)}(\beta)$ maximising $\lambda_1^{(i)}$ is the structural relationship slope estimator for the $i$th group, $\hat{\beta}_i^{(i)}$.

The common slope estimate can be found by maximising the simplified likelihood expression

$$\log \hat{l}_{\text{com}}(x, y) = C - \sum_{i=1}^{g} \frac{n^{(i)}}{2} \log \hat{\lambda}_1^{(i)} \hat{\lambda}_2^{(i)},$$

where $C$ does not depend on $\beta$. In order to estimate $\beta$, note

$$-\frac{1}{V} \frac{\partial \hat{\lambda}_1^{(i)}}{\partial \beta} = \frac{\partial \hat{\lambda}_2^{(i)}}{\partial \beta} = \frac{2}{(V + \beta^2)^2} \left( \beta^2 S_{xy}^{(i)} - \beta S_{xy}^{(i)} + \beta V S_{xx}^{(i)} - V S_{xy}^{(i)} \right)$$

$$= \frac{-2}{(V + \beta^2)} C_V^{(i)}(\beta),$$

hence

$$\frac{\partial \log \hat{l}_{\text{com}}(x, y)}{\partial \beta} = \frac{1}{(V + \beta^2)} \sum_{i=1}^{g} n^{(i)} \left( \frac{1}{\hat{\lambda}_2^{(i)}} - \frac{V}{\hat{\lambda}_1^{(i)}} \right) C_V^{(i)}(\beta).$$

The common slope estimator found by setting (5) to zero can be considered as the structural relationship slope estimator from a pooled variance/covariance matrix with a weight for each group of $n^{(i)}(1/\hat{\lambda}_2^{(i)} - V/\hat{\lambda}_1^{(i)})$, a generalisation of Flury (1984).

For each common slope estimate from (5) of the form $\hat{\beta}$, there is another solution $-V/\hat{\beta}$. Of such a pair of solutions, the common structural relationship slope estimator has the same sign as the weighted sum of $S_{xy}^{(i)}$ terms, the other slope estimate is its residual axis. The solutions for the common slope can not be written explicitly, but can be readily found by iteration, with any group slope estimate providing an adequate initial estimate. One suitable algorithm involves using the current estimate of $\hat{\beta}$ to find the weights $n^{(i)}(1/\hat{\lambda}_2^{(i)} - V/\hat{\lambda}_1^{(i)})$, and hence compute a new slope estimate based on the current pooled variance/covariance matrix. Slope estimation is complicated by the possibility of multiple solutions within the positive or negative domain. Simulations have suggested that as in Dougherty (1997), multiple local maxima typically occur where observed slopes vary markedly across groups relative to error variation from the lines.

The alternative hypothesis $H_{\text{dif}}$ is simply a reparameterisation of the bivariate normal distribution, and so the likelihood under $H_{\text{dif}}$ is as usual for $g$ groups. A likelihood ratio test for $H_{\text{com}}$ against $H_{\text{dif}}$ can be constructed:

$$-2 \log \Lambda = -\sum_{i=1}^{g} n^{(i)} \log \left( \frac{|S^{(i)}|}{\hat{\lambda}_1^{(i)} \hat{\lambda}_2^{(i)}} \right).$$
In terms of the transformed variables, for each $i$, $\lambda_{1}^{(i)} = S_{uu}^{(i)}$, $\lambda_{2}^{(i)} = S_{vv}^{(i)}$, and $|S^{(i)}| = S_{uu}^{(i)}S_{vv}^{(i)} - S_{uv}^{(i)}$, so the likelihood expression can be written as

$$-2 \log \hat{\Lambda} = -\sum_{i=1}^{g} n^{(i)} \log \left(1 - (\hat{r}_{uv}^{(i)})^2\right).$$  \hspace{1cm} (6)

The test statistic is an increasing function of the $(\hat{r}_{uv}^{(i)})^2$, as is the one-sample test (Appendix 1). If $\beta$ were known, this test would be a multiple group generalisation of the one sample test. However, when $\beta$ is estimated from the sample $(\hat{u}, \hat{v})$ are not bivariate normal, and so only an asymptotic null distribution for (6) is known. Under $H_{\text{com}}$, (6) is asymptotically distributed as a chi-squared with $(g - 1)$ degrees of freedom. Under $H_{\text{dif}}$, $\hat{u}$ and $\hat{v}$ are correlated for some groups, and so large values of (6) argue against $H_{\text{com}}$.

The approximation of (6) to a chi-squared distribution under $H_{\text{com}}$ may be improved through the use of a Bartlett correction (BARTLETT, 1954). Use of Bartlett corrections is standard for multivariate normal likelihood tests – see, for example, MARDIA, KENT and BIBBY (1979). The correction factor in the case when $\beta$ is known is $(1 - 2.5/n^{(i)})$ for each of the $g$ groups (Appendix 1), and indeed one would expect this correction also to be useful in the case where $\beta$ is estimated. Hence in place of (6) $H_{\text{com}}$ can be tested against $H_{\text{dif}}$ using

$$-2 \log \hat{\Lambda} = -\sum_{i=1}^{g} (n^{(i)} - 2.5) \log \left(1 - (\hat{r}_{uv}^{(i)})^2\right).$$  \hspace{1cm} (7)

Like the one-sample test, this common slope test does not distinguish between the true slope and its residual axis. The test statistic is distributed as under $H_{\text{com}}$ if all structural relationship slopes are either $\beta$ or $-V/\beta$.

3. A Common Slopes Test with $V$ Unknown

The error variance ratio assumptions described in (1.2) are now considered, where each $V^{(i)} = V$, an unknown constant.

Recall that the structural relationship is not identifiable without an additional constraint to those stated in the original model. Under $H_{\text{com}}$, there are $(g - 1)$ additional constraints, which allow estimation of $V$ if it is unknown. The ensuing slope estimate is found by assuming only that the grouped data share a common structural relationship, the same assumption necessary for so-called “grouping methods” due to WALD (1940) and NAIR and SHRIVASTAVA (1942). Bivariate normality must also be assumed to derive the present method, and estimation is numerically more involved than for other grouping methods.

First maximising the likelihood with respect to each $\lambda_{j}^{(i)}$, then maximising with respect to $V$ and $\beta$, results up to (5) can be applied, giving
\[
\frac{\partial \log(\hat{\lambda}_1^{(i)} \hat{\lambda}_2^{(i)})}{\partial V} = \frac{\partial}{\partial V} \log (V^2 S_{xx}^{(i)} + 2V\beta S_{xy}^{(i)} + \beta^2 S_{yy}^{(i)}) - \frac{\partial}{\partial V} 2 \log (V + \beta^2) \\
= 2\beta (\beta^2 S_{xy}^{(i)} - \beta (S_{yy}^{(i)} - VS_{xx}^{(i)}) - VS_{xy}^{(i)}) \\
= \frac{-2\beta}{(V + \beta^2)\hat{\lambda}_1^{(i)}} C_V^{(i)}(\beta),
\]

so
\[
\frac{\partial \log \hat{L}_{com}(x, y)}{\partial V} = \frac{\beta}{(V + \beta^2)} \sum_{i=1}^{g} \frac{n^{(i)}}{\hat{\lambda}_1^{(i)}} C_V^{(i)}(\beta).
\] (8)

The maximum likelihood solutions for the common slope and error variance ratio can be found by setting to zero (5) and (8), then solving simultaneously. One solution has \( V = \infty \) and any \( \beta \), although other solutions are possible. Multiple stationary points for \( V \) are sometimes found, and maxima can occur at either endpoint. A saddle point on the likelihood surface as in SOLARI (1969), however, is not expected, as this is more typically encountered when a model is overparameterised. The parameters \( \beta \) and \( V \) can be estimated simultaneously by iteration. The cases where \( V \) approaches its endpoints should also be considered as possible maxima.

As for common slope estimation from (5), the solution to (8) can be thought of as coming from a pooled variance/covariance matrix, although with weights \( n^{(i)}/\hat{\lambda}_1^{(i)} \). Groups relatively poorly fitted by a common slope are of greater interest in estimating \( V \) as it is a measure of the error variances.

Under \( H_{diff} \), \( V \) can not be identified. However, for all possible values of \( V \), the observed variables for each group are bivariate normal with unconstrained parameters. Hence the maximised likelihood under \( H_{diff} \) remains as previously, and the corrected likelihood ratio statistic takes the minimum possible value of (7):
\[
-2 \log \hat{\Lambda} = \min_V \left\{ -\sum_{i=1}^{g} (n^{(i)} - 2.5) \log \left( 1 - \left( \hat{r}_{uv}^{(i)} \right)^2 \right) \right\}. \quad (9)
\]

Taking critical values from \( \chi^2_{(g-1)} \) would provide an “asymptotically conservative” test. The procedure is analogous to one-sample testing according to Moran (1956), as again \( V \) is chosen to provide the most conservative possible test statistic.

4. A Common Slopes Test for Standardised Major Axes

We now consider the standardised major axis model, as in (1.3); \( V^{(i)} = \beta^{(i)}^2 \), i.e. the ratio of error variances equals the ratio of variances of the unobserved variables \( X \) and \( Y \). Using this substitution in (2), the procedure is essentially as previously. First applying a change of scale, \( H_{com} \) specifies that for each group there
is a common transformation

\[(x) = (2 |\beta|)^{-\frac{1}{2}} \begin{pmatrix} 1 & -1 \\ \beta & \beta \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix}\]

to independent variables \(s\) and \(t\). A change of scale was suggested here to remain consistent with the transformation used in the analogous one-sample test by PITMAN (1939). Proceeding as in section 2, the \(\lambda_j^{(i)}\) are defined as variances of the transformed variables \((s, t)\) and can be estimated using

\[
\begin{pmatrix}
\hat{\lambda}_1^{(i)} \\
\hat{\lambda}_2^{(i)}
\end{pmatrix}
= \frac{1}{2|\beta|} \begin{pmatrix} \beta & 1 \\ -\beta & 1 \end{pmatrix}
\begin{pmatrix}
S_{xx}^{(i)} & S_{xy}^{(i)} \\
S_{xy}^{(i)} & S_{yy}^{(i)}
\end{pmatrix}
\begin{pmatrix} \beta & -\beta \end{pmatrix}.
\]

The likelihood function can then be expressed as in (4). Estimation of the slope does not proceed as for Section 2, as in this case \(V\) is also a function of \(\beta\). Instead, observing that

\[
\frac{\partial \hat{\lambda}_1^{(i)}}{\partial \beta} = \frac{\partial \hat{\lambda}_2^{(i)}}{\partial \beta} = \frac{\text{sign}(\beta)}{2} \left( S_{xx}^{(i)} - \frac{S_{xy}^{(i)}}{\beta^2} \right) = -\frac{1}{|\beta|} C_{\beta^2}^{(i)}(\beta),
\]

\[
\frac{\partial \log \hat{t}_{com}(x, y)}{\partial \beta} = \frac{1}{2 |\beta|} \sum_{i=1}^g n^{(i)} \left( \frac{1}{\hat{\lambda}_1^{(i)}} + \frac{1}{\hat{\lambda}_2^{(i)}} \right) C_{\beta^2}^{(i)}(\beta).
\]

Setting this derivative to zero provides the common slope estimate for the standardised major axis of the \(g\) groups. As in previous cases, the common slope estimate can be considered as the slope estimate obtained from a pooled variance/covariance matrix, although with weights \(n^{(i)}(1/\hat{\lambda}_1^{(i)} + 1/\hat{\lambda}_2^{(i)})\). As previously, the weighting is higher for groups with larger sample sizes or a greater proportion of variance explained by the common standardised major axis, as measured by the sum of reciprocals.

The likelihood is an even function of \(\beta\), so solutions arise in pairs of equal magnitude, and as previously can only be found in general by iteration.

The corrected likelihood ratio can be expressed as

\[
-2 \log \hat{\Lambda} = -\sum_{i=1}^g (n^{(i)} - 2.5) \log \left( 1 - \left( \hat{r}_{st}^{(i)} \right)^2 \right).
\]

The test statistic is scale independent, hence it is a particular case of (7) when \(V = \beta^2\). This implies that (10) provides a less conservative test than (9).

5. Common Slopes Tests for Functional Relationships

Now consider the \(X_j^{(i)}\) fixed. In this case the tests described in Sections 2–4 are applicable, although they are no longer likelihood ratio tests.
The reasoning for extending the test to the functional relationship case is essentially as in CReasy (1956) – following basic properties of the correlation coefficient, the distribution of \( r_{xy} \) when \( x \) and \( y \) are independent and normal is identical to the distribution of \( r_{xy} \) when \( x \) and \( y \) are independent but only one of \( x \) and \( y \) is normally distributed. Hence this test may be used in the absence of knowledge of the distributions of \( (X_j^{(i)}, Y_j^{(i)}) \), provided it is still assumed that errors in observing the \( (X_j^{(i)}, Y_j^{(i)}) \) are normally distributed. Simulations presented in Section 7 provide some confirmation of the claim that the null distributions of the test statistics (7) and (10) are largely unaffected by the distribution of the \( X_j^{(i)} \).

6. Worked Example

The use of these tests will be exemplified using data arising from seed allometry research. For 287 naturalised plant species of the Sydney region, a measure was obtained of mass of the seed coat and seed reserve, i.e. everything contained within the seed coat (Westoby et al. unpublished manuscript). Each plant species was classified according to presence of a morphological adaptation for seed dispersal (as one of unassisted, wind, vertebrate, ant and ballistic dispersal). As illustrated in Figure 1, the logarithms of masses were approximately bivariate normal in distribution, and data for all species treated as independent. It is of interest to test whether or not there is evidence for the slopes of allometric lines varying with dispersal mode.

Likelihood ratio tests are constructed for three common models of such data: the major axis model, where \( V \) is taken to be 1; the standardised major axis model; the structural relationship model where \( V \) is considered unknown. The asymptotic null distribution of the test statistic in the first two cases is a chi-square with four degrees of freedom, and this is conservatively taken to be the case where \( V \) is considered unknown. Simulations presented in the following section suggest that tail probabilities are quite accurate, hence the test statistics all suggest there is reasonable evidence that the slopes of each group are not identical (Table 2a). For each model the slope estimate for wind-dispersed species is considerably lower.

<table>
<thead>
<tr>
<th>Dispersal mode</th>
<th>( n^{(i)} )</th>
<th>var (coat)</th>
<th>var (res.)</th>
<th>cov (coat, res.)</th>
<th>( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unassisted</td>
<td>87</td>
<td>0.903</td>
<td>0.694</td>
<td>0.732</td>
<td>0.856</td>
</tr>
<tr>
<td>Wind</td>
<td>32</td>
<td>0.480</td>
<td>0.701</td>
<td>0.481</td>
<td>0.688</td>
</tr>
<tr>
<td>Vertebrate</td>
<td>54</td>
<td>0.964</td>
<td>0.780</td>
<td>0.506</td>
<td>0.340</td>
</tr>
<tr>
<td>Ant</td>
<td>96</td>
<td>0.403</td>
<td>0.332</td>
<td>0.272</td>
<td>0.554</td>
</tr>
<tr>
<td>Ballistic</td>
<td>18</td>
<td>0.590</td>
<td>0.343</td>
<td>0.429</td>
<td>0.907</td>
</tr>
</tbody>
</table>

Numbers are rounded to three significant figures.
than the others, and there is little evidence against $H_{\text{com}}$ when this group is excluded from analysis (Table 2b).

7. Small Sample Simulations

Simulations have been conducted to consider the small-sample properties of the test statistics (7) and (10), for models with $V$ known (Table 3). All cases reported consider bivariate data with variances one and squared correlation 0.5, and provide the proportion of times in 100 000 simulations that the test statistic exceeded critical points of 0.01 and 0.05 from a chi-squared distribution with $(g - 1)$ degrees of freedom.

Fig. 1. Seed coat mass versus seed reserve mass for plant species with dispersal characterised as (a) unassisted (b) wind (c) vertebrate (d) ant (e) ballistic. Fitted lines are standardised major axes estimated under $H_{\text{com}}$ (dotted line) and $H_{\text{dif}}$ (solid line)
For bivariate normal data (Table 3a), there is reasonable agreement with the chi-squared distribution even when \( \min(n^{(i)}) = 5 \) for \( g \leq 8 \), although for the major axis model the critical values were exceeded as infrequently as half the intended significance levels in samples of average size 10. In small samples the standardised major axis agreed about as closely with its tabulated null distribution as the test due to CLARKE (1980).

In additional simulations, data were generated where the \( X^{(i)} \) were proportional to a gamma distribution with mean 0.5, variance 1. This distribution is more strongly skewed than the exponential. The null tail distributions of the test statistics (Table 3b) appear quite similar to those obtained for bivariate normal data – even for small, unbalanced samples – suggesting that the tests outlined in sections 24 will be useful in the absence of distributional knowledge of \( X^{(i)} \).

Simulations for test statistics without a Bartlett correction confirmed that the chi-squared approximation was considerably improved by the use of a Bartlett correction.

8. Conclusions

Each of the tests derived here is a multiple group generalisation of the analogous one-sample test. It is anticipated that if these tests were generalised to problems of greater dimensionality, the test statistics would be functions of the appropriate multiple correlation coefficients used for one-sample testing as in JOLICOEUR (1984).

In cases where departures from fitted lines show strong non-normality, more computationally intensive methods of inference might be relevant (EFRON and TIBSHIRANI, 1993).
Results presented here have ramifications for testing for common principal component vectors. This literature has not yet investigated the possibility of Bartlett corrections nor of relaxing the multivariate normality assumption, areas that require future research.

Acknowledgements

Discussions with Mark Westoby and Brian McArdle at different stages during this work were of value. We also thank referees for improvements to the presentation.
Appendix 1 – One Sample Test for a Structural Relationship Slope

Consider one bivariate normal sample of size $n$, fitted by a structural relationship model with known error variance ratio $V$. There is interest in testing $H_b : \beta = b$ against $H_{alt} : \beta \neq b$.

By considering the transformation to $(u, v)$ which satisfies

$$
\begin{pmatrix}
x \\
y
\end{pmatrix} = \left( V + b^2 \right)^{-\frac{1}{2}} \begin{pmatrix} 1 & -b \\ b & V \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix},
$$

The variance/covariance matrix $\Sigma$ under $H_b$ is

$$
\Sigma = \frac{1}{V + b^2} \begin{pmatrix} 1 & -b \\ b & V \end{pmatrix} \lambda_1 \begin{pmatrix} 0 & 1 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} -b & V \end{pmatrix},
$$

for some constants $\lambda_1$ and $\lambda_2$.

The transformed variable $u$ represents distance along a line of slope $b$, and $\lambda_1$ represents the variance of samples along this line. Conversely $v$ represents distance along residual axes, and the variance of such residuals is $\lambda_2$. Under $H_b$

$$
\log \hat{\lambda}_b(x, y) = C - \frac{n}{2} \log \hat{\lambda}_1 \hat{\lambda}_2
$$

and under the alternative

$$
\log \hat{\lambda}_{alt}(x, y) = C - \frac{n}{2} \log |\hat{\Sigma}|,
$$

leading to

$$
-2 \log \hat{\lambda} = -n \log (1 - r_{uv}^2).
$$

The likelihood ratio statistic is a monotonically increasing function of the correlation coefficient between $u$ and $v$, so we can use the correlation coefficient of transformed variables to test $H_b$ against $H_{alt}$. The likelihood ratio statistic for a one-sample test of the standardised major axis was first derived by Morgan (1939) in the case where $b = 1$, and Pitman (1939), in the context of testing for equal variances. In the case of a known error variance ratio $V$ Creasy (1956) and others derived this test.

Creasy (1956) noted that the correlation test is also appropriate for the functional relationship model, although it would not be the likelihood ratio test of $H_b$ against $H_{alt}$.

Note that from Bartlett (1954), Section IIIb, with $p = 2$ and degrees of freedom $n - 1$, the statistic $-(n - 2.5) \log (1 - r_{uv}^2)$ is to second order distributed as $\chi_1^2$ under $H_b$, although constructing an $F$-ratio provides an exact test.

References


WALD, A., 1940: The fitting of straight lines if both variables are subject to error. *Annals of Mathematical Statistics* 11, 284–300.


WESTOBY, M., HOWELL, J., RICE, B., and WARTON, D., 2001: The costs of seed dispersal. Unpublished manuscript, Division of Environmental and Life Sciences, Macquarie University.