Belg. J. Zool. - Volume 127 (1997) - issue 1 - pages 99-102 - Brussels 1997

SHORT NOTE

STATISTICAL COMMENTS ON A BIOMETRICAL STUDY

ROBBY STOKS

Evolutionary Biology Group, Department of Biology, University of Antwerp (RUCA), Groenenborgerlaan 171, 2020 Antwerpen and Laboratory for Animal Ecology, Department of Biology, University of Antwerp (UIA), Universiteitsplein 1, 2610 Wilrijk. e-mail: stoks@ruca.ua.ac.be

Key-words: Biometry, multicollinearity, type II-regression.

SKOUFAS *et al.* (1) presented biometrical data on Anthozoa. My concern is in the statistical analysis and interpretation of these data. Many statistical incorrectnesses will probably not lead to interpretation errors due to the robustness of the used techniques or to the strength of the analysed effects, others however certainly will do so. This study gives the opportunity to discuss a number of statistical errors of which some occur regularly in the literature.

SKOUFAS *et al.* (1) provide the first biometrical data on the gorgonian *Eunicella singularis*. There are several problems with the tests for normality in (1). The frequency distribution of for example S (Fig. 2 in [1]) shows significant skewness, although their test shows no deviation from normality. Probably this is due to the low power of their normality tests. One tends to accept normality and use normality-based methods when samples are small. Yet, with small samples many normality-based methods have limited value (2). The authors report a single t-test statistic for the normality test for the distribution. To the best of my knowledge there are no tests for normality using a single t-value. The preferred test for normality is the Shapiro-Wilks'W test because of its good power properties as compared to a wide range of alternative tests (3).

In addition to the dubious validity of the normality tests, the p-values are misinterpreted. The authors conclude from the non-significant t-values of the normality tests for variables H and S (1 : p. 87) that the corresponding distributions are respectively not normal and normal (1 : p. 90). On the other hand it is rather strange that after acceptance of the normality of, for example the distribution of P.S., and while explicitly stating that the distributions are unimodal (1 : p. 85), the authors continue with the detection of several groups based on the «normal» frequency distribution of the character.

SKOUFAS *et al.* (1) correctly argued that a multiple regression procedure should be used to determine the relative contribution of the variables height/length (H/L), number of dichotomies (nb. dich.) and surface (S) in explaining the variance in dry weight (P.S.). It follows from the name «multiple regression» that there have to be linear relationships between each independent variable (i.c. H/L, nb. dich. and S) and the dependent variable (P.S.) (4). The relationship between H/L (S) and P.S. however is only linear in the log scale (Fig. 4A in [1]) while this transformation was not used in the multiple regression procedure. Furthermore, the dependent variable has to be normally distributed (3), which is not the case (see above). The authors report the F-statistic of the overall significance test of the multiple regression as 58.903 with 52 df. The shape of the F- distribution is however only fully determined by two values for degrees of freedom, in this case (k, n-k-1) with k the number of independent variables (i.c. 3), and n the number of observations (i.c. 55) (5). The correct significance test should therefore be based on F (3,51).

Finally, perhaps the most serious complication is the existence of correlations among the independent variables, the so-called multicollinearity problem (6). Multicollinearity implies that part of the variation in the dependent variable may be attributed to more than one independent variable (6) and therefore one can not exactly determine the unique contribution of each independent variable to the variance of P.S. (the goal of the multiple regression in this study). Multicollinearity is severe when independent variables are more related to each other than they are to the dependent variable (7). Statistically this means that the coefficient of determination of the regression when the ith independent variable is regressed against all other independent variables (r^2_i) is greater than the coefficient of multiple determination for the total model (r^2_{tot}). A log transformation of the variables (to induce linear relationships between independent and dependent variables, see higher) would not solve the problem of multicollinearity. For example the r^2 of logS with only one of the two other independent variables (nb. dich.) is already 0.8 (see Fig. 5 in [1]). The value of r^2_i for logS will be even greater because it will also include the unique variance of logS explained by H/L. As a result r^2_i will almost certainly be greater than r^2_{tot} (model with untransformed data: $r^2_{tot}=0.72$) indicating severe multicollinearity. The problem of multicollinearity can be solved by combining independent variables into principal components or by using biased estimation methods such as ridge regression (6). In addition, there is another problem in the determination of the variance explained by the independent variables. To assess the fraction of variance explained by a single factor, one compares r^2_{tot} to r^2 for the model with all except the factor of interest: the difference (r^2_i) is the variance uniquely explained by that single factor (6). As a logical result the coefficient of determination of the total model should equal the sum of the r^2_is. The r-square for the total model they report is 0.72 while the sum of the r-squares associated with each independent variable sums to 1.86 (this is also theoretically impossible because $>$ 1.00). This means that the given r-squares do not answer the purpose of the analysis, namely to show «de quelle manière interviennent les trois paramètres... dans la détermination de la variance de P.S.» (1). Therefore the explaining capacity of the independent variables will be smaller than reported.

The allometric relationships between the measured variables were determined with ordinary least squares linear regressions on log-transformed data (1, 8). However, when both variables are subject to measurement error, as is the case, a critical assumption of ordinary least squares regression has been violated and the estimates of slopes are biased (5, 9, for some examples of bias introduced in the context of allometry see 10). This is especially important if one attempts to compare scaling exponents with values expected under the null hypothesis of isometry. In such cases one should use type II regression (5, 9). In morphometric work logarithmically transformed variables are often employed, and their functional relationship should be estimated by the slope of the major axis of the bivariate sample (5). The differences between slopes estimated by both techniques decrease as the correlation coefficients increase (9). Therefore especially the slope of log H/L against log P.S. will be biased.

Even after ignoring the bias in the calculated scaling exponents, there are some problems with their interpretation. Isometry is specified by the ratio of the dimensions of the variables (9). Allometry is the departure from isometry. SKOUFAS *et al.* (1) found a scaling exponent of 0.7 between P.S. and S (after log transformation). Because this value is smaller than three, they conclude that there is a positive allometric relationship between both variables. The calculated scaling exponent should however be tested against the appropriate null hypothesis of isometry. The expected slope of a plot of log area versus log dry weight is $L^2/L^3=2/3$ (11) and not three. Therefore it is highly probable that the claimed positive allometric relationship is in fact a purely isometric one. Also the allometry of P.S. against H/L and nb. dich. against S is checked by comparing the scaling exponents at sight (and not statistically) with values generated by a wrong null hypothesis (respectively with a scaling exponent under isometry of three and two).

I would like to thank Thierry Backeljau, Luc De Bruyn and Stefan Van Dongen for useful discussions and an anonymous referee for comments on the manuscript. Wim Jacobs and Solange Goethals provided me with some key articles. The author is a Research Assistant of the Fund for Scientific Research – Flanders (Belgium FWO).

REFERENCES

1. SKOUFAS, G., M. POULICEK & C.CH. CHINTIROGLOU (1996) – Etude préliminaire de la biométrie d'*Eunicella singularis* (Esper, 1974) (Gorgonacea, Anthozoa) à la mer Egée. *Belg. J. Zool.*, **126** (2) : 85-92.
2. STEWART-OATEN, A. (1995) – Rules and judgements in statistics : three examples. Ecology, **76** (6) : 2001-2009.
3. STATSOFT (1994) – Statistica. Volume I: general conventions & statistics I. Statsoft, Tulsa.
4. MANLY, B.F.J. (1992) – The design and analysis of research studies. Cambridge University Press, Cambridge, 353 pp.
5. SOKAL, R.R. & F.J. ROHLF (1995) – Biometry. W.H. FREEMAN and Co., New York, 887 pp.
6. PHILIPPI, T.E. (1994) – Multiple regression: herbivory. In : SCHEINER, S.M. & GUREVITCH, J. (eds.) – *Design and analysis of ecological experiments*. Chapman and Hall, London : 183-210.
7. FREUND, R.J. and LITTELL, R.C. (1986) – SAS system for regression, 1986 Edition. SAS Institute Inc., Cary.

8. GOULD, S.J. (1966) – Allometry and size in ontogeny and phylogeny. *Biol. Rev.*, **41** : 587- 640.
9. LABARBERA, M. (1989) – Analyzing body size as a factor in ecology and evolution. *Annu. Rev. Ecol. Syst.*, **20** : 97-117.
10. GREENE, A.J. (1992) – Positive allometry is likely with mate choice, competitive display and other functions. *Anim. Behav.*, **43** : 170-172.
11. MCNEILL, A.R. (1971) – Size and shape. Arnold, London, 51 pp.