

Correction to the statistical note in 'Gulliver, R., 2011. Patterns of flowering on continuously-grazed dune and machair on Colonsay. The Glasgow Naturalist 25 (3) 19-28'.

Richard Gulliver^{1,2}

¹ Carraig Mhor, Imeravale, Port Ellen, Isle of Islay, Argyll PA42 7AL, UK

² Geographical and Earth Sciences, University of Glasgow, Glasgow G12 8QQ, UK

E-mail: richardgulliver@gmail.com

INTRODUCTION

The data analyses in the article on the phenology of dune and machair communities on Colonsay were correct. However the statistical note in the Methods section was incorrect. The author apologises for this error. A revised set of notes follows.

MAIN TEXT

No overlap between samples; t, z and U test

When there is no overlap between samples (that is, where all the values in one sample are greater in magnitude than all the values in the other) and the data is parametric (that is, the distribution of sample values is well approximated by a Normal distribution), the t or z test should be applied.

For sample sizes of 5 to 20 of non parametric data, and using the form of Mann Whitney U test where the lower of the two U values is the test statistic, it is advisable procedure that a test be applied. However, in these cases the outcome of operating the test is known in advance. The lower value of U will be zero. Reference to the tabulated values of U will show that significance has been obtained and that the null hypothesis can be rejected. For sample sizes of above 20 a formula exists for converting the lower U value to z (Campbell, 1974 p61). For non overlapping samples $n_1=n_2=21$, z has a probability of less than 0.1% using the formula. Hence a very highly significant difference will be obtained in all cases where both n_1 and n_2 are above 20 for non overlapping samples.

For the Mann Whitney U test some tabulated values use the higher of the two U values. Use of the lower value means there is always the same value of U which shows the maximum difference between samples i.e. 0. Use of the upper value means that the values of U associated with maximum difference between samples varies with sample size.

Paired data: the case when the trend in every pair of values is the same throughout; paired t, paired z and Wilcoxon tests

When the trend in every pair of values is the same throughout (i.e. the larger value in each pair always belongs to the same one of the two conditions) for parametric data (where the differences between the two values in each pair give a distribution which is well approximated by a Normal curve), the paired t, or paired z test should be applied.

Where the trend is the same throughout all the pairs of values of non parametric data, for sample sizes of 7 to 25, it is advisable procedure that a Wilcoxon test be applied. However, in these cases the outcome of operating the test is known in advance. The test statistic W (T) i.e. the lower value of R_+ or R_- will be zero. Reference to the tabulated values of W will show that significance has been obtained and that the null hypothesis can be rejected. For sample sizes of above 25 a formula exists for converting the lower W (T) value to z (Campbell, 1974 p66). For pairs of values when the trend is the same throughout for $n=26$, z has a probability of less than 0.1% using the formula ($n=26$ excludes zero differences). Hence a very highly significant difference will be obtained for all paired values of n above 26 when the trend is the same throughout.

Some tabulated values of W (T) use the higher value of R_+ or R_- . Use of the lower value means there is always the same value of W (T) which shows the maximum difference between the paired replicates i.e. 0. Use of the upper value of R_+ or R_- means that the values associated with maximum difference between the paired replicates varies with sample size.

Biologists do not always agree on whether data are paired or not. In case of doubt, assume data are not paired.

ACKNOWLEDGMENT

The helpful advice provided by Dr Tim Sparks is gratefully acknowledged. However, the total responsibility for the text rests with the author.

REFERENCE

Campbell, R. C. (1974, second edition). *Statistics for Biologists*. Cambridge University Press, Cambridge.

SUPPLEMENT

The application of a Mann Whitney U test to the case of non overlap on p21 of Gulliver 2011 results in a significant difference being generated, as does the application of a Wilcoxon test to the case of the trend being the same throughout in all pairs of values on p22. NB the median of 23 for 4m² machair quadrats on p21 & p23 is correct, the value in Table 1 should be 23 not 25, author's error.