

Nonparametric tests—

some comments on an *Ecology* Special Feature

Ecology, 76(6), 1995, pp. 1997–1998
© 1995 by the Ecological Society of America

DISTRIBUTION-FREE AND ROBUST STATISTICAL METHODS: VIABLE ALTERNATIVES TO PARAMETRIC STATISTICS?¹

STEPHEN M. SMITH

Department of Biology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Abstract. In much ecological hypothesis testing, Type II errors may be more serious than Type I errors; ecologists are encouraged to make an a priori assessment of the relative costs of both types of error. Distribution-free statistics such as the Mann-Whitney procedure are not necessarily more powerful nor more robust than their parametric analogues.

Key words: *hypothesis testing; nonparametric vs. parametric tests; statistical power; Type I and Type II errors.*

The lead-off paper (Potvin and Roff 1993) in the recent *Special Feature*—“Statistical Methods: An Upgrade for Ecologists”—contains some errors and misinterpretations.

Potvin and Roff (1993:1621) state: “Outliers might lead to the rejection of a true null hypothesis in ANOVA's and, to biologists, this is the most [*sic*] costly type of error.” It is alarming to see such a statement in a leading ecological journal, many readers of which are involved in applied ecological research; it is doubly alarming in a series of papers devoted to the statistical education of ecologists. Many of the ecological and environmental problems we face can be viewed as partly the result of Type II errors (Peterman 1990). Whether a Type I or a Type II error is more serious is context specific. Ecologists should be encouraged to make an a priori assessment of the relative costs of *both* types of error and to set α and β accordingly. A blanket statement that Type I error is the more costly of the two errors can serve only to undermine the recent progress we've been making in encouraging biologists to think about Type II error and issues of power and effect sizes.

There are serious errors in the interpretation of the tree-diameter data illustrated in Fig. 1 of Potvin and Roff (1993). Potvin and Roff's Table 2 (1993:1619), which summarizes the (flawed) statistical tests of these data, will leave readers with the erroneous impression that nonparametric procedures such as the Mann-Whitney test are orders of magnitude more powerful than

analogous parametric procedures, even when the parametric procedures are carried out on transformed data. Only the *t* test on the untransformed data is correct; all other statistics and their interpretations are wrong. The correct statistics are provided in an erratum by Potvin and Ross on page 2000 of this issue. All three parametric tests are in close agreement, a result that could have been anticipated from the figure alone—none of the transformations substantially altered the distribution of the data by comparison to the distribution in the original metric.

The nonparametric procedure (the Mann-Whitney test) on these same data yields essentially the same result as the three parametric procedures, but the proper conclusion is not the one the authors might draw. The conclusion that should have been reached with these data is not that the Mann-Whitney test is a preferred or more powerful test but that the data set is marginal for testing the hypothesis at hand—we need more data, not different tests! What the example does show is how dependable both parametric and nonparametric procedures are—both approaches lead to the same interpretation.

Ecologists often may have data sets for which nonparametric approaches might be preferred because such tests make fewer distributional assumptions than their parametric analogues, or data sets that are so small that the distributional assumptions cannot be reliably checked. Unfortunately, the exploration by Potvin and Roff (1993) of an individual data set can provide little guidance on preferred analytical procedures. For most data sets the Mann-Whitney and the *t* test will lead to

¹ Manuscript received 15 July 1994; revised 26 October 1994; accepted 27 January 1995.

a similar qualitative conclusion. If one suspects a violation of the assumptions of parametric methods, then the nonparametric procedures may be preferred in order to have greater trust in the resulting P value, but one should not expect a nonparametric procedure to fundamentally alter the interpretation of an experiment.

LITERATURE CITED

- Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. *Canadian Journal of Fisheries and Aquatic Sciences* **47**:2–15.
- Potvin, C., and D. A. Roff. 1993. Distribution-free and robust statistical methods: viable alternatives to parametric statistics? *Ecology* **74**:1617–1628.

Ecology, 76(6), 1995, pp. 1998–2000
© 1995 by the Ecological Society of America

STATISTICAL SIRENS: THE ALLURE OF NONPARAMETRICS¹

DOUGLAS H. JOHNSON

Northern Prairie Science Center, National Biological Service, Jamestown, North Dakota 58401 USA

Abstract. Although nonparametric statistical methods have a role to play in the analysis of data, often their virtues are overstated and their deficiencies overlooked. A recent Special Feature in *Ecology* advocated nonparametric methods because of an erroneously stated advantage that they require no assumptions regarding the distribution underlying the observations. The present paper points out some often-ignored features of nonparametric tests comparing two means, and advocates parameter estimation as a preferred alternative to hypothesis testing in many situations.

Key words: hypothesis testing; nonparametric methods; normal distribution; parameter estimation; Wilcoxon-Mann-Whitney test vs. t test.

How often have you read something like, “Our data were not normally distributed, so we used nonparametric methods.”? The reasoning is that nonparametric methods require few, if any, assumptions. In a recent Special Feature article in *Ecology*, Potvin and Roff (1993:1619) made the point explicit: “The main advantage of nonparametric methods over their parametric counterparts is the *absence of assumptions regarding the distribution underlying the observations*” (emphasis added). Numerous authors have made similar statements, but I focus on the Potvin and Roff article because it was intended as an update for ecologists. My purpose here is to indicate that their characterization is incorrect and the implied advice is misleading.

The situations for which nonparametric methods are commended vary, and include correlation, regression, and more, but I concentrate on the comparison of two means. The usual (parametric) method is the t test, which can be employed either when variances within the two groups are the same (ordinary Student's t) or when they differ (Welch-Satterthwaite modification).

The nonparametric counterpart is the Wilcoxon rank sum or the equivalent Mann-Whitney test (WMW test).

Several points bear emphasis. First, data do not need to be normally distributed in order to apply the t test. Only the means need to be, and that property is assured by the Central Limit Theorem, even for relatively small samples, for all but the most perverse data. This is exemplified in Fig. 1, which shows at the upper left a very nonnormal (in fact, a uniform) distribution of original data. Random samples of size $N = 2, 4,$ and 8 demonstrate that the distribution of averages based on even those small sample sizes rapidly approaches normality.

Second, statements are often made about means of distributions differing, based on nonparametric tests such as WMW, although Potvin and Roff did not make this mistake. The WMW test actually tests the hypothesis that the two distributions are *identical*, not that they have the same mean (e.g., Gibbons 1985). In particular, variances must be the same if the test is to compare means; as Hollander and Wolfe (1973:71) stated, “we assume that the two populations do not differ in dispersion.” To compare means, the WMW test requires the assumption that the two distributions are

¹ Manuscript received 23 May 1994; revised 16 December 1994; accepted 27 January 1995; final version received 21 February 1995.