Ecology, 92(1), 2011, pp. 3–10

The arcsine is asinine: the analysis of proportions in ecology

David I. Warton[1,2,3] and Francis K. C. Hui[1]

[1]*School of Mathematics and Statistics, The University of New South Wales, Sydney, NSW 2052 Australia*
[2]*Evolution and Ecology Research Centre, The University of New South Wales, Sydney, NSW 2052 Australia*

Abstract. The arcsine square root transformation has long been standard procedure when analyzing proportional data in ecology, with applications in data sets containing binomial and non-binomial response variables. Here, we argue that the arcsine transform should not be used in either circumstance. For binomial data, logistic regression has greater interpretability and higher power than analyses of transformed data. However, it is important to check the data for additional unexplained variation, i.e., overdispersion, and to account for it via the inclusion of random effects in the model if found. For non-binomial data, the arcsine transform is undesirable on the grounds of interpretability, and because it can produce nonsensical predictions. The logit transformation is proposed as an alternative approach to address these issues. Examples are presented in both cases to illustrate these advantages, comparing various methods of analyzing proportions including untransformed, arcsine- and logit-transformed linear models and logistic regression (with or without random effects). Simulations demonstrate that logistic regression usually provides a gain in power over other methods.

Key words: arcsine transformation; binomial; generalized linear mixed models; logistic regression; logit transformation; overdispersion; power; Type I error.

Introduction

Proportional data are widely made use of in ecology. During 2008–2009, over one-third of papers published in *Ecology* analyzed proportions, according to a review of every fourth article published during this period (Table 1, 51 of 134 papers analyzed proportions). The simplest approach, used in 19 of 51 papers (37%; Table 1), was to analyze untransformed sample proportions using linear models (e.g., ANOVA, linear regression, or some generalization). The most common method of analysis was to utilize the arcsine square root ("arcsine" henceforth) transform followed by a linear model, with 20 (39%) adopting this procedure. This transform is recommended in statistical texts for biologists and ecologists (Sokal and Rohlf 1995, Zar 1998, Gotelli and Ellison 2004), although is notably absent from many applied regression texts aimed at practicing statisticians (Myers 1990, Venables and Ripley 2002). The likely reason behind this disparity is that the arcsine transform has been superseded by more modern methods of analysis such as logistic regression, used in only 13 (25%) papers, and not used at all with ANOVA designs.

There are two situations where the arcsine transform is used, which are very different from a statistical point of view. First, data might be binomial in the sense that they are of the form "x out of n," e.g., 18 of 25 bats survived. Such data may or may not be overdispersed, as discussed later. Second, data might be collected on a proportional scale (values between 0 and 1) but are not actually binomial. Instead the data are usually continuous and often considered as percentages, e.g., 78% nitrogen content. Currently the arcsine transform is used in the analysis of both these situations (Table 1).

In this paper, we argue that the arcsine transform should no longer be used in either situation. Where data are binomial, logistic regression should be used instead, with random effects if overdispersion is found. Where data are non-binomial, there is no motivation to use the arcsine transform at all, and instead efforts should be placed into searching for a transform that satisfies linearity assumptions, while if possible, being useful for interpretation.

In the context of analyzing sex-ratio data, Wilson and Hardy (2002) produced an extensive discussion as to why logistic regression should replace the arcsine

Manuscript received 18 February 2010; revised 16 June 2010; accepted 15 July 2010. Corresponding Editor: B. D. Inouye.
[3] E-mail: David.Warton@unsw.edu.au

TABLE 1. Literature review revealing the main methods of analyzing proportional data in ecology.

Method	Binomial†	Non-binomial	Number of papers	Percentage of papers
Untransformed linear model	7 (7)	12 (8)	19 (15)	37 (54)
Arcsine linear model	14 (8)	6 (4)	20 (12)	39 (43)
Logistic regression/GLMM	13 (0)	0	13 (0)	25 (0)
Logit linear model	0	1 (1)	1 (1)	2 (4)
Total	34 (15)	19 (13)	53 (28)‡	104 (100)‡

Notes: Results are based on 51 papers analyzing proportional data, found in a sample of 134 papers published in *Ecology* as articles, from 2008 to 2009, ignoring the CORONA issue of November 2008. Results for ANOVA designs only are expressed in parentheses.

† Articles are classified according to whether or not they report any analyses in which the response variable is a binomial (x out of n) proportion, or a non-binomial proportion.

‡ Note that two articles analyzed proportions using both an arcsine linear model and logistic regression.

transform. This paper extends those results in several ways: describing the use of generalized linear mixed models (GLMM) when the data are binomial and overdispersed; emphasizing the use of logistic regression for ANOVA designs where it is currently rarely used (Table 1); considering the case of non-binomial proportions; discussing the issue of interpretability and the limitation of the arcsine transform in this respect; and broadening the focus to all types of proportional data collected in ecology as opposed to just sex-ratio data.

METHODS

Binomial data

A variable x is known to be binomial if it counts the number of times some outcome (usually dubbed a "success") occurs out of n independent trials, each having the same probability of "success" p. If the response variable is a proportion calculated as $y = x/n$, the first-choice model is the binomial distribution with parameters n and p. The variance of such binomial proportions will be a quadratic function of the mean: $\mathrm{var}(y) = p(1 - p)/n$. Hence the equal variance assumption required in linear models is not satisfied, which can (among other things) reduce power at detecting differences between proportions, as seen in *Simulations*.

The motivation behind using the arcsine transform comes from its ability to stabilize the variance of binomial data, in the sense that it becomes approximately constant after transformation. This result, proven using a Taylor series expansion (known as a "delta method" result; Shao 1998), is

$$\mathrm{var}(\arcsin\sqrt{y}) \approx \frac{\mathrm{var}(y)}{4p(1-p)} = \frac{p(1-p)}{4np(1-p)} = \frac{1}{4n}. \quad (1)$$

The result is approximate however, and typically does not work well if p is close to 0 or 1. Note also from Eq. 1 that the transform only achieves approximately equal variance when the same number of trials n is used to estimate each sample proportion. In other circumstances, the arcsine transform fails to achieve its goal of variance stabilization. The transform also has the effect of making the transformed proportions roughly normally distributed, although how well normality is achieved depends upon sample size. Normalization is a desirable goal but, contrary to Wilson and Hardy (2002), it should *not* be the main goal of transformations, given that linear models are robust to non-normality but sensitive to heteroscedasticity (Faraway 2006).

The arcsine transform reached its zenith in the mid-20th century due to its simple application, and hence is suggested in texts whose first editions appeared around the 1960s and 1970s (e.g., Sokal and Rohlf 1995, Zar 1998). However with the advent of faster computers and the increasing functionality of statistical programs, logistic regression has emerged as an alternative approach. Logistic regression is a special case of generalized linear models (Dobson 2002) that broadens the linear model in two ways: by assuming the data are binomial rather than normal; and by assuming the "logit" of proportions, $\log(p/[1-p])$, is a linear function of the predictors (as opposed to applying this assumption to the sample proportions themselves).

We propose the use of logistic regression over an arcsine transformed linear model to binomial data for three reasons: (1) The logit-link function is monotonic and maps [0,1] to the whole real line, ensuring first that predicted proportions will always be between 0 and 1, and second that if the relationship between logit(p) and x is determined to be increasing, so will the relationship between p and x over all possible x. One might think these criteria to be a minimal requirement of any transformation of proportions, and yet the arcsine transform does not satisfy the latter due to the sine function's periodicity. This is especially a major concern if extrapolation is a primary purpose of the constructed model, although not the only situation in which non-monotonicity can result in problems. (2) The coefficients in logistic regression have a natural interpretation: a one unit increase in the predictor x leads to an increase in the odds $p/(1-p)$ by a factor of e^{β}, where β is the relevant regression parameter. (3) Logistic regression correctly models the mean–variance relationship of binomial data as $\mathrm{var}(y) = p(1-p)/n$. Fitting a linear model instead with an arcsine-transformed response does not use this relationship exactly, only approximately so. This can potentially lead to incorrect conclusions (as standard errors may be biased) and inefficiency (sample estimates of parameters may be not as precise). Note that of the three advantages listed above, (1) and (2) relate to the logit function itself. Therefore even for non-binomial data, a logit transformation can be useful for these very two reasons, as discussed later.

Sometimes the data are of the form "x out of n" but not exactly binomial, due to additional factors which result in variability exceeding that expected by the binomial distribution. This is commonly referred to as overdispersion. One simple method of diagnosing overdispersion is to consider the ratio of the residual deviance (deviance being a likelihood-based measure of goodness of fit analogous to sums of squares) to its degrees of freedom: if substantially greater than one then data are overdispersed compared to the binomial (Dobson 2002). More formal methods of assessing overdispersion are available (Collett 2002). A useful approach for modeling overdispersion is to add a normally distributed random intercept term to the model for each binomial count (Jiang 2007). The subsequent model is a mixed effects logistic regression, a special case of generalized linear mixed models (GLMM). While the mathematics of GLMM estimation are challenging (Jiang 2007), the approach has been implemented in many statistics packages nowadays e.g., glmer( ) from the lme4 package in R (Bates et al. 2009).

Quasi-binomial logistic regression, in which the variance of x is $\text{var}(x) = \phi np(1 - p)$ where ϕ is known as the overdispersion parameter, is also sometimes used to model overdispersed data. However, this is an ad hoc approach which does not correspond to any known distribution, and provides a plausible model for overdispersion only in restricted settings (Collett 2002).

Non-binomial data

If data are non-binomial, i.e., not of the form "x out of n," then logistic regression is no longer applicable, and usually the distribution of the data is no longer known. Instead an alternative approach is to transform proportions in order to (approximately) fulfill linear modeling assumptions. Consequently, there is no justification to prefer the arcsine transform over any other transform, as variance stabilization is now no longer a goal.

One criterion that should be considered in choosing a transformation is interpretation, and parameters from an arcsine-transformed linear model fit are not simple to interpret. Further, it is desirable to choose a transform that maps proportions $y \in [0, 1]$ monotonically to the whole real line $(-\infty, \infty)$. If this is not done then it becomes possible to obtain nonsensical predicted values, as illustrated in *Examples*.

One transform that does satisfy both the above criteria is the logit transform: $\log(y/[1 - y])$. This both maps proportions to the whole real line and has a natural interpretation as described in reason 2. One difficulty, though, with using this transform is that sample proportions equal to 0 and 1 transform to undefined values $-\infty$ and ∞, respectively. An ad hoc solution to this problem is to add some small value ε to both the numerator and denominator of the logit function, which introduces minimal bias while still satisfying the criteria above. This approach is a modification of the "empirical logistic transform" (Collett 2002) for data that are not discrete. We propose taking as ε the minimum non-zero proportion y, or if proportions are large, the minimum non-zero value for $1 - y$. This approach is compared with alternatives in Appendix B, where results suggest the qualitative interpretation of results might not be sensitive to the value of n, but slope coefficients may be. Users are encouraged to experiment with different values of ε.

Whatever transform is ultimately used, it is important to check diagnostic plots (Faraway 2006) to assess how well the transformed values satisfy linear modeling assumptions. A particularly important plot is that of residuals vs. fits (or residuals vs. predictor variables), which is used to check for absence of any pattern, i.e., no evidence of nonlinearity or heteroscedasticity. In small sample sizes (e.g., $n < 30$), a normal probability plot should also be considered. Likewise in the case of GLMM, it is important to check that the random effects component of the model has no evidence of a systematic trend and is roughly normally distributed (Faraway 2006). See the Supplement for illustrations of how to do this in R.

Examples

In this section, we present two examples from the ecology literature illustrating the methods above, and the advantages of using logistic regression or logit-transform. All analysis and graphs were done in R 2.9.0 (R Development Core Team 2009).

Something fishy

The first example comes from an article examining the patterns of energy balance in teleost fish (Arrington et al. 2002). It set out to determine whether or not the proportion of fish with empty stomachs, sampled from various geographic locations, differed across trophic levels (Fig. 1a). The response variable was binomial, with success being that a fish had an empty stomach. In the original analysis, the proportions were analyzed using the arcsine transform followed by a two-way ANOVA. The arcsine transform is commonly used to analyze this type of data, with 8 of 14 (57%) articles in our survey utilizing some form of ANOVA after arcsine transforming a binomial response. There appears to be a lack of recognition that logistic regression can be used in such situations, and hence this example emphasizes that logistic regression/GLMM can handle ANOVA designs just as effectively as they handle the standard regression setup. See the Supplement for related R code.

Results suggest there was significant variation across trophic levels within each geographic location, when

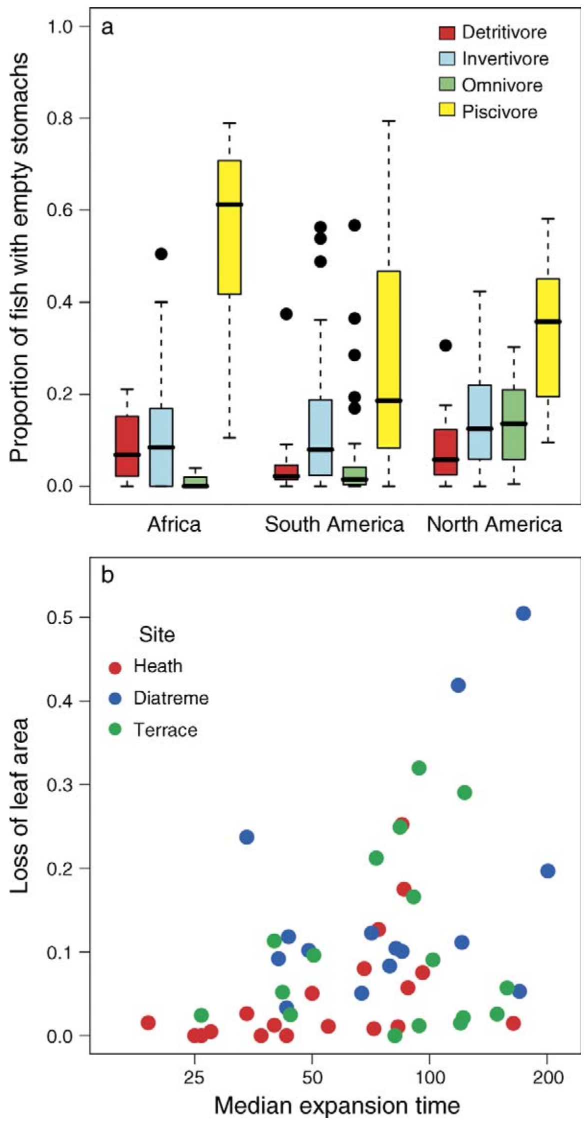

FIG. 1. Data sets used in *Examples*. (a) Proportion of fish with empty stomachs by geographic location and trophic group (Arrington et al. 2002). The limits of a box denote the upper and lower quartiles, the horizontal bar is the median, and the 1.5IQR criterion has been used to classify outliers. (b) Percentage loss of leaf area in relation to median expansion time and site (Moles and Westoby 2000). Note the log scale in panel (b).

applying ANOVA to untransformed proportions ($F_{9,242} = 14.3$, $P < 0.001$) or arcsine-transformed proportions ($F_{9,242} = 13.5$, $P < 0.001$; Appendix A). When logistic regression was used instead, with a random intercept, the same conclusion was reached ($\chi^2_9 = 88.8$, $P < 0.001$). A random intercept was included in the model to account for overdispersion, for which there was very strong evidence (residual deviance: 5005 on 242 df), meaning that there was considerable species to species variation unexplained by the two predictors (location and trophic level).

Although there were no qualitative differences in results across the three methods presented above, this will not always be the case. To illustrate, we reanalyzed the data of Arrington et al. (2002) considering nocturnally feeding fish only. While the effect of trophic level was marginally nonsignificant when using ANOVA on untransformed proportions ($F_{6,47} = 2.25$, $P = 0.055$) or arcsine-transformed proportions ($F_{6,47} = 2.19$, $P = 0.061$), it was significant when using GLMM ($\chi^2_6 = 14.8$, $P = 0.022$). This trend, where results are sometimes significant when using logistic regression but not when using the arcsine transform, is not unusual, as will be shown in *Simulations*.

Residuals vs. fits plots showed strong evidence of heteroscedasticity in the untransformed linear fit (Fig. 2a, left), as indicated by a fan-shaped pattern. Presence of heteroscedasticity means that the proportions fail to satisfy the linear modeling assumption of constant variance, invalidating subsequent model inferences. Heteroscedasticity was lessened but still apparent in the arcsine transform residual plot (Fig. 2a, center), whereas the GLMM fit did not exhibit any pattern (Fig. 2a, right), indicating a better model fit. Normal probability plots suggest the normality assumption was reasonably satisfied in all cases (Appendix A).

Expanding leaves

This second example forms part of a paper asking whether plant species with small leaves have shorter expansion times than large leaved counterparts (Moles and Westoby 2000). It was chosen as it involved a non-binomial response, namely percentage loss of leaf area (LLA), which was regressed against median expansion time ($\log_{10}$-transformed) and site (Fig. 1b). The arcsine transformation was used for analysis, as is common for this type of data (6 of 19, or 32%; Table 1), unless no transformation is performed at all (12 of 19, or 63%).

LLA increased significantly with expansion time for untransformed ($F_{1,47} = 5.26$, $P = 0.026$) and arcsine-transformed proportions ($F_{1,47} = 6.19$, $P = 0.016$; Appendix B), after adjusting for site. The same regression was then performed but with logit-transformed LLA as the response. Some of the proportions were equal to 0, so the smallest non-zero percentage response (0.48%) was added to the logit function (see *Methods*). Conclusions were as previously, i.e., expansion time was significantly related to LLA after adjusting for site ($F_{1,47} = 6.11$, $P = 0.017$).

A residuals vs. fits plot for the linear model with untransformed percentages indicated non-constant variance, as illustrated by the funnel-shaped pattern (Fig. 2b, left). While still present, this trend was much reduced by the arcsine transform (Fig. 2b, center), and was not evident after the logit transform (Fig. 2b, right). Normal probability plots indicated that both the arcsine and the logit transform approximately normalized the data (Appendix B).

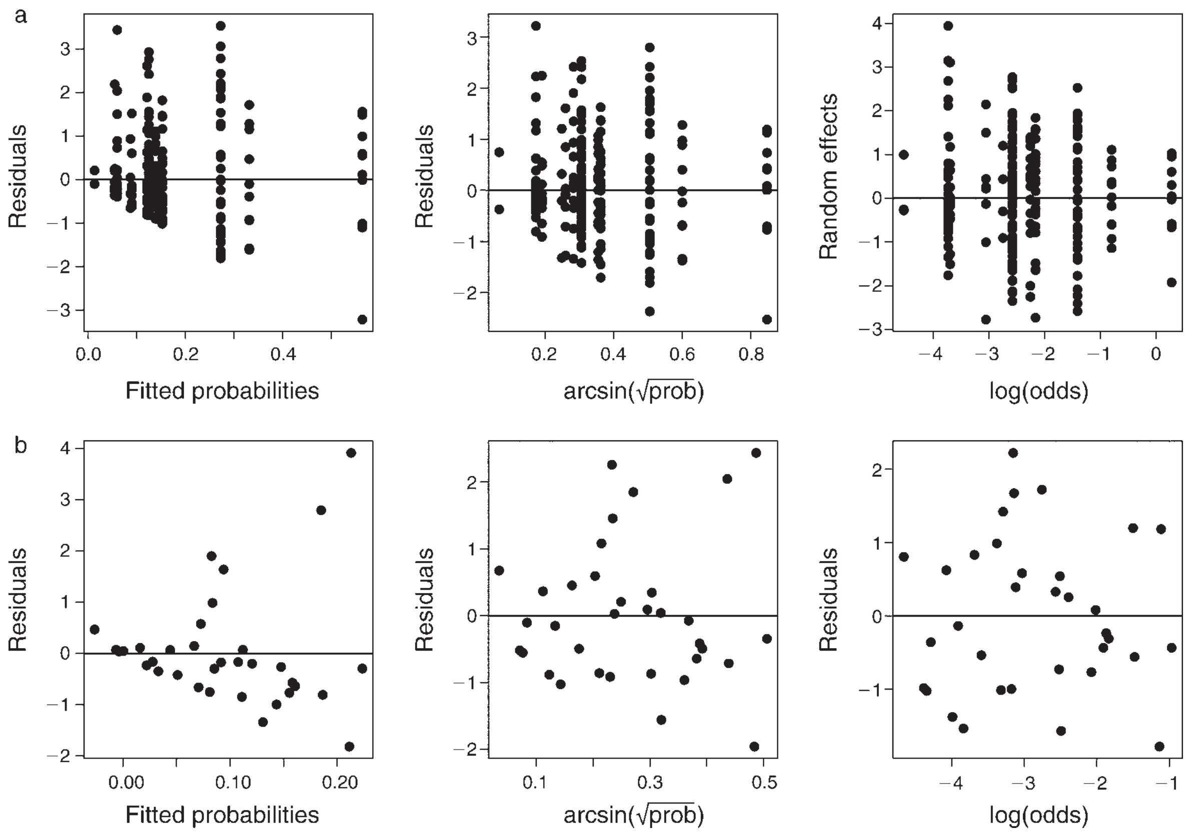

FIG. 2. Residual vs. fitted values plots for three different analyses of (a) the "something fishy" data shown in Fig. 1a, and (b) the "expanding leaves" data shown in Fig. 1b. Methods of analysis include: untransformed proportions linear model (left column), arcsine-transformed proportions linear model (center column); and in the right column of panel (b) logit-transformed linear fit. The right column of panel (a) presents estimated random effects vs. fitted values from a GLMM fit. Note the fan-shaped pattern in the untransformed residual plot, suggesting a violation of the homoscedasticity assumption. This is evident to a lesser extent after arcsine transformation and is no longer evident when using logistic methods.

The advantages for interpretation of analyzing logit-transformed proportions can be illustrated by considering the slope parameter for the effect of expansion time. For logit-transformed LLA, $b = 1.71$, meaning a one unit increase in $\log_{10}$(median expansion time) (i.e., a 10-fold increase in median expansion time) resulted in a predicted increase in the "odds of LLA" (lost leaf area divided by leaf area remaining) by a factor of $e^{1.71} = 5.5$. In contrast, parameters from the arcsine-transformed model have no simple interpretation. It should be noted however that the estimate of the slope for logit-transformed data was sensitive to choice of ε (Appendix B).

Making predictions of leaf area lost at different expansion times demonstrate further weaknesses of analyzing untransformed or arcsine-transformed proportions. When modeling untransformed proportions, predicted LLA at an expansion time of 19 days was negative (−1.3%, all predictions at Site = Heath) and thus biologically impossible. The arcsine-transformed linear fit obtained a plausible prediction at 19 days (0.42%), but because of the non-monotonicity of the transform, this was also the predicted value at some other expansion times (e.g., 5.35 days), despite the significant increasing relationship between expansion time and arcsine-transformed LLA. Furthermore, if expansion time were less than 5.35 days, then the predicted LLA would be *greater* than 0.42% despite the slope of expansion time being positive. Fortunately no leaves in this data set expanded this quickly (Moles and Westoby 2000), but it is plausible (e.g., Aide and Londono 1989). In contrast, a linear fit to logit-transformed LLA yields two sensible predicted values (0.005% and 0.76% at expansion time = 5.35 and 19 days, respectively).

SIMULATIONS

Design

Using simulations, the approaches discussed in *Methods* were compared in terms of their accuracy at maintaining a significance level of 0.05 when there is no effect to be detected (Type I error) and in their ability to detect an effect that is present (power). A method is performing well in simulations if it has Type I error near 0.05 and comparably high power.

A two-sample design was used to mimic data from a study by Sanford and Worth (2009), which set out to

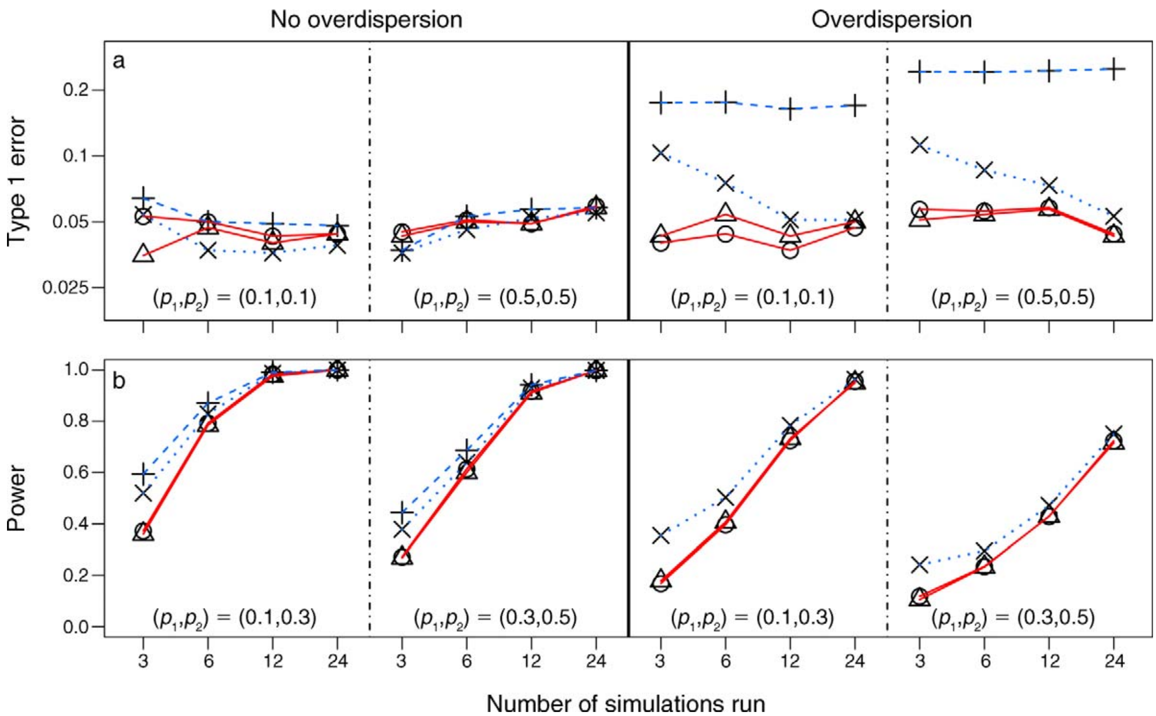

FIG. 3. Summary of simulations results (for balanced designs only) from (a) Type I error simulations and (b) power simulations. Simulations considered a two-sample design, with true proportions in the two samples denoted as p_1 and p_2, and data generated as binomial without overdispersion (left) and with overdispersion (right). Results are reported for tests using a linear fit to untransformed proportions (circles), a linear fit to arcsine proportions (triangles), logistic regression (+), and GLMM (×) likelihood ratio tests. Logistic regression power is not reported for overdispersed data (b, right), because of unacceptably high Type I error (a, right). Note the gain in power of logistic regression and GLMM as compared to untransformed and arcsine-transformed methods, but also note that in more computationally intensive simulations that keep closer control of Type I error (Appendix F), the difference in power was not as large. Note log scale on *y*-axis of panel (a).

determine whether there were geographic differences in the behavior of a predatory snail species. Snails of 64 different lineages, originating from two regions in the United States (Oregon and California), were compared in their ability to drill mussels. The response variable, measured for each lineage, was the proportion (out of 12 snails) that drilled at least one mussel.

In our simulations, there were two groups of observations, each coming from a bin(12, p) distribution. In Type I error simulations, p was held constant across both groups at 0.05, 0.1, or 0.5, whereas in power simulations p varied across the groups (0.05 vs. 0.1, 0.1 vs. 0.3, 0.3 vs. 0.5). Other properties varied included (1) number of observations in each group ($n = 3, 6, 12, 24$); (2) presence or absence of overdispersion, implemented by adding a normally distributed random effect with mean 0 and standard deviation 0.94 (this being the value obtained from a GLMM fit to Sanford and Worth 2009); (3) whether the design was balanced or unbalanced, i.e., equal or unequal number of observations per group. For each configuration, 1000 simulated data sets were generated. In total, six tests were compared: untransformed proportions linear fit (equivalent to a two-sample t test, as in van Belle et al. 2004: section 2.6), arcsine transformed proportions linear fit, logistic regression (both a Wald test and a likelihood ratio test), and GLMM (Wald and likelihood ratio test). Wald tests are the default output from most logistic regression software, although likelihood ratio tests often have better properties (Faraway 2006).

It was predicted a priori that logistic regression and GLMM would have higher power than the arcsine method. The reason for this is that maximum likelihood theory dictates that in large sample sizes no unbiased estimation procedure is more efficient than maximum likelihood estimation, provided that the model is correct (Shao 1998).

While only major results are highlighted here, full details of Type I error simulations are available in Appendix C, and of power simulations in Appendix D. A simulation using a 2 × 2 design, which produced similar results, is in Appendix E.

Type I error: results

In simulations with no overdispersion, there was little difference across statistics in Type I error (Fig. 3a, left). For $n = 3$ or 6 logistic regression likelihood ratio tests tended to be slightly liberal and linear model statistics tended to be slightly conservative, but only to a practically significant extent in extreme cases ($p = 0.05$, $n = 3$, Appendix C).

A similar pattern was seen in simulations with overdispersed data, but with more noticeable Type I error inflation for GLMM at small sample sizes (Fig. 3a, right). For $n = 3$, Type I error for GLMM was as high as 0.11, but for $n = 12$ it was much closer to 0.05 (0.051–0.073). Note that when overdispersion was present, Type I error for logistic regression failed to approach 0.05 for increasing n unless a random effect was included in the model to account for overdispersion (Fig. 3a, right). Failing to account for overdispersion leads to underestimation of sample uncertainty and hence overestimation of statistical significance.

Having an unequal number of observations per group had little effect on results (Appendix C).

In all Type I error simulations, Wald tests for logistic regression and GLMM became conservative as the probability p approached 0 or 1, although this effect lessened as sample size increased (Appendix C). This effect is well known, and it is recommended that the likelihood ratio or "analysis of deviance" test be used routinely instead (as in Faraway 2006).

Power: results

The most striking result in power simulations was that logistic regression and GLMM always had higher power than untransformed and arcsine transformed linear models (Fig. 3b). This trend was apparent in all simulations at all sample sizes. Unbalanced designs did not appear to affect the rank order of differences in power across methods, although the magnitude of the differences in power did vary depending on which group had the larger sample size (Appendix D).

It should be noted however that much of the power advantage at small sample sizes was due to differences in Type I error (with logistic regression being too liberal and linear models too conservative). This can be corrected for using resampling, as in the simulations of Appendix F. Those results demonstrate that after correcting for differences in Type I error, GLMM maintains a power advantage usually in the order of 5–15% for small sample sizes, much smaller than that suggested by Fig. 3. Logistic regression had higher power than the arcsine transform in most simulations, but often had lower power for very small sample sizes ($n = 3$).

Discussion

In this paper, we demonstrated using theory, examples, and simulations that logistic regression and its random-effects counterpart have advantages over analysis of arcsine-transformed data in power and interpretability. For binomial data, power tended to be higher when using a logistic regression approach than arcsine-transformed linear models. In addition, the logit function has a much simpler interpretation, while avoiding the possibility of nonsensical predicted values. For non-binomial proportions, there was never any theoretical reason to use the arcsine transform in the first place, and we instead suggest using the logit transform. It is important to recognize that these ideas apply equally well to proportions collected in ANOVA designs as to those collected in a regression context.

When applying logistic regression to binomial data, it is always important to check data for overdispersion. If it is present but not accounted for, standard errors will be underestimated and statistical significance overestimated, often considerably so.

If data are binomial and the number of observations is very small, we saw that Type I error may become a problem when applying logistic regression. The reason for this is that logistic regression uses only large-sample methods of inference (Shao 1998). In such cases, one might consider using resampling (Davison and Hinkley 1997) or Markov chain Monte Carlo (MCMC) functions to make more accurate inferences (Jiang 2007). The Supplement contains R code for bootstrap-based hypothesis testing in one-factor designs, as well as simulation results demonstrating the effectiveness of this technique.

Significant progress has been made in the last few decades in ecological statistics, and some methods popular decades ago have been superseded by more effective approaches. The application of logistic regression and GLMM to binomial data is a case in point. These methods are implemented in many statistical packages nowadays, and many helpful references are available which introduce scientists to the fitting and interpretation of such models (Faraway 2006, Bolker et al. 2009). Readers are encouraged to use modern alternatives to the arcsine transform, and we hope that textbook revisions will no longer suggest it as a transformation for proportional data, as it can be considered an historical antiquity.

Acknowledgments

We are grateful to Eric Sanford and Angela Moles for data, and to the latter for the witty title. F. K. C. Hui was supported by a UNSW Faculty of Science Vacation Scholarship and D. I. Warton by the Australian Research Council Discovery Projects funding scheme (project number DP0987729).

Literature Cited

Aide, T. M., and E. Londono. 1989. The effects of rapid leaf expansion on the growth and survivorship of a lepidopteran herbivore. Oikos 55:66–70.

Arrington, D., K. Winemillet, W. Loftus, and S. Akin. 2002. How often do fishes "run on empty"? Ecology 83:2145–2151.

Bates, D., M. Maechler, and B. Dai. 2009. lme4: linear mixed-effects models using S4 classes. R package version 0.999375-31. ⟨http://cran.r-project.org/web/packages/lme4⟩

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 24:127–135.

Collett, D. 2002. Modelling binary data. Second edition. Chapman and Hall, Boca Raton, Florida, USA.

Davison, A. C., and D. V. Hinkley. 1997. Bootstrap methods and their application. Cambridge University Press, Cambridge, UK.

Dobson, A. J. 2002. An introduction to generalized linear models. Chapman and Hall, London, UK.

Faraway, J. J. 2006. Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. Chapman and Hall, Boca Raton, Florida, USA.

Gotelli, N. J., and A. M. Ellison. 2004. A primer of ecological statistics. Sinauer Associates, Sunderland, Massachusetts, USA.

Jiang, J. 2007. Linear and generalized linear mixed models and their applications. Springer, New York, New York, USA.

Moles, A., and M. Westoby. 2000. Do small leaves expand faster than large leaves, and do shorter expansion times reduce herbivore damage? Oikos 90:517–524.

Myers, R. H. 1990. Classical and modern regression with applications. Second edition. PWS-Kent, Boston, Massachusetts, USA.

R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨www.r-project.org⟩

Sanford, E., and D. J. Worth. 2009. Genetic differences among populations of a marine snail drive geographic variation in predation. Ecology 90:3108–3118.

Shao, J. 1998. Mathematical statistics. Springer, New York, New York, USA.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. Third edition. W. H. Freeman, New York, New York, USA.

van Belle, G., L. D. Fisher, P. J. Heagerty, and T. Lumley. 2004. Biostatistics: a methodology for the health sciences. Second edition. John Wiley and Sons, Hoboken, New Jersey, USA.

Venables, W. N., and B. D. Ripley. 2002. Modern applied statistics with S. Fourth edition. Springer, New York, New York, USA.

Wilson, K., and I. C. W. Hardy. 2002. Statistical analysis of sex ratios: an introduction. Pages 48–92 *in* I. C. W. Hardy, editor. Sex ratios: concepts and research methods. Cambridge University Press, New York, New York, USA.

Zar, J. H. 1998. Biostatistical analysis. Fourth edition. Prentice Hall, Englewood Cliffs, New Jersey, USA.

APPENDIX A

Tables and graphs for the analysis of the "something fishy" example: proportion of fish on empty stomachs by geographic location and trophic group (*Ecological Archives* E092-001-A1).

APPENDIX B

Tables and graphs related to the analysis of the "expanding leaves" example: percentage loss of leaf area (lla) by median expansion time and site (*Ecological Archives* E092-001-A2).

APPENDIX C

Results of Type I error simulations for binomial data, and binomial data with overdispersion (*Ecological Archives* E092-001-A3).

APPENDIX D

Results of power simulations for binomial data, and binomial data with overdispersion (*Ecological Archives* E092-001-A4).

APPENDIX E

Results of simulations using a balanced 2×2 design, both for binomial data and binomial data with overdispersion (*Ecological Archives* E092-001-A5).

APPENDIX F

Results of simulations using resampling based hypothesis testing to control Type I error in small samples (*Ecological Archives* E092-001-A6).

SUPPLEMENT

R code demonstrating how to fit a logistic regression model, with a random intercept term, and how to use resampling-based hypothesis testing for inference (*Ecological Archives* E092-001-S1).