Discussion of Paper by Bendel Fygenson

Mark S. Kaiser and Daniel J. Nordman
Department of Statistics
Iowa State University
June 2007
Professor Fygenson has produced a thought-provoking paper that contains many particulars worthy of consideration. We confine ourselves to an examination of the concept of pessimistic and optimistic outlooks, and the way that this concept translates into the behavior of response curves for binomial generalized linear models. Such response curves are connected with distribution functions, and we believe the pessimistic/optimistic notion can be formulated in terms of entire distributions, rather than only on specified intervals as done in the paper. This greatly simplifies the conceptual basis for consideration of both desired behavior in extrapolation and uncertainty in model forms.

1 Tolerance Distributions and Link Functions

We wish to cast this discussion in the context of traditional short-term toxicity tests that have a statistical structure identical to the one used in the paper. In particular, short-term toxicity tests produce sets of binomial outcomes at different levels of a covariate, generally a toxic substance. A typical analysis would be conducted using a generalized linear model with binomial random component and a chosen link function (e.g., logit, probit, or complementary log-log, just as in the paper). In this context, there is a correspondence between the link function chosen and the assumed distribution of “tolerances” in the population of organisms being tested. The tolerance of an individual organism is defined to be that value of the covariate such that the organism will not respond (e.g., not die) for any value less than the tolerance, but will respond (e.g., will die) at any value of the covariate greater than or equal to the tolerance. As is well known, link functions for binomial random component models correspond to inverse distribution functions of standardized tolerance distributions. For example, under a (standardized) tolerance distribution $F$, the probability that an organism dies at a (standardized) exposure level $x$ is
\[ \mu = F(x) \] or, in terms of the link function, \( g(\mu) = x \) so that \( g = F^{-1} \). We may use \( F \) as the parent distribution to define a location-scale family with parameters \( \xi \) and \( \sigma \) (i.e., \( \mu = F\{ (x - \xi)/\sigma \} \)) whereby \( g(\mu) = -\xi/\sigma + x/\sigma \) is linear in the covariate; note that the location parameter \( \xi \) may represent mode rather than expectation in the resultant family of tolerance distributions. For brevity, we consider only standardized forms of tolerance distributions in this discussion. The relation between link function and tolerance distribution is given in the paper as expressions (1.1) and (1.2), so this structure also implicitly forms the starting point for the presentation of Fygenson.

2  Pessimism/Optimism and Distributions

The basic notion of pessimistic versus optimistic outlooks underlies a good deal of the approach to extrapolation taken in the paper. We consider these notions in terms of what the paper calls attributable risk (AR) because this is the most direct measure of the probability of an adverse outcome, the central concern in problems of the type considered. We do not find the contrast between pessimistic and optimistic distributions presented in the paper entirely convincing. Fygenson defines only portions of distribution functions as representing pessimistic or optimistic structures, related to convexity or concavity of the function over certain intervals of the line (Definition 2.1 and Theorem A.1). Any distribution with a continuous unimodal density on the entire line must then contain both pessimistic (convex) and optimistic (concave) portions, so that it seems odd to proclaim a distribution as one or the other. The fact that many distributions contain both types of behavior then motivates construction of a piecewise model as given by expression (3.1) of the paper. While the component distributions of model (3.1), namely \( F_0, F_{01} \) and \( F_1 \) are all restricted to be absolutely continuous with densities having at least one contin-
uous derivative (see below Definition 2.1), we did not see restrictions imposed that guarantee the entire model determines an absolutely continuous distribution function. This might not be a great issue if all one is concerned with is extrapolation from a regression function for a single value of the covariate, but it seems lacking as a conceptualization of an overall problem involving tolerance distributions (and densities).

We think of pessimism and optimism as a continuum (with no absolute zero) that may go by either name, and that distributions might be classified only in a relative manner, such as $F$ is more pessimistic (or less optimistic) than $G$. Thus, the pessimistic versus optimistic contrast defined by Fygenson would be replaced entirely by his notion of degree of pessimism (or optimism) as reflected in stochastic orderings, and detailed in Appendix A of the paper. The question then becomes whether there exist useful classes of distributions that allow stochastic ordering over the entire line, that may be used to define link functions in generalized linear models with binomial random components, and that may be estimated on the basis of observed information. We demonstrate in the next section that at least one such class of distributions can be identified.

The comparison of distributions based on stochastic ordering should be conducted only after the distributions have been “matched” in one or more aspects of their behavior. Without such matching, stochastic ordering may become a relatively uninteresting result of differing parameter values. For example, logistic distribution functions with different location parameters but the same scale parameter are simply translations along the real line. This fact is what underlies the definition of relative potency in parallel line bioassays (e.g., Finney (1978)), but is a rather trivial case of stochastic ordering as it relates to the behavior of distributions.
3 Parameterized Link Functions

An alternative approach to accounting for uncertainty in link functions (and hence also tolerance distributions) is to make use of a parameterized family of links. To our knowledge, the use of families of parameterized link functions was first suggested by Pregibon (1980) who used such families to form alternatives for testing the adequacy of a specific “hypothesized” link. This idea was also used in the analysis of toxicity tests by Aranda-Ordaz (1981). We consider only the simplest among a number of possible families of link functions for binomial responses, given by Pregibon (1980) as

\[ g(\mu|\lambda) = \log \left[ \frac{(1 - \mu)^{-\lambda} - 1}{\lambda} \right], \quad (1) \]

where \( \mu \) is the expected value of a binomial random variable (written in terms of proportions), that is, the probability of a response at level \( x \) of some covariate. In the paper, this is the combination of expressions (1.1) and (1.2). For the present we consider only \( \lambda > 0 \) in (1). The family of standardized tolerance distributions implied by this link function are given by

\[ F_\lambda(t) = 1 - \frac{1}{\{\lambda \exp(t) + 1\}^{1/\lambda}}, \quad (2) \]

with corresponding densities

\[ f_\lambda(t) = \frac{\exp(t)}{\{\lambda \exp(t) + 1\}^{1+1/\lambda}}; \quad -\infty < t < \infty. \quad (3) \]

If \( \lambda < 0 \) in (1) through (3), the effect is that the support of the density (3) is restricted to \( t < -\log(-\lambda) \) and is thus determined by the value of \( \lambda \).

In the terminology of Fygenson, the distributions (2) are pessimistic for \( t < 0 \), which is the mode, and optimistic for \( t > 0 \). But we also have the following.

Result 1.

If \( F_{\lambda_1} \) and \( F_{\lambda_2} \) are two distribution functions of the form (2) such that \( 0 < \lambda_1 < \lambda_2 \),
then $F_{\lambda_1}(t) > F_{\lambda_2}(t)$ for all $-\infty < t < \infty$.

That is, a random variable with distribution $F_{\lambda_2}$ is stochastically larger than a random variable with distribution $F_{\lambda_2}$ over its entire sample space, and both distributions essentially match in the lower portion of the real line (i.e., $\lim_{t \to -\infty} F_{\lambda_j}(t) = 0$).

Proof.

Fix $t$ and define a function $h(\lambda) \equiv \log[1 + \lambda \exp(t)], \ \lambda > 0$. It suffices to show $[1 + \lambda_2 \exp(t)]^{1/\lambda_2} < [1 + \lambda_1 \exp(t)]^{1/\lambda_1}$ or, equivalently, that $h(\lambda_2)/\lambda_2 < h(\lambda_1)/\lambda_1$.

Now, the function $h(\cdot)$ is strictly concave in $\lambda$ so that, with $\lambda_1 < \lambda_2$, the chord over $(0, \lambda_1)$ must have greater slope than the chord over $(0, \lambda_2)$, that is,

$$\frac{h(\lambda_2) - h(0)}{\lambda_2} \leq \frac{h(\lambda_1) - h(0)}{\lambda_1}.$$

Since $h(0) = 0$ the result follows.

This stochastic ordering is illustrated in Figure 1 for six different values of $\lambda$ ranging from 0.005 to 2.5. Smaller $\lambda$ correspond to steeper (i.e., leftmost) curves. What is important is that, through the link function (1) and the corresponding distributions (2), we have arrived at an ordering over the entire line so that there is no longer any need to restrict functions to particular intervals and construct a response curve as in model (3.1) of the paper. The implication is that the notions of pessimistic versus optimistic distributions as convex or concave over various intervals can be replaced with Fygenson’s degrees of pessimism (or optimism) for entire distributions. More pessimistic (or less optimistic) distribution functions pick up probability more rapidly than less pessimistic (or more optimistic) ones, as illustrated in Figure 1.

We point out that the stochastic ordering over the entire line allowed by the distributions of (2) is different than the effect of changes in “slope parameters” for
a generalized linear model with a given fixed link. As mentioned previously, the slope corresponds to a scale parameter in tolerance distributions. The effect of scale changes is illustrated for a logit link in Figure 2. Here, changes in scale produce stochastic orderings that differ in direction on either side of the mode of 0 for all of these distributions. Thus, the link function parameter $\lambda$ in (1), (2), and (3) produces a different effect in terms of Fygenson’s pessimism/optimism than does a slope (scale) change in a model with fixed link.

That more pessimistic (less optimistic) distributions accumulate probability more rapidly is clear from Figure 1, but it is illustrative to visualize this phenomenon in terms of density functions for tolerance. Density functions corresponding to the curves of Figure 1 are presented in Figure 3, where smaller values of $\lambda$ correspond...
to densities with more rapid declines to the right of the mode.

In terms of toxicity tests, or item failure as in the shuttle rocket example, the interpretation of tolerance distributions for different values of $\lambda$ relative to our modified concept of pessimism/optimism is quite intuitive. All of the distributions in Figures 1 and 3 represent populations of organisms (or test items) with about the same proportion of “highly susceptible” individuals. But more pessimistic distributions represent situations in which “total failure” occurs more rapidly after the mode, while less pessimistic distributions reflect populations with “highly resistant” individuals as well as highly susceptible ones.

Relative to the problem of extrapolation, the value of $\lambda$ will make little difference in how well data at the low end of the response range can be fitted with a model (this
Figure 3: Density functions corresponding to the distributions of Figure 1.

is the practical dilemma of Fygenson), but will give quite different predictions at the upper end of the response range. What is needed to distinguish among the possible values of \( \lambda \) are data at the low end of responses up to about the modal value. Such data provide the information necessary to determine the “steepness” of the response curve, which corresponds to the rate at which probability is accumulated in the distribution function, and hence the degree of pessimism/optimism the data would suggest. In fact, maximum likelihood estimates of link function parameters such as \( \lambda \) in (1) can be easily computed, simultaneously with all other model parameters, using an algorithm similar to that often employed with traditional generalized linear models (Kaiser (1997)).

The maximum likelihood algorithm just mentioned is a Newton-type algorithm
and so also produces the full information matrix for all model parameters, including the link function parameter \( \lambda \). As a result, one has (i) incorporated uncertainty into the response function through the unknown parameter \( \lambda \), (ii) estimated that portion of the model form on the basis of observed data, (iii) quantified uncertainty in the estimated value, and (iv) also quantified the effect of that uncertainty on uncertainty about estimates of the other model parameters. Of course, all of this is true only within the class of link functions defined but, in problems for which the response function form is of scientific interest in its own right, this approach might have considerable merit.
Additional References