An Asymptotic Confidence Region for the ED100p from the Logistic Response Surface for a Combination of Agents
Author(s): Walter H. Carter, Jr., Vernon M. Chinchilli, John D. Wilson, Eleanor D. Campbell, Fay K. Kessler, Richard A. Carchman
Published by: American Statistical Association
Stable URL: http://www.jstor.org/stable/2684870

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=astata.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Statistical Association is collaborating with JSTOR to digitize, preserve and extend access to The American Statistician.
An Asymptotic Confidence Region for the $ED_{100p}$ From the Logistic Response Surface for a Combination of Agents

WALTER H. CARTER, JR., VERNON M. CHINCHILLI, JOHN D. WILSON, ELEANOR D. CAMPBELL, FAY K. KESSLER, and RICHARD A. CARCHMAN*

1. INTRODUCTION

The $ED_{100p}$, $0 \leq p \leq 1$, in the analysis of quantal response data involving a single independent variable, is that value of the independent variable associated with a $100p\%$ response rate. For example, when the simple linear logistic regression model is used to approximate the underlying dose-response relationship, that is,

$$\ln(p/(1-p)) = \beta_0 + \beta_1 x,$$

the $ED_{50}$, the dose associated with a $50\%$ response rate, is given by the single point $x = -\beta_0/\beta_1$. When a combination of $k$ agents is being considered, however, the $ED_{100p}$ for the multiple linear logistic regression model is given by all solutions $(x_1, x_2, \ldots, x_k)$ to

$$\beta_0 - \ln(p/(1-p)) + \sum_{i=1}^{k} \beta_i x_i = 0.$$

When viewed in light of the simplicity of the solution in a single variable, it is not surprising that the estimation of the $ED_{100p}$ in combinations has received little attention in the statistical literature. As a result of the increasing interest in the effects of combinations of agents either as the result of treatment or environmental exposure, however, consideration of the estimation of the $ED_{100p}$ in combinations is warranted. [Skarin, Canellos, Rosenthal, Case, MacIntyre, Pinkus, Moloney, and Frei (1983) reported on the use of a combination of six drugs in the treatment of lymphoma, and Lang, Kurzepa, Cole, and Loper (1980) indicated that several known or suspected carcinogens were identified among the many compounds present in drinking water.]

Confidence bands on the logistic response curve were considered by Hauck (1983) and Brand, Pinnock, and Jackson (1973). One of the advantages claimed by Hauck is that his method is not restricted to a single explanatory variable, as is the approach taken by Brand et al. In their paper, however, Brand et al. discussed placing confidence bands about points on the inverse response curve, for example, the $ED_{100p}$, a topic not discussed by Hauck. It is the purpose of this article to develop and illustrate a method for estimating a large sample confidence region about the $ED_{100p}$ from the logistic curve in the case of multiple explanatory variables.

2. BACKGROUND

The logistic response curve is useful in relating a set of independent variables $x$ to values of a dichotomous variable $Z$ that takes on values of 0 or 1. For reasons that will become apparent in Section 4, the model will be written as

$$\ln[Pr(Z = 1|x)/Pr(Z = 0|x)] = \beta_0 + x'B + x'Bx,$$

(2.1)

where

$$x = [x_1, x_2, \ldots, x_k]^T,$$

$$\beta = [\beta_1, \beta_2, \ldots, \beta_k]^T,$$

and

$$B = \begin{bmatrix} \beta_{11} & \beta_{12}/2 & \cdots & \beta_{1k}/2 \\ \beta_{12}/2 & \beta_{22} & \cdots & \beta_{2k}/2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{1k}/2 & \beta_{2k}/2 & \cdots & \beta_{kk} \end{bmatrix}.$$ 

This provides a quadratic response surface for the logistic regression model. It should be emphasized, however, that the method to be developed here is not dependent on the presence of the quadratic terms in the model. Further, let $\beta^*$ be the collection of all of the parameters, that is,

$$\beta^* = [\beta_0, \beta_1, \ldots, \beta_k, \beta_{11}, \beta_{12}, \ldots, \beta_{kk}]^T,$$

and let

$$x^* = [1, x_1, x_2, \ldots, x_k, x_1^2, x_2^2, \ldots, x_k^2]^T.$$ 

Both $\beta*$ and $x*$ are vectors with $k^* = (k + 1)(k + 2)/2$ components.

Notice that (2.1) can be written as

$$\ln[Pr(Z = 1|x)/Pr(Z = 0|x)] = x^T*\beta^*.$$ 

For a sample of size $N$, the maximum likelihood principle can be invoked to estimate $\beta^*$ and to show that

$$N^{1/2}(\hat{\beta}^* - \beta^*) \xrightarrow{d} \mathcal{N}_{k^*}(\mathbf{0}, \Sigma),$$

where $\mathcal{N}_{k^*}(\mathbf{0}, \Sigma)$ denotes convergence in distribution and $\mathcal{N}_{k^*}(\mathbf{0}, \Sigma)$ denotes the $k^*$-dimensional multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\Sigma$. $\Sigma$ is consistently estimated by $N$ times the inverse of the sample information matrix, that is, $\Sigma = NJ^{-1}$, where $J$ is the sample information matrix.

*All of the authors are with the Medical College of Virginia/Virginia Commonwealth University, P.O. Box 32, MCV Station, Richmond, VA 23298. Walter H. Carter, Jr. and Richard A. Carchman are Professors and Vernon M. Chinchilli is Associate Professor in the Departments of Biostatistics and Pharmacology and Toxicology; John D. Wilson is Associate Professor in the Department of Radiology; Eleanor D. Campbell is Instructor in the Department of Biostatistics; and Fay K. Kessler is Research Assistant in the Department of Pharmacology and Toxicology. This work was supported in part by National Institute of Environmental Health Sciences Grants 1RO1ES02992 and 1NO1ES15001. Richard A. Carchman is the recipient of National Institutes of Health Research Career Development Award 5K04AM00565. Reprint requests may be addressed to Walter H. Carter, Jr.
3. CALCULATION OF THE LARGE SAMPLE CONFIDENCE REGION ABOUT THE INVERSE LOGISTIC FUNCTION

Using the Cauchy–Schwarz inequality in a manner similar to that of Hauck, it can be shown that

\[ \sup_{x \in R^k} N(x^* (\hat{\beta} - \beta^*))^2 \leq \sup_{y \in R^{k'}} N(y^* (\hat{\beta} - \beta^*))^2 = N(\hat{\beta}^* - \beta^*)^2 = X_{k',1-\alpha}^2. \]

This leads to the following conservative probability statement:

\[ \Pr[N(x^* (\hat{\beta} - \beta^*))^2 \leq X_{k',1-\alpha}^2 \text{ simultaneously for all } x \in R^k] = 1 - \alpha. \]

This can be manipulated to yield

\[ \Pr[\hat{\alpha}^* (\hat{\beta} - \beta^*) \leq x^* X_{k',1-\alpha}^2 \text{ simultaneously for all } x \in R^k] = 1 - \alpha. \]

In what follows, we assume that \( N \) is sufficiently large so that we can replace \( \Sigma \) by its consistent estimator. For a particular value of \( x \), let

\[ \hat{\beta}_L(x) = [1 + \exp(-x^* \hat{\beta}^* + (x^* J^{-1} x^* X_{k',1-\alpha}^2)]^{-1} \]

and

\[ \hat{\beta}_U(x) = [1 + \exp(-x^* \hat{\beta}^* - (x^* J^{-1} x^* X_{k',1-\alpha}^2)]^{-1}. \]

This gives us Hauck's (1983) conservative confidence interval for our model:

\[ \Pr[\hat{\alpha}_L(x) \leq 1 + \exp(-x^* \hat{\beta}^*)^{-1} \leq \hat{\beta}_U(x), \text{ simultaneously for all } x \in R^k] = 1 - \alpha. \]

The \( ED_{1000} \) set is defined as the set of all \( x \in R^k \) such that \( p = [1 + \exp(-x^* \hat{\beta}^*])^{-1} \), that is,

\[ S_p(\beta^*) = \{ x \in R^k : p = [1 + \exp(-x^* \hat{\beta}^*)]^{-1}. \]

In general, \( S_p(\beta^*) \) defines the surface of a hyperparaboloid or hyperellipsoid in \( k \)-dimensional space. For the special case of \( k = 1 \) and our quadratic model, \( S_p(\beta^*) \) may consist of zero, one, or two points, depending on the values of the model parameters.

Obviously, \( S_p(\hat{\beta}^*) \) estimates \( S_p(\beta^*) \). By using the definition of \( S_p(\hat{\beta}^*) \) and inverting Hauck's (1983) result, an approximate \( 100(1 - \alpha)\% \) conservative confidence region for \( S_p(\hat{\beta}) \) is defined by

\[ G_p(\hat{\beta}) = \{ x \in R^k : \hat{\beta}_L(x) \leq [1 + \exp(-x^* \hat{\beta}^*)]^{-1} \leq \hat{\beta}_U(x) \text{ for all } x \in S_p(\hat{\beta}^*) \}. \]

4. EXAMPLE

The technique presented in the previous section will be illustrated using cytotoxicity data from a combination of methylmethanesulfonate (MMS) and phorbol 12-myristate, 13-acetate (PMA) in the human promyelocytic leukemia cell line HL-60. MMS is a cytotoxic agent that can alkylate nucleophiles (e.g., DNA) and has demonstrable activity as a mutagen/carcinogen (Lawley 1976). PMA is a member of the family of dipterine tumor promoters that have been shown to have specific effects in a variety of systems (Blumberg 1980), including HL-60 cells (Koeffler and Golde 1980). Based on the current understanding of the dose–response (HL-60 cell cytotoxicity) curve for each individual agent, it was of interest to evaluate the dose–response surface for their combination. Hence the study was performed. HL-60 cell cytotoxicity was evaluated by determining the number of viable and dead cells, using the ability of viable cells to exclude the vital dye trypan blue. The concentrations of the two chemicals and the corresponding responses are given in Table 1.

Table 1. Treatment Combinations and the Observed and Expected Responses

<table>
<thead>
<tr>
<th>MMS (µg/ml)</th>
<th>PMA (M x 10^{-9})</th>
<th>No. viable cells</th>
<th>No. dead cells</th>
<th>Expected no. dead cells from analysis 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>79</td>
<td>19</td>
<td>20.5</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>63</td>
<td>87</td>
<td>20.6</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>37</td>
<td>84</td>
<td>48.5</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>19</td>
<td>68</td>
<td>71.2</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>75</td>
<td>16</td>
<td>17.9</td>
</tr>
<tr>
<td>250</td>
<td>0</td>
<td>73</td>
<td>17</td>
<td>13.0</td>
</tr>
<tr>
<td>250</td>
<td>1</td>
<td>75</td>
<td>19</td>
<td>24.9</td>
</tr>
<tr>
<td>250</td>
<td>10</td>
<td>69</td>
<td>19</td>
<td>20.9</td>
</tr>
<tr>
<td>250</td>
<td>100</td>
<td>14</td>
<td>79</td>
<td>75.0</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>10</td>
<td>73</td>
<td>13.7</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>36</td>
<td>41</td>
<td>35.8</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>21</td>
<td>62</td>
<td>61.6</td>
</tr>
<tr>
<td>250</td>
<td>1</td>
<td>1</td>
<td>36</td>
<td>27.9</td>
</tr>
<tr>
<td>250</td>
<td>10</td>
<td>56</td>
<td>56</td>
<td>57.2</td>
</tr>
<tr>
<td>250</td>
<td>100</td>
<td>10</td>
<td>74</td>
<td>75.1</td>
</tr>
</tbody>
</table>

Table 2. Parameter Estimates and Their Significance

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate Analysis 1</th>
<th>Estimate Analysis 2</th>
<th>p Value for test of ( H_0 ): parameter = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>-1.343</td>
<td>-1.330</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-.084</td>
<td>-.084</td>
<td>.0024</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>.160</td>
<td>.159</td>
<td>.0001</td>
</tr>
<tr>
<td>( \beta_{11} )</td>
<td>3.865 x 10^{-3}</td>
<td>3.883 x 10^{-3}</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>( \beta_{22} )</td>
<td>-1.309 x 10^{-3}</td>
<td>-1.308 x 10^{-3}</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>( \beta_{12} )</td>
<td>-6.273 x 10^{-4}</td>
<td>—</td>
<td>.0754</td>
</tr>
</tbody>
</table>

The American Statistician, May 1986, Vol. 40, No. 2 125
portions of dead cells than a less toxic treatment. To account for such curvature in the dose–response relationship, the following model was fit to the data by using PROC LOGIST in SAS (1983):

\[
\ln(p/1-p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2, \]

where \( p \) = proportion of dead cells, \( x_1 \) = concentration of MMS/10, and \( x_2 \) = concentration of PMA.

The estimated coefficients and their significance are given in Table 2. Since \( \beta_{12} \), the coefficient of the cross-product term in the model, cannot be shown to differ significantly from zero, it was deleted from the model and the data were reanalyzed. The results of the second analysis also appear in Table 2. The estimated large sample covariance matrix associated with the vector of parameter estimates is given in Figure 1. The predicted, or expected, number of dead cells for each treatment combination is given in the last column of Table 1.

Although the logistic function provides a reasonable model for the expected number of dead cells, common experimental errors—for example, pipetting errors—could possibly cause the variance to exceed \( np(1-p) \) as assumed. Such an extra component of variance, if it existed, would invalidate our confidence region, although the point estimate of \( \beta \) would still be appropriate. Fortunately, the chi-squared test of fit not only tests the adequacy of the quadratic model for the logit but also tests for an extra variance component. Each summand of the chi-squared statistic is (observed−expected)\(^2\), standardized by the expected value of this quantity under the null hypothesis of only binomial variability. Thus an extra variance component would lead to a large value of the test statistic associated with a test of model adequacy. For example the test of fit yields a value of \( X^2 = 13.94 \) with 11 df. The associated \( p \) value is .24. Thus there is no indication of model inadequacy.

A plot of the estimated contours of constant response for the fitted response surface appears in Figure 2. For the values of \( p \) used in generating this figure the appropriate contour(s), restricted to the experimental region, is the \( ED_{100p} \) set. From this figure it can be seen that for fixed levels of PMA, MMS has little effect on cell viability and this is true across levels of PMA, which is a graphical indication of no interaction between the two agents. Conservative 95% confidence regions for the \( ED_{100p} \), \( p = .4, .5, \) and .6, are given in Figures 3, 4, and 5, respectively. Since these three confidence regions overlap, we conclude that we cannot confidently distinguish among the respective \( ED_{100p} \) sets, even though there is a highly significant dose–response relationship with both variables.
5. CONCLUSION

Various statistical procedures used in the risk assessment of a single agent make use of confidence limits on points of the inverse response curve. The methodology discussed in this article, therefore, would be useful in the risk assessment of combinations of agents—a topic that is yet to receive attention in the statistical literature.

Although this article has dealt exclusively with a model useful in the analysis of binary data, the procedure proposed is general and could also be applied in the situation considered in the analysis of response surfaces with a continuous dependent variable. There it is quite common to consider estimated contours of constant response without an indication of the variability associated with these estimates. With a straightforward modification of the method presented here, it would be possible to estimate the confidence regions associated with these contours.

[Received October 1983. Revised February 1985.]
Figure 5. Confidence Region About the ED_{60} for the MMS–PMA Combination. (The MMS axis is the MMS concentration divided by 10.)

REFERENCES
Blumberg, P. M. (1980), "In Vitro Studies on the Mode of Action of the Phorbol Esters, Potent Tumor Promoters" (Parts 1 and 2), Critical Reviews in Toxicology, 8, 153–234.