Chapter 575

Probit Analysis

Introduction

Probit Analysis is a method of analyzing the relationship between a stimulus (dose) and the quantal (all or nothing) response. Quantitative responses are almost always preferred, but in many situations they are not practical. In these cases, it is only possible to determine if a certain response (such as death) has occurred. In a typical quantal response experiment, groups of animals are given different doses of a drug. The percent dying at each dose level is recorded. These data may then be analyzed using Probit Analysis.

The Probit Model assumes that the percent response is related to the log dose as the cumulative normal distribution. That is, the log doses may be used as variables to read the percent dying from the cumulative normal. Using the normal distribution, rather than other probability distributions, influences the predicted response rate at the high and low ends of possible doses, but has little influence near the middle. Hence, much of the comparison of different drugs is done using response rates of fifty percent. The probit model may be expressed mathematically as follows:

\[ P = \alpha + \Phi[\log_{10}(Dose)] \]

where \( P \) is five plus the inverse normal transform of the response rate (called the Probit). The five is added to reduce the possibility of negative probits, a situation that caused confusion when solving the problem by hand.

The popularity of the method is due in large part to the work of Finney (1971), in his book Probit Analysis. He explains the proper use and analysis of quantal response data. In NCSS, we have coded the algorithms given in his book, and we refer you to it for further information and background.

Data Structure

The data below are suitable for analysis by this procedure. Note that the first variable, Dose, gives the dose level of the treatment. The second variable, Subjects, gives the number of individuals receiving a specific dose level. The third variable, Response, gives the number of treated individuals who exhibited the response of interest.

These data are contained on the Survival dataset.

<table>
<thead>
<tr>
<th>Dose</th>
<th>Subjects</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>102</td>
<td>19</td>
</tr>
<tr>
<td>60</td>
<td>121</td>
<td>26</td>
</tr>
<tr>
<td>70</td>
<td>111</td>
<td>24</td>
</tr>
<tr>
<td>80</td>
<td>105</td>
<td>31</td>
</tr>
<tr>
<td>90</td>
<td>117</td>
<td>54</td>
</tr>
<tr>
<td>100</td>
<td>108</td>
<td>83</td>
</tr>
</tbody>
</table>
Procedure Options
This section describes the options available in this procedure.

Variables Tab
This panel specifies the variables used in the analysis.

Count Variable
R: Count Variable
This variable contains the number of individuals with the desired response. It must be less than the number of animals. The analysis adds one-half to zero and subtracts one-half if the R = N. This slight modification avoids division by zero in the calculations.

Sample Size Variable
N: Sample Size Variable
This is the variable containing the total number of individuals sampled at a particular dose level.

Dose Variable
X: Dose Variable
This option contains the name of the variable containing the dose levels. Note that the analysis uses the log (base 10) transformation of dose levels.

Group Variable
Group Variable
An optional categorical (grouping) variable may be specified. If it is used, a separate analysis is conducted for each unique value of this variable.

Reports Tab
The following options control the display of reports and plots.

Select Reports
Probit Estimation Report ... Dose Percentiles Report
These options specify whether to display the corresponding report.

Percentiles
A separate row in the Dose Percentile report is created for each percentage value given here. This is a list of numbers between 0 and 100 separated by blanks or commas.
Report Options

Precision
Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

Variable Names
This option lets you select whether to display only variable names, variable labels, or both.

Value Labels
This option lets you select whether to display only values, value labels, or both. Use this option if you want to automatically attach labels to the values of the group variable (like 1=Yes, 2=No, etc.). See the section on specifying Value Labels elsewhere in this manual.

Plots Tab
These options control the attributes of the corresponding plots.

Select Plots

Dose - Response Plot ... Probit Plot
These options specify whether to display the corresponding plot. Click the plot format button to change the plot settings.

Example 1 – Probit Analysis
This section presents an example of how to perform a probit analysis using the data that were shown earlier and found in the Survival dataset.

You may follow along here by making the appropriate entries or load the completed template Example 1 by clicking on Open Example Template from the File menu of the Probit Analysis window.

1 Open the Survival dataset.
   • From the File menu of the NCSS Data window, select Open Example Data.
   • Click on the file Survival.NCSS.
   • Click Open.

2 Open the Probit Analysis window.
   • Using the Analysis menu or the Procedure Navigator, find and select the Probit Analysis procedure.
   • On the menus, select File, then New Template. This will fill the procedure with the default template.

3 Specify the variables.
   • On the Probit Analysis window, select the Variables tab.
   • Double-click in the R: Count Variable box. This will bring up the variable selection window.
   • Select Response from the list of variables and then click Ok.
   • Double-click in the X: Dose Variable box. This will bring up the variable selection window.
   • Select Dose from the list of variables and then click Ok.
   • Double-click in the N: Sample Size Variable box. This will bring up the variable selection window.
   • Select Subjects from the list of variables and then click Ok.
4 Run the procedure.
   • From the Run menu, select Run Procedure. Alternatively, just click the green Run button.

Probit Estimation Section

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>-4.545974</td>
<td>1.032341</td>
</tr>
<tr>
<td>Beta</td>
<td>4.901165</td>
<td>0.5483724</td>
</tr>
<tr>
<td>LD50</td>
<td>1.947695</td>
<td>1.304145E-02</td>
</tr>
<tr>
<td>Dose50</td>
<td>88.65325</td>
<td>2.662173</td>
</tr>
</tbody>
</table>

**Alpha**
The estimated value of the intercept, with its associated standard error.

**Beta**
The estimated value of the slope, with its associated standard error.

**LD50**
The estimated value, on the log10(dose) scale, at which 50% responded.

**Dose50**
The estimated value, on the dose scale, at which 50% responded.

Probit Detail Section

<table>
<thead>
<tr>
<th>Dose</th>
<th>Actual Percent</th>
<th>Probit Percent</th>
<th>N</th>
<th>R</th>
<th>E(R)</th>
<th>Difference</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>18.63</td>
<td>11.14</td>
<td>102</td>
<td>19.00</td>
<td>11.36</td>
<td>7.64</td>
<td>5.77</td>
</tr>
<tr>
<td>60</td>
<td>21.49</td>
<td>20.30</td>
<td>121</td>
<td>26.00</td>
<td>24.56</td>
<td>1.44</td>
<td>0.11</td>
</tr>
<tr>
<td>70</td>
<td>21.62</td>
<td>30.75</td>
<td>111</td>
<td>24.00</td>
<td>34.14</td>
<td>-10.14</td>
<td>4.35</td>
</tr>
<tr>
<td>80</td>
<td>29.52</td>
<td>41.35</td>
<td>105</td>
<td>31.00</td>
<td>43.41</td>
<td>-12.41</td>
<td>6.05</td>
</tr>
<tr>
<td>90</td>
<td>46.15</td>
<td>51.28</td>
<td>117</td>
<td>54.00</td>
<td>60.00</td>
<td>-6.00</td>
<td>1.23</td>
</tr>
<tr>
<td>100</td>
<td>76.85</td>
<td>60.12</td>
<td>108</td>
<td>83.00</td>
<td>64.93</td>
<td>18.07</td>
<td>12.62</td>
</tr>
</tbody>
</table>

**Dose**
The dose level.

**Actual Percent**
The ratio of the count to the sample size (R/N).

**Probit Percent**
The estimated ratio (R/N) based on the probit model.

**N**
The sample size.
**Probit Analysis**

**R**
The count (number responding).

**E(R)**
The expected count based on the probit model.

**Difference**
The difference between the actual and the expected counts.

**Chi-Square**
The Chi-Square statistic for testing the significance (non-zero) of the difference. Since these are single degree of freedom tests, the value should be greater than 3.81 to be significant at the 0.05 level.

**Total Chi-Square**
The total of the Chi-Square values, used to test the overall significance of the differences from the model.

**D.F.**
The degrees of freedom of the Chi-Square test.

**Prob Level**
The probability to the right of the above Chi-Square value. The significance level of the Total Chi-Square test.

### Dose Percentile Section

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Probit</th>
<th>Log(Dose)</th>
<th>Std. Error Log(Dose)</th>
<th>Dose</th>
<th>Std. Error Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.6737</td>
<td>1.4730</td>
<td>0.0468</td>
<td>29.7196</td>
<td>3.2008</td>
</tr>
<tr>
<td>5</td>
<td>3.3551</td>
<td>1.6121</td>
<td>0.0318</td>
<td>40.9346</td>
<td>2.9993</td>
</tr>
<tr>
<td>10</td>
<td>3.7184</td>
<td>1.6862</td>
<td>0.0242</td>
<td>48.5530</td>
<td>2.7013</td>
</tr>
<tr>
<td>20</td>
<td>4.1584</td>
<td>1.7760</td>
<td>0.0158</td>
<td>59.7002</td>
<td>2.1685</td>
</tr>
<tr>
<td>25</td>
<td>4.3255</td>
<td>1.8101</td>
<td>0.0132</td>
<td>64.5768</td>
<td>1.9640</td>
</tr>
<tr>
<td>30</td>
<td>4.4756</td>
<td>1.8407</td>
<td>0.0115</td>
<td>69.2946</td>
<td>1.8364</td>
</tr>
<tr>
<td>40</td>
<td>4.7467</td>
<td>1.8960</td>
<td>0.0108</td>
<td>78.7052</td>
<td>1.9529</td>
</tr>
<tr>
<td>50</td>
<td>5.0000</td>
<td>1.9477</td>
<td>0.0130</td>
<td>88.6533</td>
<td>2.6622</td>
</tr>
<tr>
<td>60</td>
<td>5.2533</td>
<td>1.9994</td>
<td>0.0171</td>
<td>99.8587</td>
<td>3.9219</td>
</tr>
<tr>
<td>70</td>
<td>5.5244</td>
<td>2.0547</td>
<td>0.0222</td>
<td>113.4200</td>
<td>5.8064</td>
</tr>
<tr>
<td>75</td>
<td>5.6745</td>
<td>2.0853</td>
<td>0.0253</td>
<td>121.7063</td>
<td>7.0888</td>
</tr>
<tr>
<td>80</td>
<td>5.8416</td>
<td>2.1194</td>
<td>0.0288</td>
<td>131.6477</td>
<td>8.7309</td>
</tr>
<tr>
<td>90</td>
<td>6.2816</td>
<td>2.2092</td>
<td>0.0383</td>
<td>161.8727</td>
<td>14.2814</td>
</tr>
<tr>
<td>95</td>
<td>6.6449</td>
<td>2.2833</td>
<td>0.0463</td>
<td>191.9991</td>
<td>20.4873</td>
</tr>
<tr>
<td>99</td>
<td>7.3263</td>
<td>2.4223</td>
<td>0.0616</td>
<td>264.4519</td>
<td>37.5022</td>
</tr>
</tbody>
</table>

This report displays the dose levels yielding various predicted response rates.

**Percentile**
The response rate times 100.

**Probit**
The normal transform of the percentage plus five. (The five is added to avoid the possibility of a negative probit. This practice was helpful when calculations were done by hand, but is based solely on tradition now that calculations are carried out by computer.)

**Log Dose**
The logarithm of the dose level (base 10).

**Std. Error Log(Dose)**
The standard error of the estimated log dose level.
Dose
The dose level.

Std. Error Dose
The standard error of the estimated dose level.

---

**Dose-Response Plot**

This plot lets you look at the relationship between percent response and dose. Usually, this plot will be nonlinear.

---

**Log(Dose) - Response Plot**

This plot lets you look at the relationship between percent response and log dose. Usually, this plot will be nonlinear.
This plot presents the probit model. If the probit model is to be a good approximation, this plot should show a linear relationship. Obviously, in this example, the relationship is quadratic, indicating that the probit model should be modified—perhaps by using the square of Log dose.