The Versatility of Logit Over Probit Regression Analyses Estimating the Strength of Gear Teeth

By:

A.A. Shariff and A. Zaharim

(Paper presented at the ***13th WSEAS International Conference on Applied Mathematics*** held on 15-17 December 2008 in Canary Islands, Spain)

The Versatility of Logit Over Probit Regression Analyses in Estimating the Strength of Gear Teeth

A. A. SHARIFF [1] AND A. ZAHARIM [2]

[1] Centre For Foundation Studies In Science
University of Malaya
50603 Kuala Lumpur
MALAYSIA
asma@um.edu.my

[2] Faculty of Engineering
University Kebangsaan Malaysia
Bangi, Selangor
MALAYSIA

Abstract: Logit and probit are two regression methods which are categorised under Generalized Linear Models. Both models can be used when the response variables in the analyses are categorical in nature. For the case of the strength of gear teeth data, it can be in terms of counted proportions, such as r teeth fail out of n teeth tested. In this paper, the two models, logit and probit are discussed and the methods of analysis are compared for simulated data sets obtained from experimental procedure called staircase design (SCD) experiment. For the analysis, the response variable is the proportion failing and the explanatory variable is the corresponding load. The analysis is also compared with the explanatory variable of logarithm of load. The population distributions of strengths considered are normal and Weibull distribution and 1000 SCD experiments are simulated. The sampling distributions of the various estimators are then compared for bias, standard deviation, and mean squared error for the two contrasting population distributions of strength. It is found that, a regression of the logit on the logarithm of load seems to be the most robust approach if normality of strengths is in doubt.

Key-words: logit, probit, regression analysis, counted proportion, gear teeth, staircase design.

1 Introduction

For ordinary linear regression, the response variable is always quantitative and continuous in nature. When the response variables are categorical and in particular binary, that is, it can assume only two values (a 'yes-no' or 'fail-survive') or in terms of counted proportions (r fail out of n tested) we are led to consider some other models which are more appropriate than ordinary linear regression. An important characteristic of data in which the response variables are binary is that the response variables must lie between 0 and 1. Therefore fitting these data using ordinary linear regression can give prediction for the proportion of above one or less than zero, which would be meaningless. On the other hand what we actually need in this situation is a regression model which will predict the proportion of occurrences, p (let us call them p instead of y) at certain levels of x.

For this type of data, in particular, when the response variables are in terms of counted proportions, the relationship between response variables p and explanatory variables x is a non-linear curved relationship (S-shaped) curve which is usually called ***sigmoid***. The S-shaped behaviour is very common in modeling binomial responses as a function of predictors and also makes use of the assumption that the responses are from underlying binomial or binary distribution. The purpose of logistic modelling (and also probit) is the same as other modelling techniques used in statistics, that is, to find a model that fits the data best and is the simplest, yet physically reasonable in describing the relationship between the response and the explanatory variables [1]. There is little to distinguish between logit and probit models. Both curves are so similar as to yield essentially identical results. It was found that probit and logit analysis applied to the same set of data produce coefficient estimates which differ approximately by a factor of proportionality, and that factor should be about 1.8 [2].

2 Materials and Methods

2.1 Logit Versus Probit Regression Techniques

Logit model can be presented as

$$p = \frac{\exp(Z)}{1+\exp(Z)} \qquad (1)$$

where p is the proportion of occurrences, $Z = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k$ and $x_1 \ldots x_k$ are the explanatory variables. The inverse relation of equation (1) is

$$Z = \ln\left(\frac{p}{1-p}\right) \qquad (2)$$

that is, the natural logarithm of the odds ratio, known as the logit. It transforms p which is restricted to the range [0, 1] to a range $[-\infty, \infty]$.

Probit regression analysis involves modeling the response function with the normal cumulative distribution function. The probit of a proportion p is just the point on a normal curve with mean 0 and standard deviation 1 which has this proportion to the left of it. The model can be presented as

$$\Phi^{-1}(p) = Z = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k \qquad (3)$$

where p is the proportion and Φ^{-1} is the inverse of the cumulative distribution function of the standard normal distribution. That is,

$$p = \Phi(Z) = \int_{-\infty}^{Z} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)\,du \qquad (4)$$

is the cumulative distribution function of the standard normal distribution.

For logistic and probit regression, the binomial, rather than the normal distribution describes the distribution of the errors and will be the statistic upon which the analysis is based. The principles that are used for ordinary linear regression analysis could be adapted to fit both regressions. However, instead of using least square method to fit the model, for logistic and probit regressions, it is more appropriate to use maximum likelihood estimate. The likelihood function is given as

$$L = \prod_{i=1}^{m} p_i^{r_i} (1-p_i)^{n_i - r_i},$$

where the p_i are defined in terms of the parameters $\beta_0, \ldots, \beta_k$ and the known values of the predictor variables. This has to be maximized with respect to the parameters.

2.2 Experimental Design

Gear teeth are commonly tested by applying oscillatory loads, using a special machine called pulsator-test machine. In the experiment, the test specimen, in this case the gear tooth is subjected to vibrations of a resonant spring/mass system. When this happens, it experiences stresses and crack propagation takes place. Eventually, after certain number of cycles the tooth fails. The number of cycles to failure can then be recorded. If the tooth does not fail after a certain fixed number of cycles, it is considered to have survived in the experiment. The experimental procedure used is the well-known staircase design (SCD). SCD experiment is also known as sensitivity testing or 'up-and-down' method [3] where the testing of specimens is made close to the anticipated mean level. In the experiment the first test piece should be tested at a load level assumed to be near the mean value of the fatigue strength. If failure occurs before N cycles, the next test piece is tested at one step, a fixed change in load, below the first load level. Otherwise, the next test at the load one step above the first level. This procedure is continued until all the pieces have been tested. The increment between load levels should be equal for steps up and down and should be approximately one standard deviation of the fatigue strength distribution. Since the data obtained are categorical in nature, particularly in terms of counted proportions, fatigue strength of a gear is then determined by analysing the data obtained using appropriate statistical techniques, in this case logit and probit.

2.3 Analysis For SCD

The results obtained in the experiment are then analysed using logit and compared with probit. The logit transformation of p is defined by $\ln(\frac{p}{1-p})$, and the lines

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

are fitted, using maximum likelihood. A comparison is made with fitting the line,

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \ln x$$

which is equivalent to assuming a log-normal distribution of strengths.

For probit, results of SCD experiment are analysed by fitting the line

$$\Phi^{-1}(p) = \beta_0 + \beta_1 x$$

where p is the proportion failing and x is the corresponding load. This is equivalent to assuming a normal distribution of strengths. Then a comparison is made with

ISSN: 1790-2769

ISBN: 978-960-474-034-5

References:

[1] Hosmer, D.W., & Lemeshow, S. (1989). *Applied Logistic Regression*. Wiley Series in Probability and Mathematical Science. Wiley-Interscience Publication.

[2] Aldrich, J.H. & Nelson, F.D. (1984). *Linear Probability, Logit and Probit Models*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-045. Beverly Hills and London: Sage Pubns.

[3] Lloyd, D. K., & Lipow, M. (1989). *Reliability: managements, methods, and mathematics* (Second ed.). American Society for Quality Control.

[4] ISO/CD 12107. (1997). Draft for Public Comment, *Metallic Materials - Fatigue Testing - Statistical Planning and Analysis of Data*, British Standard Institution.

[5] Weibull, W. (1961). *Fatigue Testing and the Analysis of Results*. Pergamon Press. Oxford.

[6] Crowder, M.J., Kimber, A.C., Smith, R.L. & Sweeting, T.J. (1991). *Statistical Analysis of Reliability Data*. Chapman and Hall.

Table 1: Results of Staircase Experiments Analysed by Probit and Logit Regression Techniques with the load on a linear scale for Each Distribution.

Distribution	ΔL	Statistics	Probit	Logit
Normal μ=20.0 σ=2.0 $x_{0.99}$=15.34	2	$\bar{\hat{\mu}}$	20.01 (0.39) [0.39]	20.02 (0.40) [0.40]
		$\bar{\hat{\sigma}}$	1.90 (0.53) [0.54]	2.24 (0.65) [0.69]
		$\bar{\hat{x}}_{0.99}$	15.59 (1.33) [1.35]	14.80 (1.60) [1.69]
Weibull (c=2) μ=20.0 σ=10.46 $x_{0.99}$= 2.26	6	$\bar{\hat{\mu}}$	19.42 (1.98) [2.06]	19.39 (2.00) [2.09]
		$\bar{\hat{\sigma}}$	9.45 (3.16) [3.32]	11.33 (3.93) [4.03]
		$\bar{\hat{x}}_{0.99}$	-2.60 (7.37) [8.82]	-7.00 (9.13) [13.0]

( ) - standard deviation, [ ] - root mean square error (RMSE).

Table 2: Results of Staircase Experiments Analysed by Probit and Logit Regression Techniques with the load on a logarithmic scale for Each Distribution

Distribution	ΔL	Statistics	Probit (ln scale)	Logit (ln scale)
Normal μ=20.0 σ=2.0 $x_{0.99}$=15.34	2	$\bar{\hat{\mu}}$	19.97 (0.40) [0.40]	19.97 (0.40) [0.40]
		$\bar{\hat{\sigma}}$	1.69 (0.42) [0.52]	1.95 (0.49) [0.49]
		$\bar{\hat{x}}_{0.99}$	16.02 (1.10) [1.29]	15.42 (1.25) [1.25]
Weibull (c=2) μ=20.0 σ=10.46 $x_{0.99}$=2.26	6	$\bar{\hat{\mu}}$	18.53 (2.05) [2.52]	18.58 (2.06) [2.50]
		$\bar{\hat{\sigma}}$	5.35 (0.95) [5.20]	5.82 (0.98) [4.74]
		$\bar{\hat{x}}_{0.99}$	6.07 (2.44) [4.52]	5.01 (2.38) [3.64]

( ) - standard deviation, [ ] - root mean square error (RMSE).

ISSN: 1790-2769

ISBN: 978-960-474-034-5

$$\Phi^{-1}(p) = \beta_0 + \beta_1 \ln x$$

The estimated mean fatigue strength, $\hat{\mu}$, and the lower 1% point of the distribution of fatigue strength, $\hat{x}_{0.99}$, are the values of x corresponding to $p = 0.5$ and $p = 0.01$ respectively. The standard deviation can be estimated from $\hat{\sigma} = (\hat{\mu} - \hat{x}_{0.99}) / 2.33$.

The methods of analysis have been compared for simulated data sets. The population distribution of strengths is specified and 1000 SCD experiments are simulated. The sampling distributions of the various estimators can thus be compared for bias, standard deviation, and mean squared error. Two contrasting population distributions of strength are considered:

(i) normal distribution with mean of 20.0 and standard deviation of 2.0;
(ii) Weibull distribution [4 – 6], which has a cumulative distribution function F(x) defined by

$$\Pr(X < x) = F(x) = 1 - \exp[-(x / b)^c]),$$

with shape parameter, c = 2, and scale parameter, b = 22.56. These parameter values correspond to a mean of 20 and standard deviation of 10.45. The probability density functions of both distributions are plotted in Figure 1. The Weibull distribution has a substantial area near zero. This might be realistic for the strength of a component being tested under extreme conditions, as in an accelerated testing programme. It could also be interpreted as strength above some minimum value.

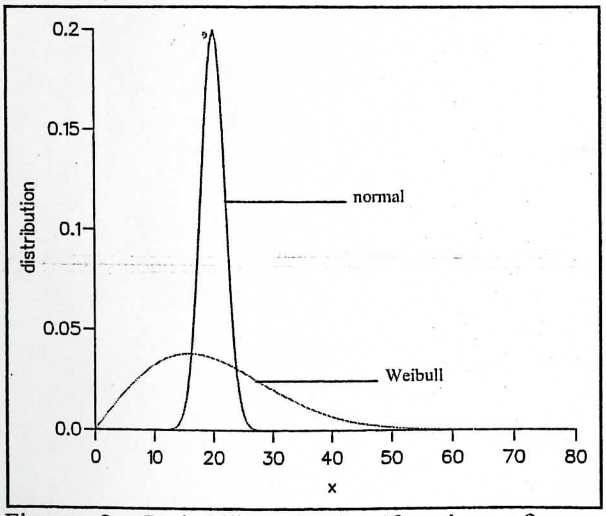

Figure 2: Probability density function of normal ($\mu = 20$, $\sigma = 2.0$) and Weibull ($\mu = 20$, $\sigma = 10.45$) distribution

In the simulation, each SCD used 50 test specimens. There are 1000 independent SCD experiments within a simulation. Each specimen put on test is randomly selected from the 50 specimens. For the normal distribution the load increment is chosen to be 2, while for the Weibull distribution it is chosen to be 6, since the standard deviation for this distribution is larger.
Results obtained from the above experiments are analysed using logit and then compared with probit analyses for each distribution. The means and standard deviations of the estimated mean, standard deviation and lower 1% point of the strength distribution are computed from 1000 SCD experiments for each distribution.

These results are tabulated in Table 1; the standard deviations of these statistics are shown in brackets, and the root mean square error (RMSE) is also calculated using the formula

$$\text{RMSE} = \sqrt{(\text{standard deviation})^2 + (\text{bias})^2}$$

where, bias = (actual value - mean of estimated value). These RMSE values are presented in square brackets. Table 2 shows results for the logit and probit analysis using a logarithmic scale for load.

3 Results and Discussion

Table 1 indicates that for load on linear scale, results obtained by probit analysis are more realistic with less error as compared to logit for normal distribution of strength. Logit analysis appears to overestimate the standard deviation and hence underestimate the lower one percent point of the distribution. However, for the Weibull distribution, which has a large standard deviation, both methods predict negative lower 1% points which are physically impossible.

Regressing the sample probit against the logarithm of load (refer to Table 2) gives estimates with a smaller standard deviation and, somewhat surprisingly, a slightly smaller mean squared error. Regressing logit of the sample proportion against the logarithm of load is a slight improvement on the probit analysis and a considerable improvement on a regression of logit against load.

When sampling from both normal and Weibull distributions the regression of the logit against the logarithm of load gives an estimate of the lower 1% point with the smallest mean squared error. Overall, a regression of the logit on the logarithm of load seems to be the most robust approach if normality of strengths is in doubt.

ISSN: 1790-2769

ISBN: 978-960-474-034-5