

Jeffrey F. Bromaghin¹ and Richard M. Engeman¹

Head-to-Head Comparison of SAS and ASTM-Proposed Probit Computer Programs

REFERENCE: Bromaghin, J. F. and Engeman, R. M., "Head-to-Head Comparison of SAS and ASTM-Proposed Probit Computer Programs." *Aquatic Toxicology and Environmental Fate: Eleventh Volume, ASTM STP 1007*, G. W. Suter II and M. A. Lewis, Eds., American Society for Testing and Materials, Philadelphia, 1989, pp. 303-307.

ABSTRACT: Estimates of the median lethal dose (LD_{50}) [or the median lethal concentration (LC_{50})] are used in toxicological and hazard assessment studies to indicate the dosage of a substance that produces a 50% response (death) rate in a population. The calculation of LD_{50} estimates is frequently accomplished through probit analysis. Probit analysis is based on certain underlying assumptions about the tolerance of the population in question to the substance being tested. The response rate is assumed to follow a cumulative normal distribution with respect to the dosage or the logarithms of the dosage.

Recently we reviewed ASTM's proposed standard practice for using probit analysis (Draft No. 2). A BASIC computer program for performing probit analysis was included with the proposed standard. We evaluated this program using the PROBIT procedure in the well-known SAS Institute package as a standard of comparison. The SAS procedure offers a default option that assumes the response rate of the population follows the cumulative normal distribution. In addition, two logarithmic options are also offered (natural and base 10), whereby the response rate is assumed cumulative normal relative to the logarithm of the dose. The BASIC and SAS programs give very similar results when the appropriate logarithmic option is used with the SAS procedure. Thus, the ASTM program would be a subset of the SAS procedure in that it is comparable to one of the SAS options, employing one set of possible assumptions. We believe that the assumptions one would make about the distributional response of a population are not necessarily standard, and that careful examination of the data relative to analytical assumptions prior to analysis is recommended.

KEY WORDS: probit computer programs, aquatic toxicology

Estimates of the median lethal dose (LD_{50}) [or median lethal concentration (LC_{50})] are used or required in toxicological and hazard assessment studies to indicate the dosage of a substance that produces a 50% response (death) rate in a population. Various analytical methods are available for calculating LD_{50} estimates [1]. Probit analysis is one of the most frequently used of the parametric methods and is based on the assumption that the response rate of the population in question to the substance being tested follows a cumulative normal distribution with respect to dosage or the logarithm of the dosage.

Recently, we reviewed ASTM's proposed standard practice for using probit analysis (Draft No. 2). A BASIC computer program for performing probit was included with the proposed standard. This program was installed on a Radio Shack Model 16 microcomputer. Because estimates from the ASTM program calculated LD_{50} values that were different than those calculated in our routine data analyses, we proceeded to evaluate this pro-

¹ Statisticians, U.S. Department of Agriculture/Animal and Plant Health Inspection Service, Animal Damage Control Program, Denver Wildlife Research Center, Denver, CO.

gram using the PROBIT procedure in the well-known SAS package [2] as a standard of comparison.

The SAS procedure provides three options for analyses corresponding to two assumptions about the distributional relationship between the response rate and dosage level. The default option assumes that the response rate follows the cumulative normal distribution with respect to the dosage. The other two options, which must be specified to be implemented, assume that the response rate is cumulative normal with respect to a logarithm of dosage. One option uses the base 10 logarithm of the dosage while the other uses the natural (base e) logarithm.

Our initial attempts at using the ASTM program to analyze data previously analyzed using SAS indicated that for calculating an LD_{50} , the ASTM program probably was equivalent to one of the SAS options. To determine this, exact data generated from known distributions with known LD_{50} s were used for comparing the ASTM-proposed program to the PROBIT procedure in SAS.

Test Data

Two artificial data sets were generated based on specified normal distributions. These data sets provide a controlled basis of comparison between the two programs because the properties (including the LD_{50}) of the data are known.

One data set was constructed using a normal distribution with a mean (LD_{50}) of 4.0 and a standard deviation of 1.25. Eleven dosage levels, ranging from 1.5 to 6.5 in increments of 0.5, were selected for inclusion in the analyses.

The second data set was based on a normal distribution with a mean (LD_{50}) equal to 8 and a standard deviation equal to 10. Five dosage levels, ranging from 2 to 32 by multiples of 2, were considered for the analyses.

Specification of distributions and dosage levels of interest allows exact population response (death) rates for each dosage to be calculated. Calculation of these response rates for theoretical populations requires the use of the cumulative normal distribution function (see, for example, Ref 3). The standard normal cumulative distribution function is well tabulated in many statistics texts and is denoted by $\Phi(z)$, the probability that a standard normal random variable, Z , is less than or equal to z .

Response rates for populations with normal tolerance distributions with respect to dosage are calculated using the following equation

$$\text{response rate} = \Phi \left[\frac{\text{dosage} - \text{mean}}{\text{standard deviation}} \right] \quad (1)$$

For example, in our first data set (mean = 4.0, standard deviation = 1.25) we calculate the response rate for the 5.5 dosage by

$$\Phi \left[\frac{5.5 - 4.0}{1.25} \right] = \Phi(1.2) = 0.88$$

To calculate the population response rates when distributed cumulative normal with respect to the logarithm (base 10 or natural) of the dosage, Eq 1 is employed using logarithms of dosage, mean, and standard deviation. For the purposes of our analyses we assumed that 100 subjects were tested at each dosage. The data that we produced for comparing the two probit programs are given in Table 1.

TABLE 1—Test data generated for comparing two probit programs (the number responding is based on 100 subjects tested at each dosage).

Data Set	Dosage Level	Number Responding		
		Normal with Respect to Dosage	Normal with Respect to \log_{10} (Dosage)	Normal with Respect to \ln (Dosage)
1	1.5	2	0	0
	2.0	5	0	0
	2.5	12	2	2
	3.0	21	10	10
	3.5	34	27	27
	4.0	50	50	50
	4.5	66	70	70
	5.0	79	84	84
	5.5	88	92	92
	6.0	95	97	97
	6.5	98	99	99
2	2	38	32	32
	4	42	41	41
	8	50	50	50
	16	66	59	59
	32	88	68	68

Program Runs

The variables input into both of the programs are dosage, sample size at each dose, and number responding at each dose. For each run on each data set, the dosage values given in Table 1 were input to the program. We always assumed a sample size of 100 for each dosage level in both data sets. Three runs were made on each data set using the ASTM program, one run for each category of response rates (corresponding to the two underlying assumptions about the tolerance distributions). The same three runs were made on each data set for each of the three options in the SAS PROBIT procedure (thus, three times as many runs were made using SAS). The output information from each run that was used for comparative purposes included the point estimate of the LD_{50} , the 95% fiducial limits, and the chi-square value for testing goodness-of-fit.

As an additional test of the ASTM program, we input a new set of dosage levels with the response rates that are normal with respect to dosage. The new dosage levels were calculated as 10 to the power of the original dosages (10^{dosage}). Base 10 logarithms of the LD_{50} estimate and fiducial limits were calculated for comparison with the results produced by the analysis of the original dosage levels and the same response rates when the default in the SAS PROBIT procedure was used.

Results

Analysis of the sets of response rates that are normally distributed with respect to the base 10 and natural logarithms of dosage yielded identical results, because the response rate data in Table 1 input into the programs were identical. Also, the results from using the base 10 and natural logarithm options in the SAS PROBIT procedure always gave identical results for each data set analyzed using these options. For this reason, we will

only examine the results from using the default and base 10 logarithm options in SAS. Similarly, we will not present separate results for data analyses where the response rates were calculated for the response rates that are cumulative normal with respect to the natural logarithm of dosage. The results from the analyses of the data in Table 1 are presented in Table 2.

The ASTM program and the option in SAS that assumes the response rate to be normally distributed with respect to the logarithm of dosage produce nearly identical results. When the sets of response rates that satisfied this distributional assumption were submitted to these routines, both produced the correct point estimate of the LD₅₀, small chi-square goodness-of-fit values, and fiducial limits that contained the true LD₅₀ (see Table 2). Results produced by these routines were always identical to at least two decimal places. When the data normally distributed with respect to the base 10 logarithm of dosage were submitted, the default option in SAS overestimated the LD₅₀. Correspondingly, larger goodness-of-fit values were also produced, and the 95% fiducial limits just missed containing the true LD₅₀.

When the sets of response rates normally distributed cumulatively with respect to dosage were analyzed, the default option in the SAS PROBIT procedure was the only routine to produce correct results. The ASTM program and the base 10 logarithm option in SAS underestimated the LD₅₀ in this case. Again, the wrong underlying assumption produced larger chi-square goodness-of-fit values. The 95% fiducial limits barely contained the true LD₅₀.

In all cases the choice of routine incorporating the appropriate distributional assumption produced the correct LD₅₀ estimates. If the appropriate program (or option) was not used, the results ranged from being biased a small amount in the first data set to a large amount of relative bias in the second data set.

The data using 10 to the power of the dosage and the set of response rates cumulatively normally distributed with respect to dosage were analyzed with the ASTM program (Table 3). Transforming the estimate of the LD₅₀ and the fiducial limits in the 10 to the power of dose scale back to the original scale produced estimates nearly identical to those obtained using the default option in the SAS PROBIT procedure. The chi-square goodness-of-fit statistics produced in these two runs were also nearly identical.

Discussion

The ASTM-proposed probit program produces results very similar to those given by the option in the SAS PROBIT procedure that assumes the response rate to be normally distributed with respect to the logarithm of the dosage. Both of these routines appear to follow the standard logarithmic transformation as found in Ref. 1. Use of this transformation appears to be the most common approach to probit analyses. SAS is one of the most widely used and respected program packages available, and the use of a default option in the PROBIT procedure, where no transform is calculated, indicates a widespread application of the assumption that response rate is frequently considered cumulative normal with respect to dosage rather than the logarithm of dosage. A standard program incorporating only one of the possible distributional assumptions can be considered too restrictive. This limitation is emphasized by the poor results obtained in the present study when the distributional assumption is violated.

The results in Table 3 indicate that the ASTM program can be "tricked" into producing the same estimates as the default option in the SAS PROBIT procedure. If the response rates are suspected to be normally distributed with respect to dosage, then raising 10 to the power of the dosage and inputting these values as the dosages into the ASTM program will produce results that can be back transformed to the values that the SAS default option

TABLE 2—Results from analyses of the data in Table 1.

Data Set	True LD ₅₀	Normal with Respect to Dose or Log ₁₀ Dose	Program	Estimated LD ₅₀	95% Fiducial Limits	Chi-Square
1	4.0	dose	SAS default	4.000	3.88, 4.12	0.1772
			SAS log 10	3.800	3.59, 4.02	20.5966
			ASTM	3.800	3.58, 4.02	20.5946
		log ₁₀ dose	SAS default	4.112	4.01, 4.21	5.9737
			SAS log 10	3.999	3.90, 4.10	0.3706
			ASTM	3.999	3.90, 4.10	0.3706
2	8.0	dose	SAS default	8.021	5.35, 10.36	0.0277
			SAS log 10	5.365	1.46, 10.91	8.8362
			ASTM	5.365	1.46, 10.85	8.8364
		log ₁₀ dose	SAS default	12.283	8.04, 16.57	4.1715
			SAS log 10	8.000	5.61, 11.42	0.0039
			ASTM	8.000	5.61, 11.42	0.0039

TABLE 3—Results from using 10 to the power of dosage and response rates distributed normally with respect to dosage.

Data Set	True LD ₅₀	Value Estimated	ASTM Estimate 10 ^{dosage} Scale	ASTM Estimate Transformed to Original Scale	SAS Default Estimate (Table 2)
1	4	LD ₅₀	10 ⁴	4,000	4,000
		fiducial limits chi-square	7628.5, 13 108.7 0.1771	3.88, 4.12 ...	3.88, 4.12 0.1772
2	8	LD ₅₀	1,0494 × 10 ⁸	8,021	8
		fiducial limits chi-square	223 450, 2,308 × 10 ¹⁰ 0.0277	5.35, 10.36 ...	5.35, 10.36 0.0277

would have produced. The options available in the SAS PROBIT procedure can be duplicated in the ASTM program through prior data manipulation.

Although the SAS PROBIT procedure gives more information on output (such as LD values from LD₀₁ to LD₉₉ and plots), we feel that the primary limitation of the ASTM program is the lack of options corresponding to different possible distributional assumptions. The rigid use of a particular routine without considering the underlying assumptions can produce poor results. Our analyses indicate that these methods, at least in these large sample cases (100 subjects per dose), may not be particularly robust to departures from the underlying assumptions. We recommend using analytical methods based on examination of the data to determine the most appropriate underlying assumptions. When using the ASTM program this would require a run for the original dosages and one using transformed dosages. Examination of the goodness-of-fit results helps to provide a means of comparing the underlying assumptions prior to commitment to a final analytical technique.

References

- [1] Finney, D. J., *Statistical Method in Biological Assay*, Griffin, London, 1971.
- [2] SAS Institute Inc., *SAS User's Guide: Statistics*, Cary, NC, 1985.
- [3] Mood, A. M., Graybill, F. A., and Boes, D. C., *Introduction to the Theory of Statistics*, McGraw-Hill, New York, 1974.