NOTICE: This standard has either been superseded and replaced by a new version or withdrawn.
Contact ASTM International (www.astm.org) for the latest information

Designation: E 1847 – 96 (Reapproved 2003)

Standard Practice for Statistical Analysis of Toxicity Tests Conducted Under ASTM Guidelines[1]

This standard is issued under the fixed designation E 1847; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This practice covers guidance for the statistical analysis of laboratory data on the toxicity of chemicals or mixtures of chemicals to aquatic or terrestrial plants and animals. This practice applies only to the analysis of the data, after the test has been completed. All design concerns, such as the statement of the null hypothesis and its alternative, the choice of alpha and beta risks, the identification of experimental units, possible pseudo replication, randomization techniques, and the execution of the test are beyond the scope of this practice. This practice is not a textbook, nor does it replace consultation with a statistician. It assumes that the investigator recognizes the structure of his experimental design, has identified the experimental units that were used, and understands how the test was conducted. Given this information, the proper statistical analyses can be determined for the data.

1.1.1 Recognizing that statistics is a profession in which research continues in order to improve methods for performing the analysis of scientific data, the use of statistical methods other than those described in this practice is acceptable as long as they are properly documented and scientifically defensible. Additional annexes may be developed in the future to reflect comments and needs identified by users, such as more detailed discussion of probit and logistic regression models, or statistical methods for dose response and risk assessment.

1.2 The sections of this guide appear as follows:

Title	Section
Referenced Documents	2
Terminology	3
Significance and Use	4
Statistical Methods	5
Flow Chart	6
Flow Chart Comments	7
Keywords	8
References	

1.3 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:* [2]

E 178 Practice for Dealing with Outlying Observations
E 380 Practice for Use of the International System of Units (SI) (the Modernized Metric System)
E 456 Terminology Relating to Statistics
E 1241 Guide for Conducting Early Life Stage Toxicity Tests with Fishes
E 1325 Terminology Relating to Design of Experiments

3. Terminology

3.1 *Definitions of Terms Specific to This Standard*—The following terms are defined according to the references noted:

3.1.1 *analysis of variance (ANOVA)*—a technique that subdivides the total variation of a set of data into meaningful component parts associated with specific sources of variation for the purpose of testing some hypothesis on the parameters of the model or estimating variance components (**1**).[3]

3.1.2 *categorical data*—variates that take on a limited number of distinct values (**2**).

3.1.3 *censored data*—some subjects have not experienced the event of interest at the end of the study or time of analysis. The exact survival times of these subjects are unknown (**3**).

3.1.4 *central limit theorem*—whatever the shape of the frequency distribution of the original populations of X's, the frequency distribution of the mean, in repeated random samples of size n tends to become normal as n increases (**2**).

3.1.5 *central tendency measure*—a statistic that measures the central location of the sample observations (**4**).

3.1.6 *concentration-response testing*—the quantitative relation between the amount of factor X and the magnitude of the effect it causes is determined by performing parallel sets of

[1] This practice is under the jurisdiction of ASTM Committee E47 on Biological Effects and Environmental Fate and is the direct responsibility of Subcommittee E47.06 on Terminology and Technical Services.

Current edition approved Dec. 10, 1996. Published February 1997.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] The boldface numbers given in parentheses refer to a list of references at the end of the text.

Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States.

Documents
Document Sharing Hub
Full document details: ASTM E1847-96(2003)

operations with various known amounts, or doses, of the factor and measuring the result, that is called the response (**5**).

3.1.7 *continuous data*—a variable that can assume a continuum of possible outcomes (**4**).

3.1.8 *control*—an experiment in which the subjects are treated as in a parallel experiment except for omission of the procedure or agent under test and that is used as a standard of comparison in judging experimental effects (**6**).

3.1.9 *dichotomous data*—variates that have only 2 mutually exclusive outcomes, binary data, success or failure data (**3**).

3.1.10 *dispersion measure*—a statistic that measures the closeness of the independent observations within groups, or relative to a sample's central value (**4**).

3.1.11 *distribution*—a set of all the various values that individual observations may have and the frequency of their occurrence in the sample or population (**1**).

3.1.12 *duplication*—the execution of a treatment at least twice under similar conditions (**1**).

3.1.13 *experimental unit*—a portion of the experimental space to which a treatment is applied or assigned in the experiment (**1**).

3.1.14 *homogeneity*—lack of significant differences among mean squares of an analysis (**2**).

3.1.15 *hypothesis test*—a decision rule (strategy, recipe) which, on the basis of the sample observations, either accepts or rejects the null hypothesis (**4**).

3.1.16 *independence*—having the property that the joint probability (as of all events or samples) or the joint probability density function (as of random variables) equals the product of the probabilities or probability density functions of separate occurrence (**6**).

3.1.17 *mean*—a measure of central tendency or location that is the sum of the observations divided by the number of observations (**1**).

3.1.18 *model*—an equation that is intended to provide a functional description of the sources of information which may be obtained from an experiment (**1**).

3.1.19 *nonparametric statistic*—a statistic which has certain desirable properties that hold under relatively mild assumptions regarding the underlying populations (**4**).

3.1.20 *normality*—having the characteristics of a normal distribution (**2**).

3.1.21 *outlier*—an outlying observation is one that appears to deviate markedly from other members of the sample in which it occurs (see Practice E 178).

3.1.22 *parametric statistic*—a statistic that estimates an unknown constant associated with a population (**4**).

3.1.23 *probit logit*—when the response Y in binary, the probit/logit equation is as follows:

$$p = Pr(Y = 0) = C + (1 - C)\, F(\boldsymbol{x}'\boldsymbol{b}) \tag{1}$$

where:

$\boldsymbol{b}$ = vector of parameter estimates,
F = cumulative distribution function (normal, logistic),
$\boldsymbol{x}$ = vector of independent variables,
p = probability of a response, and
C = natural (threshold) response rate.

The choice of the distribution function, F, (normal for the probit model, logistic for the logit model) determines the type of analysis (**7**).

3.1.24 *regression analysis*—the process of estimating the parameters of a model by optimizing the value of an objective function (for example, by the method of least squares) and then testing the resulting predictions for statistical significance against an appropriate null hypothesis model (**1**).

3.1.25 *replication*—the repetition of the set of all the treatment combinations to be compared in an experiment. Each of the repetitions is called a replicate (**1**).

3.1.26 *residual*—Y_{obs} minus Y_{pred} – the difference between the observed response variable value and the response variable value that is predicted by the model that is fit to the data (**8**).

3.1.27 *scedasticity*—variance (**5**).

3.1.28 *significance level*—the probability at which the null hypothesis is falsely rejected, that is, rejecting the null hypothesis when in fact it is true (**4**).

3.1.29 *transformation*—the transformation of the observations *Xij* into another scale for purposes of allowing the standard analysis to be used as an adequate approximation (**2**).

3.1.30 *treatment*—a combination of the levels of each of the factors assigned to an experimental unit (see Terminology E 456).

3.1.31 *variance*—a measure of the squared dispersion of observed values or measurements expressed as a function of the sum of the squared deviations from the population mean or sample average (see Terminology E 456).

4. Significance and Use

4.1 The use of statistical analysis will enable the investigator to make better, more informed decisions when using the information derived from the analyses.

4.1.1 The goals when performing statistical analyses, are to summarize, display, quantify, and provide objective measures for assessing the relationships and anomalies in data. Statistical analyses also involve fitting a model to the data and making inferences from the model. The type of data dictates the type of model to be used. Statistical analysis provides the means to test differences between control and treatment groups (one form of hypothesis testing), as well as the means to describe the relationship between the level of treatment and the measured responses (concentration effect curves), or to quantify the degree of uncertainty in the end-point estimates derived from the data.

4.1.2 The goals of this practice are to identify and describe commonly used statistical procedures for toxicity tests. Fig. 1, Section 6, following statistical methods (Section 5), presents a flow chart and some recommended analysis paths, with references. From this guideline, it is recommended that each investigator develop a statistical analysis protocol specific to his test results. The flow chart, along with the rest of this guideline, may provide both useful direction, and service as a quality assurance tool, to help ensure that important steps in the analysis are not overlooked.

AVA Documents
Document Sharing Hub

Full document details: ASTM E1847-96(2003)

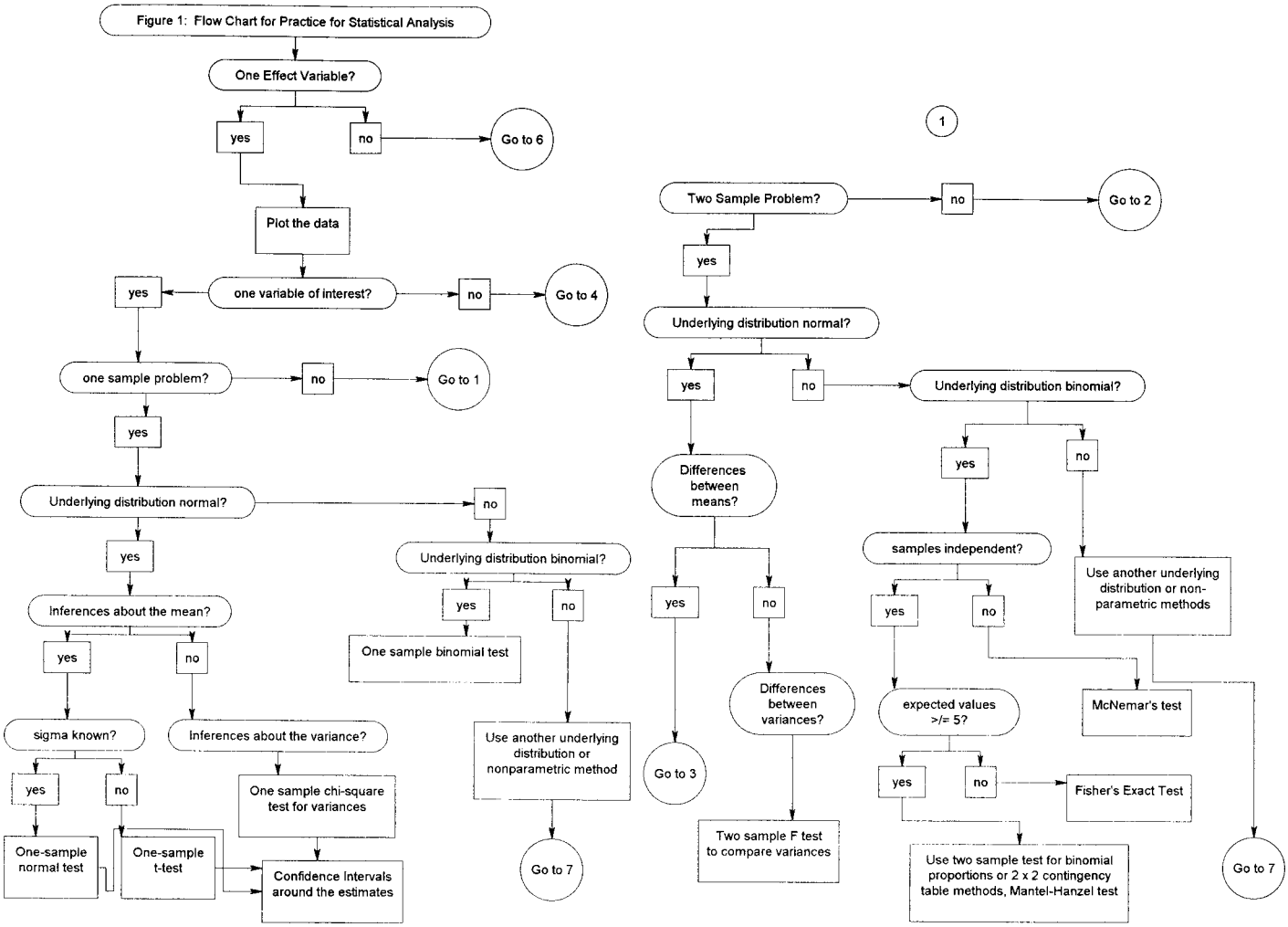

FIG. 1 Flow Chart for Practice for Statistical Analysis

5. Statistical Methods

5.1 *Exploratory Data Analysis*—The first step in any data analysis is to look at the data and become familiar with their content, structure, and any anomalies that might be present.

5.1.1 *Plots*:

5.1.1.1 Histograms are unidimensional plots that show the distributional shapes in the data and the frequencies of individual values. These diagrams allow the investigator to check for unusual observations and also visually check the validity of some assumptions that are necessary for several statistical analyses that may be used (**9**).

5.1.1.2 Scatter plots of two or more variables demonstrate the relationships among the variables, so that correlations can be observed and interactions can be studied. These plots are very useful when looking for concentration effect relationships (**9**).

5.1.1.3 Normality and box plots are additional plots that give distributional information, quantiles and pictures of the data, either as a whole or by treatment group (**9**).

5.1.2 *Outliers*—On occasion, some data points in the histogram, scatter plot, or box plot, appear to be quite different from the majority of points. These data, known as outliers, can be tested to determine if they are truly different from the distribution of the experimental data (**10**). The Z or t scores are usually used for testing, with a confidence level chosen by the investigator. If they are different and can be attributed to an error in the execution of the study (violation of protocol, data entry error, and so forth), then they can be removed from the analyses. However, if there is no legitimate reason to remove them, then they must be kept in the analyses. It is recommended that the analyses can be conducted on two data sets, the complete one and one with the outliers removed. In this way, the outliers' influence on the analyses can be studied.

5.1.3 *Non-Detected Data*:

5.1.3.1 Data that fall below a chemical analysis threshold level of detection, in an analytical technique used to measure a value, are called non-detected. Values that occur above the detection limit but are below the limit of quantitation, are called non-estimable. Occasionally, the two terms are used interchangeably. Essentially, these data are results for which no reliable number can be determined.

5.1.3.2 In analyzing a data set containing one or more non-detects, several methods can be used. If the amount of non-detects is below approximately 25 % of the entire data set, then the non-detects can be replaced by one half the detection limit (or quantitation limit, whichever is appropriate) and

AVA Documents
Document Sharing Hub

Full document details: ASTM E1847-96(2003)