

<p>Volume III Other Lab Operations Section 4 – Laboratory Applications of Statistical Concepts</p>	<p>ORA LABORATORY MANUAL</p> <p>FDA Office of Regulatory Affairs Office of Regulatory Science</p>	<p>DOCUMENT NO.: III-04 VERSION NO.: 1.4 FINAL EFFECTIVE DATE: 10-01-03 Revised:01/31/13</p>
<p>Section 4 Basic Statistics and Data Presentation Section 4</p>		

Contents

4.1	Introduction
4.2	General Considerations
4.2.1	Accuracy, Precision, and Uncertainty
4.2.2	Error and Deviation; Mean and Standard Deviation
4.2.3	Random and Determinate Error
4.2.4	The Normal Distribution
4.2.5	Confidence Intervals
4.2.6	Populations and Samples; Student's t Distribution
4.2.7	References
4.3	Data Handling and Presentation
4.3.1	Rounding of Reported Data
4.3.2	Significant Figures
4.3.2.1	Definitions and Rules for Significant Figures
4.3.2.2	Significant Figures in Calculated Results
4.4	Linear Curve Fitting
4.5	Development and Validation of Spreadsheets for Calculation of Data
4.5.1	Introduction
4.5.2	Development of Spreadsheets
4.5.3	Validation of Spreadsheets
4.6	Control Charts
4.6.1	Definitions
4.6.2	Discussion
4.6.3	Quality Control Samples
4.6.4	References
4.7	Statistics Applied to Drug Analysis
4.7.1	Introduction
4.7.2	USP Guidance on Significant Figures and Rounding
4.7.3	Additional Guidance in the USP
4.7.4	References
4.8	Statistics Applied to Radioactivity
4.8.1	Introduction

4.8.2	Sample Counting
4.8.3	Standard Deviation and Confidence Levels
4.8.4	Counting Rate and Activity
4.9	Statistics Applied to Biological Assays
4.10	Statistics Applied to Microbiological Analysis
4.10.1	Introduction
4.10.2	Geometric Mean
4.10.3	Most Probable Number
4.10.4	References
4.11	Statistics Applied to pH in Canned Foods
4.12	Document History

4.1 Introduction

Statistics in the ORA laboratory may be used to describe and summarize the results of sample analysis in a concise and mathematically meaningful way. Statistics may also be used to predict properties (ingredient, acidity, quantity, dissolution, height, weight) of a contaminant or of a regulated product as a whole based on measurements made on a subset, or sample, of the contaminant or product. All statistical concepts are ultimately based on mathematically derived laws of probability. Understanding statistical concepts will allow the ORA analyst to better convey analytical results with the maximum amount of assurance as to quality of the data.

By profession and training, the ORA laboratory analyst is motivated to carry out measurements precisely, and to report results that contain the maximum amount of useful information. Proper application of statistics gives this ability, while allowing for the fact that there is inherent error (both random and determinate) in virtually every laboratory measurement made.

This section is not meant to be an in-depth reference for the myriad ways that statistics could be applied in the laboratory, but rather as a general guide for situations commonly encountered in the ORA laboratory. The section also gives guidance on various aspects of data presentation and verification.

4.2 General Considerations

Statistical procedures used to describe measurements of samples in the ORA laboratory allow regulatory decisions to be made in as unbiased manner as possible. The following are numerically descriptive measures commonly used in ORA laboratories.

4.2.1 Accuracy, Precision, and Uncertainty

The *accuracy* of a measurement describes the difference between the measured value and the true value. Accuracy is said to be high or low depending on whether the measured value is near to, or distant from, the true value. *Precision* is concerned with the differences in results for a set of measurements, regardless of the accuracy. Applied to an analytical method as used in an ORA laboratory, a highly precise method is one in which repeated application of the method on a sample will give results which agree closely with one another. Precision is related to *uncertainty*: a series of measurements with high precision will have low uncertainty and vice versa. Terms such as accuracy, precision, and uncertainty are not mathematically defined quantities but are useful concepts in understanding the statistical treatment of data. Exact mathematical expressions of accuracy and precision (*error* and *deviation*), will be defined in the next section.

As an example of these terms, consider shooting arrows at a target, where the “bull’s eye” is considered the true value. An archer with high precision (low uncertainty) but low accuracy will produce a tightly clustered pattern outside the bull’s eye; if low precision (high uncertainty) and low accuracy, the pattern will be random rather than clustered, with the bull’s eye being hit only by chance. The best situation is high accuracy and high precision: in this case a tight cluster is found in the bull’s eye area. This example illustrates another important concept: accuracy and precision depend on both the bow and arrow, *and* the archer. Applied to a laboratory procedure, this means that the reliability of results depends on both the apparatus/instruments used *and* the analyst. It is extremely important to have a well trained analyst who understands the method, applies it with care (for example by careful weighing and dilution), and uses a calibrated instrument (demonstrated to be operating reliably). Without all of these components in place, it is difficult to obtain the reliable results needed for regulatory analysis.

4.2.2 Error and Deviation; Mean and Standard Deviation

The concepts of accuracy and precision can be put on a mathematical basis by defining equivalent terms: *error* and *deviation*. This will allow the understanding of somewhat more complicated statistical formulations used commonly in the ORA laboratory.

If a set of N replicate measurements $x_1, x_2, x_3, \dots, x_n$, were made (examples: weighing a vial N times, determining HPLC peak area of N injections from a single solution, measuring the height of a can N times, ...), then:

$$E_i = x_i - \mu$$

where E_i = error associated with measurement i ,

x_i = result of measurement i , and

μ = true value of measurement.

The definition of error often has little immediate practical application, since in many cases μ , the true value, may not be known. However, the process of calibration against a known value (such

as a chemical or physical standard) will help to minimize error by giving us a known value with which to compare an unknown.

The deviation, a measure of precision, is calculated without reference to the true value, but instead is related to the *mean* of a set of measurements. The mean is defined by:

$$\bar{X} = \sum_{i=1}^N \frac{x_i}{N}$$

where \bar{X} = mean of set of N measurements,
 $x_i = i^{\text{th}}$ measurement, and
N = Number of Measurements.

Note: this is the *arithmetic mean* of a set of observations. There are other types of mean which can be calculated, such as the *geometric mean* (see the section on “Application of Statistics to Microbiology” below), which may be more accurate in special situations.

Then, the deviation, d_i , for each measurement is defined by:

$$d_i = x_i - \bar{X}$$

Using the example of the archer shooting arrows at a target, the deviation for each arrow’s position is the distance from the arrow’s position to the calculated mean of all of the arrow’s positions.

Finally, the expression of deviation most useful in many ORA laboratory applications is s , the *standard deviation*:

$$s = \sqrt{\frac{\sum_{i=1}^N d_i^2}{N-1}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{N-1}}$$

where s = standard deviation, and other terms are as previously defined.

The standard deviation is then a measure of precision of a set of measurements, but has no relationship to the accuracy. The standard deviation may also be expressed in relative terms, as the *relative standard deviation*, or RSD:

$$\text{RSD (\%)} = \frac{(100)(s)}{\bar{X}}$$

Whereas the standard deviation has the same units as the measurement, the RSD is dimensionless, and expressed as a percentage of the mean.

Standard deviation as defined above is the correct choice when we have a *sample* drawn from a larger population. This is almost always the case in the ORA laboratory: the sample which has been collected is assumed to be “representative” of the larger population (for example, a batch of tablets, lot of canned goods, field of wheat) from which it has been taken. As it is taken through analytical steps in the laboratory (by subsampling, compositing, diluting, etc.) the representative characteristic of the sample is maintained.

If the entire population is known for measurement, the standard deviation s is redefined as σ , the *population standard deviation*. The formula for σ differs from that of s in that $(N-1)$ in the denominator is replaced by N . The testing of an entire population would be a rare circumstance in the ORA laboratory, but may be useful in a research project.

Statistical parameters such as mean and standard deviation are easily calculated today using calculators and spreadsheet formulas. Although this is convenient, the analyst should not forget how these parameters are derived.

4.2.3 Random and Determinate Error

Recall the definition of error in section 4.2.2 above. Errors in measurement are often divided into two classes: *determinate error* and *non-determinate error*. The latter is also termed *random error*. Both types of error can arise from either the analyst or the instruments and apparatus used, and both need to be minimized to obtain the best measurement, that with the smallest error.

Determinate error is error that remains fixed throughout a set of replicate measurements. Determinate error can often be corrected if it is recognized. Examples include correcting titration results against a blank, improving a chromatographic procedure so that a co-eluting peak is separated from the peak of interest, or calibrating a balance against a NIST-traceable standard. In fact, the purpose of most instrument calibrations is to reduce or eliminate determinate error. Using the example of the archer shooting arrows at a target, calibration of the sights of the bow would decrease the error, leading to hitting the bull’s eye.

Random error is error that varies from one measurement to another in an unpredictable way in a set of measurements. Examples might include variations in diluting to the mark during volumetric procedures, fluctuations in an LC detector baseline over time, or placing an object to be weighed at different positions on the balance pan. Random errors are often a matter of analytical technique, and the experienced analyst, who takes care in critical analytical operations, will usually obtain more accurate results.

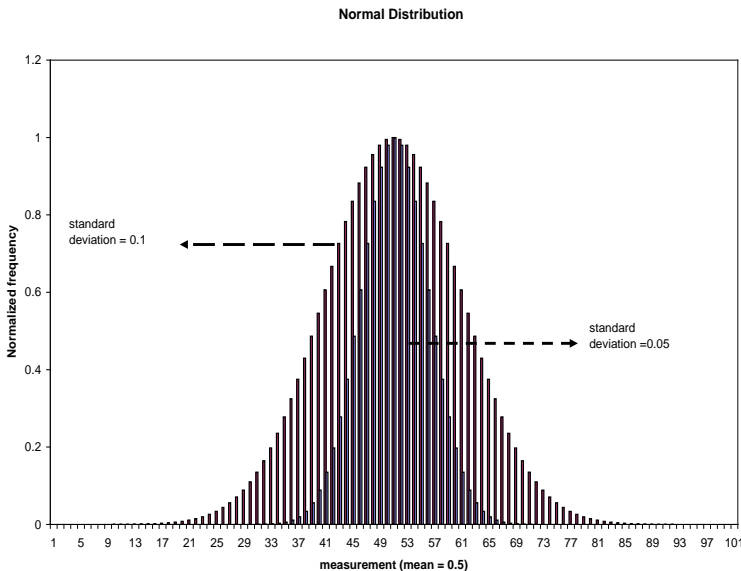
4.2.4 The Normal Distribution

In the introduction to this chapter, it was briefly mentioned that statistics is derived from the mathematical theory of probability. This relationship can be seen when we consider *probability distribution functions*, of which the *normal distribution function* is an important example. The normal distribution curve (or function) is of great value in aiding understanding of measurement statistics, and to interpret results of measurements. Although a detailed explanation is outside the scope of this chapter, a brief explanation will be beneficial. The normal distribution curve describes how the results of a set of measurements are distributed as to frequency; assuming only random errors are made. It describes the probability of obtaining a measurement within a specified range of values. It is assumed here that the values measured (i.e. variables) may vary continuously rather than take on discrete values (the *Poisson distribution*, applicable to radioactive decay is an example of a discrete probability distribution function; see discussion under “Statistics Applied to Radioactivity”). The normal distribution should be at least somewhat familiar to most analysts as the “bell curve” or *Gaussian curve*. The curve can be defined with just two statistical parameters that have been discussed: the true value of the measured quantity, μ , and the true standard deviation, σ . It is of the form:

$$Y = e^{-1/2\{(x-\mu)/\sigma\}^2}$$

Where Y= frequency of occurrence of a measurement (a value between 0 and 1),
 x = the magnitude of the measurement,
 μ = the true value of the measurement,
 σ = true standard deviation of the population, and
 e = base of natural logarithms (2.718...).

An example of two normal curves with the same true value, μ , but two different values of σ is shown below (this was calculated using an Excel[®] spreadsheet, using the formula above and an array of x values):



Some properties of the normal distribution curve that are evident by inspection of the graph and mathematical function above go far in explaining the properties of measurements in the laboratory:

- In the absence of determinate errors, the measurement with the most probable value will be the true value, μ .
- Errors (i.e. $x-\mu$), as defined previously, are distributed symmetrically on either side of the true value, μ ; errors greater than the mean are equally as likely as errors below the mean.
- Large errors are less likely to occur than small errors.
- The curve never reaches the y -axis but approaches it asymptotically: there is a finite probability of a measurement having any value.
- The probability of a measurement being the true value increases as the standard deviation decreases.

4.2.5 Confidence Intervals

The *confidence interval* of a measurement or set of measurements is the range of values that the measurement may take with a stated level of uncertainty. Although confidence intervals may be defined for any probability distribution function, the normal distribution function illustrates the concept well.

Approximately 68% of the area under the normal distribution curve is included within ± 1 standard deviation of the mean. This implies that, for a series of replicate measurements, 68% will fall within ± 1 standard deviation of the true mean. Likewise, 95% of the area under the normal distribution curve is found within about $\pm 2\sigma$ (to be precise, 1.96σ), and approximately 99.7% of the area of the curve is included within a range of the mean $\pm 3\sigma$. A 95% confidence interval for a series of measurements, therefore, is that which includes the mean $\pm 2\sigma$. An example of the application of confidence limits is in the preparation of *control charts*, discussed in Section 7.6 below.

4.2.6 Populations and Samples: Student's *t* Distribution

In the above discussion, we are using the true standard deviation, σ (i.e. the population standard deviation). In most real life situations, we do not know the true value of σ . In the ORA laboratory, we are generally working with a small sample which is assumed to be representative of the population of interest (for example, a batch of tablets, a tanker of milk). In this case, we can only calculate the sample standard deviation, s , from a series of measurements. In this case, s is an estimate of σ , and confidence limits need to be expanded by a factor, t , to account for this additional uncertainty. The distribution of t is called the *Student's t Distribution*. Further discussion is beyond the scope of this chapter, but tables of t values, which depend on both the confidence limit desired and the number of measurements made, are widely published.

4.2.7 References

The following are general references on statistics and treatment of data that may be useful for the ORA Laboratory:

Dowdy, S., Wearden, S. (1991). *Statistics for research* (2nd ed.). New York: John Wiley & Sons.

Garfield, F.M. (1991). *Quality assurance principles for analytical laboratories*. Gaithersburg, MD: Association of Official Analytical Chemists.

Taylor, J. K. (1985). *Handbook for SRM users* (NBS Special Publication 260-100). Gaithersburg, MD: National Institute for Standards and Technology.

4.3 Data Handling and Presentation

In the most general sense, analytical work results in the generation of numerical data. Operations such as weighing, diluting, etc. are common to almost every analytical procedure, and the results of these operations, together with instrumental outputs, are combined mathematically to obtain a result or series of results. How these results are reported is important in determining their

significance. As a regulatory agency, it is important that we report analytical results in a clear, unbiased manner that is truly reflective of the operations that go into the result. Data should be reported with the proper number of *significant digits* and *rounded* correctly. Procedures for accomplishing this are given below:

4.3.1 Rounding of Reported Data

When a number is obtained by calculations, its accuracy depends on the accuracy of the number used in the calculation. To limit numerical errors, an extra significant figure is retained during calculations, and the final answer rounded to the proper number of significant figures (see next section for discussion of significant figures).

The following rules should be used:

- If the extra digit is less than 5, drop the digit.
- If the extra digit is greater than 5, drop it and increase the previous digit by one.
- If the extra digit is five, then increase the previous digit by one if it is odd; otherwise do not change the previous digit.

Examples are given in the following table:

Calculated Number	Significant digits to report	Number with one extra digit retained	Reported rounded number
79.35432	4	79.354	79.35
99.98798	5	99.9879	99.988
32.9653	4	32.965	32.96
32.9957	4	32.995	33.00
0.0396	1	0.039	0.04
105.67	3	105.6	106
29	2	29	29

4.3.2 Significant Figures

Significant figures (or significant digits) are used to express, in an approximate way, the precision or uncertainty associated with a reported numerical result. In a sense, this is the most general way to express “how well” a number is known. The correct use of significant figures is important in today’s world, where spreadsheets, handheld calculators, and instrumental digital readouts are capable of generating numbers to almost any degree of apparent precision, which

may be much different than the actual precision associated with a measurement. A few simple rules will allow us to express results with the correct number of significant figures or digits. The aim of these rules is to ensure that the final result should never contain any more significant figures than the least precise data used to calculate it. This makes intuitive as well as scientific sense: a result is only as good as the data that is used to calculate it (or more popularly, “garbage in, garbage out”).

4.3.2.1 Definitions and Rules for Significant Figures

- All non-zero digits are significant.
- The *most significant digit* in a reported result is the left-most non-zero digit: 359.741 (3 is the most significant digit).
- If there is a decimal point, the *least significant digit* in a reported result is the right-most digit (whether zero or not): 359.741 (1 is the least significant digit). If there is no decimal point present, the right-most non-zero digit is the least significant digit.
- The number of digits between and including the most and least significant digit is the *number of significant digits* in the result: 359.741 (there are six significant digits).

The following table gives examples of these definitions:

	Number	Sig. Digits
A	1.2345 g	5
B	12.3456 g	6
C	012.3 mg	3
D	12.3 mg	3
E	12.30 mg	4
F	12.030 mg	5
G	99.97 %	4
H	100.02 %	5

4.3.2.2 Significant Figures in Calculated Results

Most analytical results in ORA laboratories are obtained by arithmetic combinations of numbers: addition, subtraction, multiplication, and division. The proper number of digits used to express the result can be easily obtained in all cases by remembering the principle stated above: numerical results are reported with a precision near that of the least precise numerical measurement used to generate the number. Some guidelines and examples follow.

Addition and Subtraction

The general guideline when adding and subtracting numbers is that the answer should have decimal places equal to that of the component with the least number of decimal places:

$$\begin{array}{r} 21.1 \\ 2.037 \\ 6.13 \\ \hline \end{array}$$

29.267 = 29.3, since component 21.1 has the least number of decimal places

Multiplication and Division

The general guideline is that the answer has the same number of significant figures as the number with the fewest significant figures:

$$\frac{56 \times 0.003462 \times 43.72}{1.684}$$

A calculator yields an answer of 4.975740998 = 5.0, since one of the measurements has only two significant figures.

4.4 Linear Curve Fitting

This section deals with fitting of experimental data to a mathematical function. This situation is encountered in a variety of situations in the ORA laboratory, in particular with calibration curves. In most situations, the relationship between the variables is linear, and therefore a linear function is needed:

$$y = f(x) = mx + b$$

Where x = independent variable,
 y = dependent variable,
 m = calculated slope of line, and
 b = calculated y-intercept of line.

The *independent variable*, x , is assumed to be known exactly, with no error (such as concentration, distance, time, etc.). The *dependent variable*, y , (instrument response for example) then depends on (is a function of) the value of x . Each value of the independent variable is assumed to follow a normal distribution and to have the same *variance* (i.e. square of the standard deviation). The method of *linear regression* (also known as linear least squares) is used to fit experimental data to a linear function (note: in certain cases, a non-linear relationship may be reduced to a linear equation by a transformation of variables; if so, the linear regression method is still applicable).

The aim of linear regression is to find the line which minimizes the sum of the squares of the deviations of individual points from that line. Once that is accomplished, the slope (m) and the intercept (b) of the ‘least squares’ line is determined. It should be intuitively clear that minimizing deviations of data points from the fitted line gives the best fit of data. Given a set of data points (x_i, y_i) , the equations used to determine the least squares parameters are:

$$m = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (\text{slope})$$

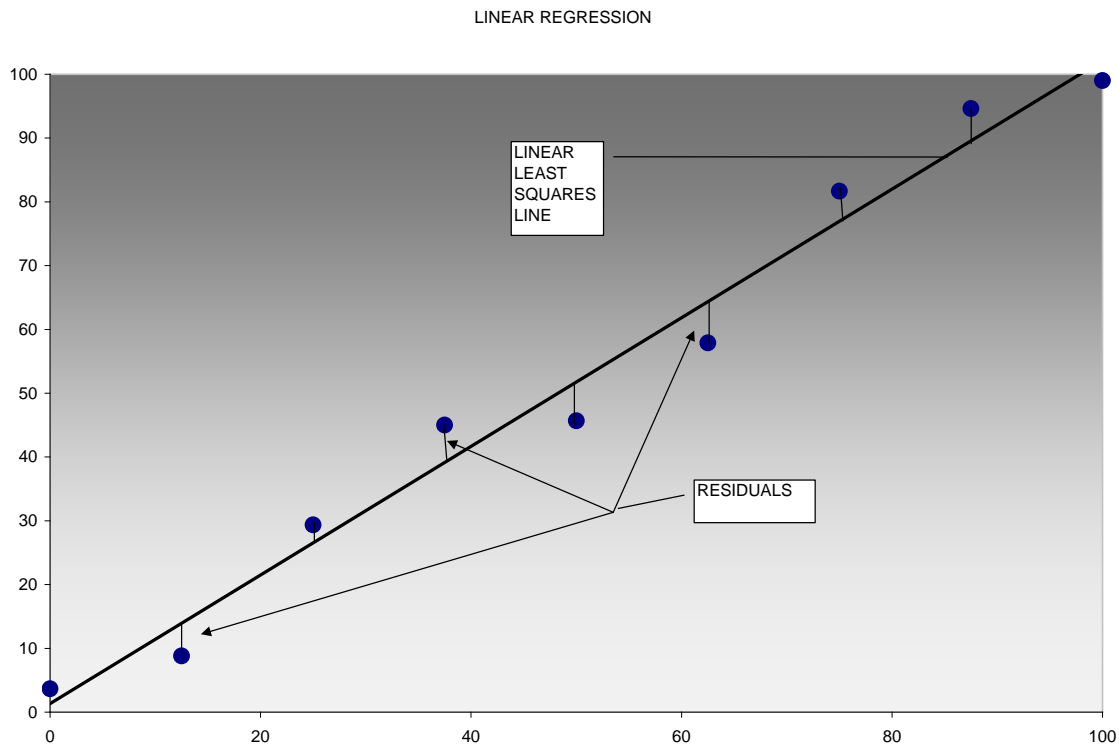
$$b = \frac{\left(\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i \right)}{n} = \bar{Y} - m\bar{X} \quad (\text{intercept})$$

An additional parameter, which is an indicator of the “goodness of fit” of the line to the data points, is the *coefficient of determination*. This coefficient denotes the strength of the linear association between x and y . The coefficient, r^2 , uses information on means and deviations of each data set to express variation numerically. If the two data sets correspond perfectly or exhibits no variation, a coefficient of 1 will be calculated. A coefficient of 0 indicates there is no relationship or no explanation of variation between the two data sets. Typically, for analytical work performed in the ORA laboratory, the coefficient should be very close to 1 (for example 0.999). The formula for the coefficient of determination is:

$$r^2 = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n(s_x)(s_y)} \right)^2$$

where terms have been defined previously.

The following figure illustrates several points relating to linear least squares curve fitting. Data was entered into an Excel[®] spreadsheet and the linear least squares regression line calculated and plotted from the data. The vertical lines indicate the distances (*residuals*) that are minimized in order to achieve the best fit.



4.5 Development and Validation of Spreadsheets for Calculation of Data

When using spreadsheets or programmable calculators for reduction of data generated by sample analyses, there should be assurance that the results are valid and usable for regulatory use. The following section provides guidance for assuring that spreadsheets will meet these criteria.

4.5.1 Introduction

Although the formulas given above for calculation of statistical parameters may seem complicated, matters are simplified by the ready availability of spreadsheets and calculators which provide these values transparently. This makes calculation of statistical parameters much more straightforward than in the past, when direct application of these formulas was used. It is still useful to have some familiarity with these formulas to understand how statistical parameters are derived. In addition, there may be a need to verify the results of statistical data generated by a spreadsheet or calculator; data can be plugged directly into the formulas above to verify these results.

4.5.2 Development of Spreadsheets

Excel[®] and other spreadsheets incorporate all of the statistical parameters discussed, as well as many others. Although individual spreadsheet functions can be considered as reliable, it is important to make sure that data is presented to the spreadsheet with the proper syntax. Also, when spreadsheets are used for multiple numerical calculations in the form of in-house developed templates, it is important to protect the spreadsheet from inadvertent changes, to verify the reliability of the spreadsheet by comparison with known results from known data, and to ensure that the spreadsheet can handle unforeseen data input needs. Spreadsheets developed in the ORA laboratory should be looked upon as in-house developed software that should be qualified before use, just as instruments are qualified before use.

4.5.3 Validation of Spreadsheets

General guidance for design and validation of in-house spreadsheets and other numerical calculation programs includes the following considerations:

- Lock all cells of a spreadsheet, except those needed by the user to input data.
- Make spreadsheets read-only, with password protection, so that only authorized users can alter the spreadsheet.
- Design the spreadsheet so that data outside acceptable conditions is rejected (for example, reject non-numerical inputs).
- Manually verify spreadsheet calculations by entering data at extreme values, as well as at expected values, to assess the ruggedness of the spreadsheet.
- Test the spreadsheet by entering nonsensical data (for example alphabetical inputs, <CTRL> sequences, etc.).
- Keep a permanent record of all cell formulas when the spreadsheet has been developed. Document all changes made to the spreadsheet and control using a system of version numbers with documentation.
- Periodically re-validate spreadsheets. This should include verification of cell formulas and a manual re-verification of spreadsheet calculations.

4.6 Control Charts

A control chart is a graph of test results with limits established in which the test results are expected to fall when the instrument or analytical procedure is in a state of “statistical control.” A procedure is under statistical control when results consistently fall within established control limits. There are a variety of uses of control charts other than identifying results that are out of control. A chart will disclose trends and cycles which will allow real time analysis of data and information for deciding corrective action prior to say an entire analytical system goes out of control. The use of control charts is strongly encouraged in regulatory science.

4.6.1 Definitions

Central line: mean value of earlier determinations, usually a minimum of twenty results

Inner control limit: the mean value ± 2 standard deviations

Outer control limit: the mean value ± 3 standard deviation

4.6.2 Discussion

Control charts are frequently used for quality control purposes in the laboratory. Control charts serve as a tool that determines if results performed on a routine basis (e.g. quality control samples) are acceptable for the intended purposes of the data.

The mean control chart consists of a horizontal central line and two pairs of horizontal control limits lines. The central line defines the mean value, the inner control limit (mean ± 2 standard deviations), and outer control limit (mean ± 3 standard deviations). Results are plotted on the y-axis against the x-axis variable (e.g. date, batch number).

Results fall within the inner control limits 95% of the time. Results falling outside the inner control limit serve as a warning that the results may be biased. Results falling outside the outer control limit indicate the results are biased and corrective action should be taken.

4.6.3 Quality Control Sample Example

The control chart for a laboratory instrument often plots the results of the calibration result (y-axis) against the date (x-axis).

Mean control chart:

- Calculate the mean calibration value

- Calculate ± 2 standard deviation, ± 3 standard deviation values
- Draw horizontal lines above and below the mean value at ± 2 deviations and the mean value ± 3 standard deviations
- Plot calibration results against the date or batch number
- Define corrective actions if the calibration results fall outside the inner and outer control limits.

4.6.4 References

Pecsok, Shields, Cairns. (1986). *Modern methods of analysis* (2nd Ed.). New York: John Wiley and Sons.

Steinmeyer, K. P. (1994). *Mathematics review for health physics technicians*. Hebron, CT: Radiation Safety Associates Publications. (Also 2nd Ed. in 1998.)

4.7 Statistics Applied to Drug Analysis

Chemists in ORA laboratories may have to analyze a wide range of human and animal drugs in a number of different dosage forms, and using differing analytical methods. Statistical evaluation of the analytical results is important for making regulatory decisions.

4.7.1 Introduction

Drug analysis, as well as most analysis performed in the ORA laboratory, relies on the statistical concepts defined above. In addition, there are references in the United States Pharmacopeia (USP) and other official references with which the drug analyst should be familiar.

4.7.2 USP Guidance on Significant Figures and Rounding

Under GENERAL NOTICES, the USP has several references, either direct or implied, to statistics, reporting of results and maintaining precision during an analysis. The drug analyst should be thoroughly familiar with the “Significant Figures and Tolerances” section of the USP. Highlights of this section are summarized as follows:

- Numerical limits specified in a monograph include the extremes of the values and all values in between, but no values outside these limits. This statement should be applied after proper rounding of numerical results. If, for example, a properly rounded result is found to lie exactly at the extreme of a limit (e.g. limits 98.0-102.0% of declared; found 102.03%, rounded to 102.0%) then the monograph limits are met. If the result lies outside the numerical limits (e.g. 98.0-102.0% of declared; found 102.05%, rounded to 102.1%), then the monograph limits are not met.
- Numerical result should be reported to the same number of decimal places as the limit expression stated in the monograph. For example, if limits are stated as 90.0-110.0% of declared, report results to 1 decimal place (e.g. 98.3%, 101.8%), after applying USP rounding rules.
- The USP has slightly different rounding rules than those commonly encountered, as discussed earlier in section 4.3.1. The difference is when a value ends in 5. USP rounding conventions are as follows:
 - Retain only one extra digit to the right of the rightmost digit of the monograph limit expression
 - If the extra digit is less than 5, drop the digit.
 - If the extra digit is greater than 5, drop it and increase the previous digit by one
 - If the extra digit is exactly five, then drop it and (always) increase the previous digit by one.
- An explicit statement is made for titrimetric procedures: essentially all factors, such as weights of analyte, should be measured with precision commensurate with the equivalence statement given in the monograph. Examples in the significant figures section above illustrate the importance of this for all analytical work.
- There is a table given in **SIGNIFICANT FIGURES AND TOLERANCES** that gives examples of USP conventions for rounding, reporting, and comparison of results with compendial limits. This should be reviewed and thoroughly understood by all ORA drug analysts. A few additional examples are given in the following table:

Compendial Requirement	Unrounded Result	Rounded Result	Conforms? (Y/N)
Assay not less than 95.0 And not more than 105.0% of Declared	94.95%	95.0%	Y
	94.94%	94.9%	N
	105.65%	105.7%	N
Limit Test LTE 0.2%	0.24%	0.2%	Y
	0.25%	0.3%	N

4.7.3 Additional Guidance in the USP

Also under GENERAL NOTICES, TESTS AND ASSAYS, is additional guidance. An important section is “Test Results, Statistics, and Standards,” which is of particular regulatory significance. Important points to understand include:

- USP compendial instructions or guidelines are not to be applied “statistically,” meaning the conformance or non-conformance of a product is determined by a single test which may be applied to any portion of a sample, at any time throughout its stated shelf life. The monograph limits are chosen so that inherent uncertainty in the method is taken into account, and system suitability tests verify that the analytical system is reliable; therefore “any specimen tested as directed in the monograph complies”(FDA’s practice, nonetheless, is to perform a check analysis to confirm non-compliance with a monograph limits).
- To emphasize the “singlet determination” viewpoint of the USP, the following statement is made: “Repeats, replicates, statistical rejection of outliers, or extrapolations of results to larger populations are neither specified nor proscribed by the compendia.”

Finally, under GENERAL NOTICES, TESTS AND ASSAYS, the “Procedures” section includes some guidance that should be understood by the ORA Laboratory drug analyst:

- Weights and volumes of test substances and reference standards may be adjusted proportionately, provided that such adjustments do not adversely affect the accuracy of the procedure.
- Similarly, when a method calls for a standardized solution of a known concentration, a solution of a different concentration, molarity, or normality may be used, provided allowance is made for the differing concentration, and the error of measurement is not thereby increased.
- Monographs often use expressions such as “25.0 mL” for volumetric measurements. This is not to be taken literally. In practice, volumes used quantitatively (i.e. the measurement will be used in a quantitative calculation) should be measured to the higher precision specified in “Volumetric Apparatus <31>” of the USP. This generally means that class A flasks, burets, and pipets are to be used, and with proper analytical technique employed. Similarly for weights: “25.0 mg” means that the weighing should be performed with a high precision balance meeting standards set forth in “Weights and Balances <41>.”

4.7.4 References

(Current Ed.). *U. S. Pharmacopeia and national formulary*. Rockville, MD: United States Pharmacopeial Convention, Inc.

4.8 Statistics Applied to Radioactivity

ORA laboratories may be involved in the identification and quantitative measurement of radionuclides in foods, drugs, and the environment. Instrumentation varies from simple counters to solid state detectors that measure both discrete energy levels and the quantity of radiation in these samples. The correct application of statistical principles is important for arriving at the correct analytical result that will support regulatory decisions.

4.8.1 Introduction

Statistics is directly and intimately involved in measurements of radioactivity. Whereas most measurements made in the ORA laboratory are based on variables which vary continuously, radioactivity measurements are based on the counting of discrete, random events. In this case, the normal distribution probability function is replaced by the Poisson distribution, and the associated statistical parameters (mean, standard deviation) are therefore expressed differently.

4.8.2 Sample Counting

Radioactive decay is a random process that is described quantitatively in statistical terms. Therefore repeatedly counting radioactive transformations in a sample under identical conditions will not necessarily result in identical values. The result of counting sample radiations is

$$\text{number of sample counts} = N_s.$$

The standard deviation of the sample counts, based on Poisson statistics, is

$$\text{standard deviation of sample counts} = \sigma_s = \sqrt{N_s}.$$

Noise originating in the background, also a random process, simultaneously generates counts that are indistinguishable from those originating in the sample, and therefore the total or gross counts observed from counting a sample include background counts,

$$\text{gross sample counts} = N_g = N_s + N_b$$

where $N_s = \text{sample counts}$, and
 $N_b = \text{background counts}$.

It follows that the counts due to sample radioactivity are obtained by subtracting the background noise count from the sample gross counts

$$N_s = N_g - N_b .$$

The counting rate due to sample radioactivity is

$$R_s = \frac{N_s}{t_s}$$

where $t_s = \text{sample counting interval}$.

The sample counting rate can also be expressed as

$$R_s = R_g - R_b = \frac{N_g}{t_g} - \frac{N_b}{t_b} ,$$

where $R_g = \text{gross sample counting rate}$,
 $R_b = \text{background counting rate}$,
 $t_g = \text{gross sample counting interval}$, and
 $t_b = \text{background counting interval}$.

4.8.3 Standard Deviation and Confidence Levels

The standard deviation is a measure of the dispersion of values of a random variable about the mean value. For a large number of measurements, 68 percent would be expected to lie within plus and minus one standard deviation of the mean of the measurements; 96 percent would occur within plus or minus two standard deviations.

The standard deviation of the sample counting rate, σ_{R_s} is given by

$$\sigma_{R_s} = \sqrt{\sigma_{R_g}^2 + \sigma_{R_b}^2} = \sqrt{\frac{R_g}{t_g} + \frac{R_b}{t_b}}$$

where σ_{R_g} = standard deviation of the gross sample counting rate , and
 σ_{R_b} = standard deviation of the background counting rate .

The sample rate plus or minus one standard deviation is reported as

$$R_s \pm \sigma_{R_s} = R_s \pm \sqrt{\frac{R_g}{t_g} + \frac{R_b}{t_b}} .$$

If a measured value is reported within the limits of one standard deviation, there is a 68 percent certainty, or 68 percent confidence level, that the true value of the measured quantity is between the given limits. In other words, there is a 68 percent certainty that the real value lies within the limits. If the value is reported at the 96 percent confidence level, the true value is within plus or minus two standard deviations of the reported value. Several confidence levels are tabulated below:

Confidence Level (%)	Number of Standard Deviations (σ_s)
90	1.645
95	1.960
96	2.0
99	2.58

Example. A sample counted for 100 seconds yields 2300 gross counts. The background measured under identical conditions yields 100 counts in 10 seconds. Calculate the sample counting rate (counts per second) and the standard deviation of the sample counting rate. Report the results at the 96% confidence level.

$$R_s = \frac{2300 \text{ counts}}{100 \text{ s}} - \frac{100 \text{ counts}}{10 \text{ s}} = 23 \text{ cps} - 10 \text{ cps} = 13 \text{ cps} .$$

$$\sigma_{R_s} = \sqrt{\frac{23}{100} + \frac{13}{10}} = 1.2 \text{ cps} .$$

$$R_s = 13 \text{ cps} \pm 2.4 \text{ cps}$$

4.8.4 Counting Rate and Activity

The sample counting rate is proportional to sample activity, and may be converted to radioactivity units using correction factors. These may include detector efficiency in units of counts per disintegration, chemical recovery fraction, fractional radiation yield, and others. The sample activity may be obtained from the counting rate as follows:

$$\text{sample activity} = \frac{R_s \pm \sigma_{R_s}}{\varepsilon \times r \times Y} = \frac{R_s \pm \sqrt{\frac{R_g}{t_g} + \frac{R_b}{t_b}}}{\varepsilon \times r \times Y}$$

where ε = detector efficiency ,
 r = chemical recovery , and
 Y = radiation yield .

Example. A Sr-89 sample, counted using a detector having a 50% beta-particle detection efficiency for Sr-89 (0.5 counts per Sr-89 disintegration which emits one beta particle per disintegration), yields 500 gross counts in 10 seconds. The background count was 100 counts in 60 seconds. The chemical recovery of strontium was 86%. Report the approximate activity in the sample at the 68% confidence level.

$$R_s \pm \sigma_{R_s} = \frac{500 \text{ counts}}{10 \text{ s}} - \frac{100 \text{ counts}}{60 \text{ s}} \pm \sqrt{\frac{50}{10} + \frac{1.7}{60}} \text{ cps} = 48.3 \pm 2.2 \text{ cps}$$

$$\text{Sample activity} = \frac{48.3 \pm 2.2}{0.5 \times 0.86} \text{ Bq} = 112.4 \pm 5.2 \text{ Bq}$$

where 1 Bq (Becquerel) = 1 disintegration/s.

4.9 Statistics Applied to Biological Assays

Biological assays are those carried out by dosing a biological test system (such as a rat or mouse) with the substance to be determined, and measuring a response. An example is the USP monograph for Menotropins. This biological extract contains Luteinizing Hormone (LH) and Follicle Stimulating Hormone (FSH), which have effects on the reproductive organs. The assay consists of dosing male (LH) and female (FSH) rats with menotropins and observing the effects (weight) on the seminal vesicles and ovaries respectively after a multiple day incubation time.

Although this type of assay will rarely be encountered in the ORA laboratory, biological assays are instructive in the statistical complexity encountered when dealing with highly variable systems such as live animals. The interpretation of results is complicated by the fact that the total

variance of a measurement includes a large variance due to the biological component. The analyst may also encounter these assays when on team inspections.

The subject of biological assays is addressed in General Chapter <111> of the USP, “Design and Analysis of Biological Assays,” where an extensive statistical treatment is developed, based on the Analysis of Variance (ANOVA). This is also one of the rare instances in the USP where rejection of “outlier data” is allowed, under strict statistical justification. The interested reader is referred to <111> for further information.

4.10 Statistics Applied to Microbiological Analysis

Several analyses used by ORA microbiologists call for the enumeration of microorganisms by statistical means. Two commonly used procedures for estimating the number of microorganisms in a product are the plate count and the *Most Probable Number* (MPN) tube methods. To avoid fictitious impression of precision and accuracy, only 2 significant figures are reported. Many regulatory decisions pertaining to microbial contamination or time-temperature abuse of food will be based upon the level of organisms present.

4.10.1 Introduction

Many microbiological analyses involve the counting of discrete events, for example plate and tube counts for microbial growth and isolated colonies. As in the case for radioactivity, the situation is one of random, discrete, and relatively improbable events (such as growth of a colony forming unit on an agar plate), and Poisson statistics apply.

4.10.2 Geometric Mean

In microbiological assays, because of the techniques used and the fact that biological systems are being measured, a variety of unique statistical situations arise. When determining, for example, the number of colony forming units on a plate from a large number of replicate inoculations, the data often does not correspond to the expected normal distribution. That is, if the frequency of a given number of colonies is plotted against the observed number of colonies, a *non-symmetrical* frequency distribution is observed (note that the normal distribution curve is completely symmetrical, centered about the arithmetic mean). Instead the distribution is *skewed*, or tailed at the higher end. This is attributed to the fact that bacterial counts tend to favor lower counts and disfavor extremely high counts. In this situation, the arithmetic mean is not the best statistical indicator; instead the *geometric mean* is most often used:

$$\bar{x}_g = \left(\prod_{i=1}^n x_i \right)^{1/n}$$

where x_i are the individual counts, and \prod indicates that the *product* of the observations is determined rather than the sum.

For example, the arithmetic mean of the individual observations 1, 2 and 3 is:

$$\frac{(1+2+3)}{3} = 2$$

whereas the geometric mean of the same observations is:

$$(1 \times 2 \times 3)^{1/3} = 1.8$$

Question: Why would one expect lower plate counts to be more probable than higher counts, thus causing a skewed probability distribution? *Answer:* As the number of counts on a plate rises, in other words the density of colonies rises, an overcrowding error occurs from individual colonies inhibiting the formation of other colonies nearby. Another factor appears to be a “counting fatigue” error at high numbers, where the analyst may not count accurately because of the large numbers involved.

An alternative way to calculate the geometric mean, which can be easily derived from the product expression above, is to add the logarithms of individual counts rather than form the product of the counts themselves. The geometric mean is then defined as:

$$\bar{x}_g = \text{anti log} \left(\frac{\sum_{i=1}^n \log x_i}{n} \right)$$

This formula is much easier for calculation purposes, particularly when a large number of observations are involved.

4.10.3 Most Probable Number

Another statistical concept unique to microbiological observations is that of Most Probable Number (MPN). The Most Probable Number is a statistically derived estimate of the presence of microorganisms based on the presence or absence of growth in serially diluted samples. After an initial dilution, serial dilutions of the sample are made (for example, 1:10, 1:100, and 1:1000) with a number of replicate tubes (for example, 3 or 5) at each dilution. After incubation, the

presence or absence of growth in each tube is tabulated. The resulting code (number of positive tubes) is compared with published tables to give the most probable number of microorganisms per unit of original, undiluted sample. Most probable number tables are published for various numbers of tubes at a number of dilutions. The statistical derivation is beyond the scope of this discussion, but is based on Poisson counting statistics. Tables are published in the *Bacteriological Analytical Manual* (BAM), the *AOAC Official Methods of Analysis*, and General Chapter <61> of the USP.

4.10.4 References

(Current Ed.). "Microbial Limits Tests <61>," *U. S. Pharmacopeia and national formulary*. Rockville, MD: United States Pharmacopeial Convention, Inc.

Tomlinson, L. (Ed.). (1998). *Bacteriological analytical manual* (8th ed., Rev. A, in hardcopy). Washington DC: R. I. Merker, Ph.D., Office of Special Research Skills, Center for Food Safety and Applied Nutrition, U.S. Food & Drug Administration. The current updated version is available online at: <http://vm.cfsan.fda.gov/~ebam>.

4.11 Statistics Applied to pH in Canned Foods

pH is a logarithmic measure for the acidity of an aqueous solution. Since pH represents the negative logarithm of a number, it is not mathematically correct to calculate simple averages or other summary statistics. Instead, the values should be converted to hydrogen ion concentrations, averaged, and re-converted to pH values. The following guidance is provided.

1. Convert each pH value to hydrogen-ion activity (H^+), using the equation:

$$\text{Activity} = 10^{-\text{pH}}$$

In Excel, the formula is: $=10^{-(\text{pH number})}$

2. Calculate the mean of the activity values by adding the values and dividing the sum by the total number of values. Calculate the standard deviation also from the activity values.

3. Convert the calculated mean activity back to pH units, using the equation:
 $\text{pH} = -(\log_{10})(\text{mean } H^+ \text{ activity})$. Also convert the standard deviation to pH units.
In Excel, the formula is: $= -\text{LOG10}(\text{number})$

When the pH values correspond closely, there is not a significant difference between the mathematical mean and the logarithmic mean. As the pH values spread further apart from each other, the difference between the two means become more significant.

4.12 Document History

Version 1.4 Revision Approved: 01-30-13 Author: LMEB Approver: LMEB

Version 1.2 changes:

Contents – Document History added.

Section 4.3 – title corrected to Data Handling and Presentation.

Section 4.3.1. – number of significant digits to report changed to 1 for 0.0396.

Section 4.10.4 – removed date from BAM reference.

Version 1.3 changes:

Section 4.4, fourth paragraph – corrected to coefficient of determination and revised.

Version 1.4 changes:

Header – Division of Field Science changed to Office of Regulatory Science

4.11 – Section added