

OAK RIDGE  
NATIONAL LABORATORY

---

MANAGED BY UT-BATTELLE  
FOR THE DEPARTMENT OF ENERGY



ORNL-27 (4-00)

## DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge.

**Web site** <http://www.osti.gov/bridge>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source.

National Technical Information Service  
5285 Port Royal Road  
Springfield, VA 22161  
**Telephone** 703-605-6000 (1-800-553-6847)  
**TDD** 703-487-4639  
**Fax** 703-605-6900  
**E-mail** [info@ntis.fedworld.gov](mailto:info@ntis.fedworld.gov)  
**Web site** <http://www.ntis.gov/support/ordernowabout.htm>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE) representatives, and International Nuclear Information System (INIS) representatives from the following source.

Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831  
**Telephone** 865-576-8401  
**Fax** 865-576-5728  
**E-mail** [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)  
**Web site** <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

# **Statistical Methods for Material Characterization and Qualification**

*Andrew K. Kercher and John D. Hunn  
Oak Ridge National Laboratory*

January 26, 2005

Table of Contents:

1	Introduction.....	1
1.1	Purpose.....	1
1.2	Fundamental concepts.....	1
2	Attribute sampling .....	2
2.1	Binomial distribution methods for attribute sampling.....	3
2.1.1	Binomial distribution (BD) test .....	4
2.1.2	Binomial distribution tolerance limit (BDTL) approximation.....	5
2.1.3	Binomial distribution control limit (BDCL) approximation.....	6
2.2	Standard deviation methods for attribute sampling .....	8
2.2.1	Standard deviation (SD) test .....	9
2.2.2	Standard deviation tolerance limit (SDTL) approximation .....	11
2.2.3	Standard deviation control limit (SDCL) approximation .....	14
2.3	Comparison of attribute sampling methods .....	17
3	Variable sampling.....	19
3.1	Adequacy of the normal distribution assumption .....	19
3.2	Variable sampling based on a mean (VSM) .....	21
3.3	Variable sampling based on a distribution (VSD).....	24
	Appendix A. Guide for suggested use of statistical methods for attribute sampling	
	Appendix B. Derivation of the rule of thumb for “acceptable” use of SD methods	
	Appendix C. Derivation of the standard deviation tolerance limit (SDTL) approximation	
	Appendix D. Derivation of the standard deviation control limit (SDCL) approximation	
	Appendix E. z table (for one-tailed testing)	
	Appendix F. t table	
	Appendix G. $K_1$ table (one-tailed tolerance factor)	
	Appendix H. $K_2$ table (two-tailed tolerance factor)	

*List of Figures*

Figure 2-1: Confidence intervals from binomial and Poisson distributions. A Poisson distribution closely approximates a binomial distribution for a sufficiently large sample ( $n$ ) and sufficiently small tolerance limit ( $L_t$ )..... 4

Figure 2-2: Comparisons of standard deviation (SD) and binomial distribution (BD) methods for various sample sizes and tolerance levels. Data is shown for different values of the product of the sample size ( $n$ ) and the tolerance level ( $L_t$ ). The data applies to sample sizes up to 50 million and tolerance levels up to 2% (0.02). The BD methods curve for  $nL_t = 50$  was computed for up to  $n = 50$  million for confidence level  $> 0.90$ , but was only computed for up to  $n = 1.25$  million for lower confidence levels..... 9

Figure 2-3: For control limits, the true significance in differences between lots is not graphically intuitive. The significance comes from the number of defective items that occur inside the difference in confidence intervals. (Results for two hypothetical lots are shown.)..... 18

Figure 3-1: Examples of various sources of non-normality..... 20

*List of Tables*

Table 2-1: Example of an acceptance criterion based on an attribute for the AGR Fuel Program ..... 2

Table 2-2: Traditional and alternate statistical methods for attribute sampling. (italics denote alternate statistical methods)..... 3

Table 2-3: Tolerance limit “guesses” for BDTL approximation example ..... 6

Table 2-4: Confidence levels for different numbers of defective items ( $n_d$ ) for the BDCL approximation example..... 7

Table 2-5: A table of actual sphericity data for the BDCL approximation example. An ordered list of the 42 highest sphericities are listed for a sample of 2825 kernels. .... 7

Table 2-6: Data taken from a z table with associated confidence level (one-tailed). The z statistic is given to three decimal places for key confidence levels. .... 11

Table 2-7: A table of actual sphericity data for the SDCL approximation example. An ordered list of the 42 highest sphericities are listed for a sample of 2825 kernels. (same table as BDCL example) ..... 16

Table 2-8: Summary of the six methods for analyzing attribute sampling data. Calculated values are in bold type. The percent error caused by using standard deviation methods (instead of binomial distribution methods) is shown for all calculated values. .... 18

Table 3-1: Example of acceptance criterion for variable sampling based on a mean. .... 23

Table 3-2: Data taken from a t table ..... 23

Table 3-3: Example of acceptance criteria for variable sampling based on a distribution ..... 25

Table 3-4: Data from a tolerance factor table for one-tailed tests (Appendix G) ..... 25

*List of symbols:*

$l_T$	computed tolerance limit based on measured sample
$n$	sample size
$n_d$	number of defects in the measured sample
$p_p$	predicted defect fraction in the lot
$p_m$	defect fraction of the measured sample ( $p_m = n_d/n$ )
$t$	t statistic (taken from a t table)
$\bar{x}$	mean of the sample
$z$	z statistic (taken from a z table)
$N_d$	the average number of defects which would occur in hypothetical random samples of n items
$K$	tolerance factor (VSD method)
$L_c$	computed confidence level
$L_t$	tolerance limit specified in acceptance criterion
$P$	probability
$R_t$	range containing all but the tolerance limit of the sample to the specified confidence level (VSD method)
$T$	number of tails
$\alpha$	a probability associated with a rejected hypothesis
$\mu_p$	predicted mean of the lot
$\nu$	degrees of freedom
$\sigma_a$	attribute standard deviation
$\sigma_p$	standard deviation of the defect fraction
$\sigma_x$	standard deviation of the sample (variable sampling)

*List of abbreviations:*

AGR	Advanced Gas Reactor
APC	attribute property cut-off
BD	binomial distribution
BDCL	binomial distribution control limit (approximation)
BDTL	binomial distribution tolerance limit (approximation)

CL	control limit
ORNL	Oak Ridge National Laboratory
SD	standard deviation
SDCL	standard deviation control limit (approximation)
SDTL	standard deviation tolerance limit (approximation)
TL	tolerance limit
VSD	variable sampling based on a distribution
VSM	variable sampling based on a mean

# 1 Introduction

## 1.1 Purpose

This document describes a suite of statistical methods that can be used to infer lot parameters from the data obtained from inspection/testing of random samples taken from that lot. Some of these methods will be needed to perform the statistical acceptance tests required by the Advanced Gas Reactor Fuel Development and Qualification (AGR) Program. Special focus has been placed on proper interpretation of acceptance criteria and unambiguous methods of reporting the statistical results. In addition, modified statistical methods are described that can provide valuable measures of quality for different lots of material. This document has been written for use as a reference and a guide for performing these statistical calculations. Examples of each method are provided. Uncertainty analysis (e.g., measurement uncertainty due to instrumental bias) is not included in this document, but should be considered when reporting statistical results.

## 1.2 Fundamental concepts

While this statistical guide has been written to be as generally applicable as possible, it is necessary to define several terms to avoid confusion when using this guide.

A lot refers to the entire amount of a specified material. A sample is a random subset of the entire lot. An item is a single unit in a sample. Every item in a sample is measured for a property that will be used in statistical analysis. Properties can be attributes or variables. An attribute property is a characteristic that classifies an item as either defective or nondefective. A variable property describes a characteristic of an item with a number. Variable properties are quantitative.

The distinction between attributes and variables can be quite confusing. An attribute property of an item is always “defective” or “nondefective.” The defective or nondefective decision for attribute properties can be based on qualitative features (e.g. color) or can be based on numeric cut-offs, but the decision is made for each individual item. Variable properties always have numeric values for each item.

Attribute sampling is statistical analysis using attributes and is based on the fraction of items in a sample that is declared defective. The fraction of defective items in the entire lot of material is statistically calculated based on the fraction of defective items in the sample measured. The fraction of defective items is used to determine whether the lot passes the acceptance criterion.

Variable sampling is statistical analysis using variables and is based on computing values (e.g. average and standard deviation) that describe a measured sample. The computed values that describe the sample are used to statistically calculate probable ranges for the entire lot. The calculated probable range is used to determine whether a lot passes the acceptance criterion.

## 2 Attribute sampling

Attribute sampling is based on the fraction of a sample that is declared defective. In acceptance testing for an attribute property, a lot is acceptable if there is a sufficiently high probability (i.e., confidence level) that the fraction of defective material in the lot is below the tolerance limit, where “defective” material is defined by the control limit. The control limit defines how a particle is declared “defective” or “nondefective.” The tolerance limit states how much material in the lot can fall outside the control limit (how much can be declared “defective”) without rejecting the lot. Because only a sample of material is measured, the measured fraction of defective material usually differs from the actual fraction of defective material in the whole lot. The actual fraction of defective material is statistically calculated from the measured fraction to a given confidence level.

If a tolerance limit is stated, the statistical tests for attribute sampling are ALWAYS “one-tailed” even if there are multiple control limits.

**Table 2-1: Example of an acceptance criterion based on an attribute for the AGR Fuel Program**

Property	Control Limit	Tolerance Limit ( $L_t$ )	Confidence Level
Kernel Sphericity	$\leq 1.05$	0.01	0.95

As shown in Table 2-1, a proper acceptance criterion for attribute sampling specifies a control limit, a tolerance limit, and a confidence level. In traditional statistical analysis, the statistician calculates the confidence level that the lot contains less than the tolerance limit of units as defined “defective” by the control limit, and then the computed confidence level is compared to the confidence level of the acceptance criterion. Effectively, traditional statistical methods have a defined control limit, a fixed tolerance limit, and a computed confidence level that has to pass the acceptance criterion. Traditional statistical methods are tests used for material qualification. Often, a sampling plan is generated from statistical techniques for attribute sampling (such as the BDTL or SDTL approximation described later in this section) using *a priori* knowledge about the material. A sampling plan simplifies statistical data analysis by requiring a specific sample size and by defining the maximum number of defects associated with passing the acceptance criterion for that specific sample size.

As shown in Table 2-2, there are two fundamental statistical techniques for attribute sampling, herein termed: Binomial Distribution (BD) and Standard Deviation (SD). Traditional statistical methods for attribute sampling can test a lot against a given acceptance criterion, but they do not clearly indicate how good a product is. Critical control limits and tolerance limits of materials often need to be determined. Statistical approximations for attribute sampling are included in this guide to estimate the control limit and tolerance limit for which a given lot of material would be deemed acceptable (shown with italics in Table 2-2). A guide for choosing the appropriate statistical method is given in Appendix A. Details of each method will be discussed in its specific section.

**Table 2-2: Traditional and alternate statistical methods for attribute sampling. (italics denote alternate statistical methods)**

Fundamental Technique	Statistical Method	Control Limit	Tolerance Limit	Confidence Level
Binomial Distribution	binomial distribution (BD) test	defined	fixed	compared
	<i>tolerance limit (BDTL) approximation</i>	<i>defined</i>	<i>determined/compared</i>	<i>fixed</i>
	<i>control limit (BDCL) approximation</i>	<i>determined/compared</i>	<i>fixed</i>	<i>fixed</i>
Standard Deviation	standard deviation (SD) test	defined	fixed	compared
	<i>tolerance limit (SDTL) approximation</i>	<i>defined</i>	<i>determined/compared</i>	<i>fixed</i>
	<i>control limit (SDCL) approximation</i>	<i>determined/compared</i>	<i>fixed</i>	<i>fixed</i>

### 2.1 Binomial distribution methods for attribute sampling

Binomial distribution methods are based on simple probability theory. A binomial distribution equation can be used to determine the probability of having a specific number of defective items in a sample.<sup>1</sup> For binomial distribution methods for attribute sampling, a sum of binomial distribution equations is used to determine the probability (i.e., confidence level) of having less than a specific fraction (i.e., tolerance limit) of defective items in a lot (where “defective” items are defined by the control limit). The fundamental equation for the binomial distribution methods is Eq. 2-1:<sup>1</sup>

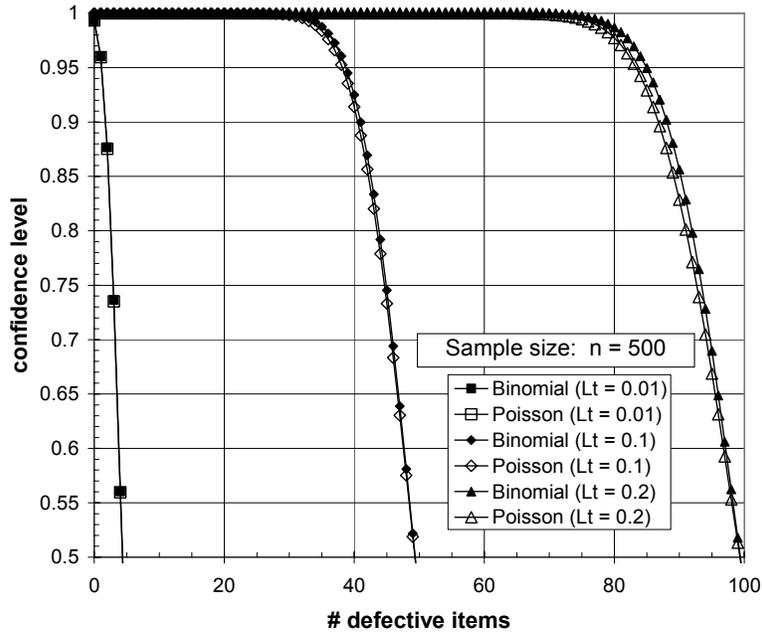
$$L_c = 1 - P(0 \leq i \leq n_d) = 1 - \sum_{i=0}^{n_d} \frac{n!}{i!(n-i)!} L_t^i (1-L_t)^{n-i} \quad (2-1)$$

where  $L_c$  is the confidence level,  $n$  is the total number of items in the sample,  $n_d$  is the number of defective items in the sample,  $P$  is the probability of having less than  $n_d$  defective items, and  $L_t$  is the tolerance limit.

For very large samples, the fundamental equation for binomial distribution methods can be impossible to solve exactly because of the resultant large factorial and exponential terms. Depending on the tolerance level, samples with millions or billions of items can be solved exactly by binomial distribution methods (if like terms in the factorials are cancelled) with conventional spreadsheet or mathematics software. For commonly used tolerance limits (<10%) and sufficiently large sample sizes (>100), the Poisson distribution can be used as a close approximation to the binomial distribution (see Eq. 2-2 and Figure 2-1). The Poisson approximation for a binomial distribution becomes increasingly accurate for larger sample sizes and smaller tolerance limits.<sup>2</sup> Compared to the binomial distribution equation, the equation for the Poisson distribution

is slightly less computationally cumbersome and can readily handle very large sample sizes.

$$L_c = 1 - \sum_{i=0}^{n_d} \frac{(nL_t)^i}{i!} \exp(-nL_t) \quad (2-2)$$



**Figure 2-1: Confidence intervals from binomial and Poisson distributions. A Poisson distribution closely approximates a binomial distribution for a sufficiently large sample (n) and sufficiently small tolerance limit (Lt).**

The greatest problem with binomial distribution methods (even when using the Poisson approximation) is that the fundamental equation is time consuming to solve, *especially for the statistical approximations (BDTL & BDCL)*. Using the fundamental equation, a confidence level can be computed from a fixed control limit and a fixed tolerance limit. The equation cannot be rearranged in order to directly solve for tolerance limit or to determine control limits. Complex software, spreadsheets, or tables can be generated for the statistical approximations, if desired. Modern statistical software can potentially be used to simplify the use of binomial distribution methods. Depending on the desired accuracy, the standard deviation methods (section 2.2) can be used to more easily calculate desired statistical results.

### 2.1.1 Binomial distribution (BD) test

The binomial distribution test directly calculates a confidence level for a fixed tolerance limit based on a defined control limit and then compares the calculated confidence level to the specified confidence level to determine whether a lot passes the acceptance criterion. If the calculated confidence level is greater than the specified

confidence level, then the lot passes the acceptance criterion. When possible, Eq. 2-1 should be used to exactly calculate the confidence level.

BD Test Method:

Directly use Eq. 2-1 to calculate the confidence level for a given sample size and measured number of defects (based on the control limit). Compare the calculated confidence level with the confidence level required in the specification to determine whether the lot passes.

BD Test Example:

This example is based on the acceptance criterion from Table 2-1. Assume a sample of 2825 kernels. A kernel is declared “defective” if it has a sphericity greater than 1.05. Assume that 12 kernels in the sample are measured to have a sphericity greater than 1.05.

To calculate the confidence level, simply insert the values into Eq. 2-1 to obtain:

$$L_c = 1 - \sum_{i=0}^{12} \frac{(2825)!}{i!(2825-i)!} (0.01)^i (0.99)^{2825-i} = 0.9995$$

The calculated confidence level (0.9995) is greater than the specified confidence level (0.95), so the lot passes the acceptance criterion. In the BD test, the calculated confidence level should be reported.

*2.1.2 Binomial distribution tolerance limit (BDTL) approximation*

The binomial distribution tolerance limit (BDTL) approximation provides a relative measure of lot quality. The tolerance limit that would allow the lot to be accepted is estimated based on a fixed definition of a “defective” item (i.e., the control limit) and a fixed confidence level. The BDTL approximation can determine, to a fixed confidence level, what fraction of a lot is defective. The BDTL approximation can be best used to compare the defective content of different lots when the pertinent definition of “defective” (i.e., the control limit) is established.

BDTL approximation method:

The BDTL approximation is based on Eq. 2-1, but it is not possible to obtain an exact, general solution for tolerance limit from Eq. 2-1. With current spreadsheets or mathematics software, it’s relatively straightforward to guess and check to obtain the tolerance limit that would have the fixed confidence level for the measured number of defective items. An example of performing a BDTL approximation with a “guess-and-check” method is given below. Since confidence levels of practical acceptance criteria are greater than 0.75, the tolerance limit must be greater than the measured fraction of defective items ( $p_m = n_d/n$ ).

BDTL approximation example:

This example is based on the acceptance criterion from Table 2-1. Assume a sample of 2825 kernels. A kernel is declared “defective” if it has a sphericity greater than

1.05. Assume that 12 kernels in the sample are measured to have a sphericity greater than 1.05. Eq. 2-3 is used to generate confidence levels for different “guesses” for tolerance level ( $l_T$ ), as shown in Table 2-3. The goal is to obtain a confidence level of 0.95 for twelve “defective” kernels ( $n_d = 12$ ). As a starting “guess”, a value slightly above the measured fraction of defective kernels (0.004248) was chosen.

$$L_c = 1 - \sum_{i=0}^{12} \frac{(2825)!}{i!(2825-i)!} l_T^i (1-l_T)^{2825-i} \quad (2-3)$$

**Table 2-3: Tolerance limit “guesses” for BDTL approximation example**

	Tolerance Limit ( $l_T$ ) “Guess”	Confidence Level	Comment
Guess 1	0.005	0.654	guess for $l_T$ is too low
Guess 2	0.0075	0.978	guess for $l_T$ is too high
Guess 3	0.007	0.957	guess for $l_T$ is slightly too high
Guess 4	0.00675	0.942	guess for $l_T$ is slightly too low
Guess 5	0.00687	0.950	This is the appropriate tolerance limit ( $l_T$ )

Based on the fixed control limit (sphericity  $\leq 1.05$ ), the estimated tolerance limit for the lot is 0.00687 to a 95% confidence level (i.e., 0.95). Although the BDTL approximation is intended for estimating tolerance limits, it also shows that the lot passes the acceptance criterion, because the calculated tolerance level ( $l_T = 0.00687$ ) is lower than the specified tolerance level ( $L_t = 0.01$ ).

### 2.1.3 Binomial distribution control limit (BDCL) approximation

The binomial distribution control limit (BDCL) approximation can be used to estimate the control limit for which the lot would be accepted given a fixed tolerance limit and confidence level. The BDCL approximation can only be used for quantitative measurements. The details of the method and an example follow.

#### BDCL approximation method:

Calculate the confidence level from Eq. 2-1 for the number of defective items ( $n_d$ ) varying from zero to about  $L_t n$ . From this table of values, find the  $n_d$  closest to the fixed confidence level without going under, which will be called the “critical  $n_d$ ”. For the measured sample, determine the control limit that would declare no more than the “critical  $n_d$ ” number of items as “defective.” Be sure not to round numbers or choose numbers so that additional items would be declared “defective.”

BDCL approximation example:

This example is based on the acceptance criterion from Table 2-1. Assume a sample of 2825 kernels. For the given sample size (2825) and fixed tolerance level, Eq. 2-1 is used to generate confidence levels for different numbers of defective items ( $n_d$ ), as shown in Table 2-4.

**Table 2-4: Confidence levels for different numbers of defective items ( $n_d$ ) for the BDCL approximation example.**

$n_d$	confidence level	$n_d$	confidence level
$\leq 9$	1	17	0.9843
10	0.9999	18	0.9735
11	0.9998	19	0.9572
12	0.9995	20	0.9342
13	0.9989	21	0.9032
14	0.9977	22	0.8633
15	0.9954	23	0.8141
16	0.9912	24	0.7561

An  $n_d$  of 19, which has a confidence level of 0.9572, is the closest to the fixed confidence level (0.95) without going under. Thus, the critical  $n_d$  is 19.

**Table 2-5: A table of actual sphericity data for the BDCL approximation example. An ordered list of the 42 highest sphericities are listed for a sample of 2825 kernels.**

1.217	1.059	1.049	1.040	1.037	1.033	1.031
1.148	1.057	1.047	1.040	1.037	1.033	1.030
1.092	1.057	1.046	1.039	1.036	1.033	1.030
1.075	1.057	1.044	1.039	1.036	1.032	1.029
1.072	1.056	1.043	1.039	1.035	1.031	1.029
1.060	1.053	1.043	1.038	1.035	1.031	1.028

Table 2-5 contains an ordered list of the 42 highest sphericities from the 2825 sphericity measurements used for this example. By looking at Table 2-5, the control limit would need to be set at 1.041 for no more than 19 kernels to be declared “defective.”

The lot would pass an estimated control limit of 1.041 at a tolerance limit of 1% (i.e., 0.01) to a 95% confidence level. Although the BDCL approximation is intended for estimating control limits, it also shows that the lot passes the acceptance criterion, because the calculated control limit (1.041) is lower than the specified control limit (1.05).

## 2.2 Standard deviation methods for attribute sampling

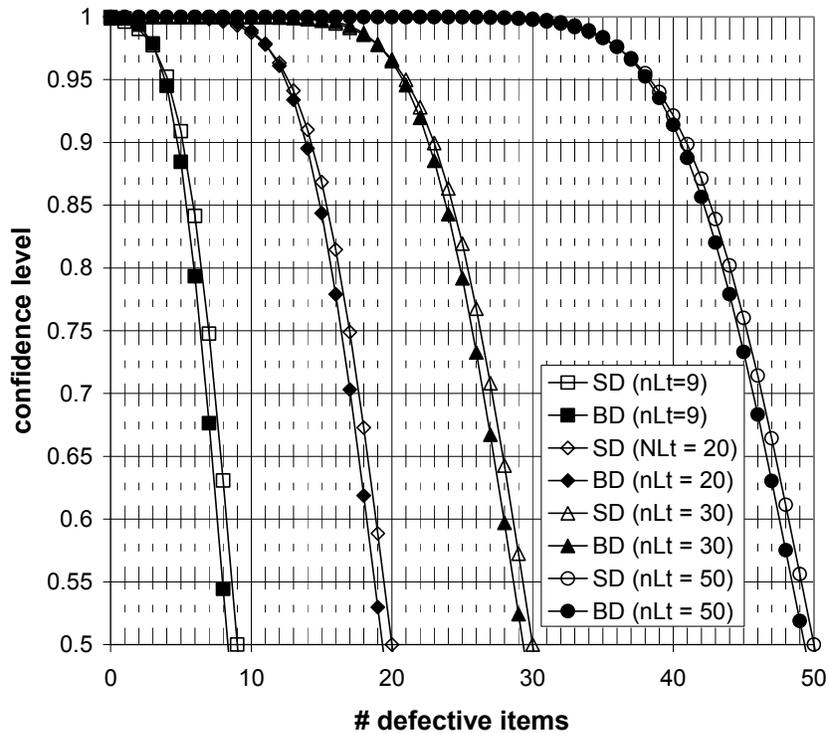
Standard deviation methods for attribute sampling use common statistical methods to determine the probability (i.e., confidence level) of having less than a specific fraction (i.e., tolerance limit) of defective items in a lot (where “defective” is defined by the control limit). Confidence levels are based on the assumption that measuring several samples for the fraction of defective items would produce results with a normal distribution.

Standard deviation methods for attribute sampling are based on the Central Limit Theorem,<sup>1</sup> which states that the frequency distribution for the means (averages) of random large samples is approximately normal. Invoking the Central Limit Theorem for a sample with a binomial distribution introduces an error in reported values. The error depends on: (1) the confidence level and (2) the relative magnitude of the sample size and the tolerance limit (Figure 2-2). For common tolerance limits (< 2%), large sample sizes (relative to the tolerance limit) with high confidence levels ( $\geq 95\%$ ) have reasonably good agreement between the standard deviation methods and the binomial distribution methods. If the tolerance limit is low and the sample size is not sufficiently large, the standard deviation methods are not considered acceptable to use. As a conservative rule of thumb, the standard deviation methods are considered acceptable if the following condition is satisfied (see Appendix B for derivation):

$$n \geq \frac{9}{L_t} \quad (2-4)$$

where  $n$  is the sample size and  $L_t$  is the tolerance limit (in decimal form). Even though the standard deviations methods may be considered “acceptable”, the error in reported values may not be tolerable for a given application. All of the cases shown in Figure 2-2 are “acceptable” for standard deviation methods, but significant error can be observed for certain ranges of confidence level.

Despite these limitations, standard deviation methods are commonly used for media polls and industrial quality control. The key advantages to standard deviation methods are their ease of use and their intuitive feel. The pertinent equations are simple to solve. A novice in statistics can quickly gain an intuitive feel for the underlying logic of standard deviation methods and the meaning of the statistical results. Standard deviation methods can be easily applied and tweaked to provide insightful interpretation for attribute sampling. Standard deviation methods should not be nonchalantly used because of their ease and intuitive feel; a brief check of expected error should be performed before their use.



**Figure 2-2: Comparisons of standard deviation (SD) and binomial distribution (BD) methods for various sample sizes and tolerance levels. Data is shown for different values of the product of the sample size ( $n$ ) and the tolerance level ( $L_t$ ). The data applies to sample sizes up to 50 million and tolerance levels up to 2% (0.02). The BD methods curve for  $nL_t = 50$  was computed for up to  $n = 50$  million for confidence level  $> 0.90$ , but was only computed for up to  $n = 1.25$  million for lower confidence levels.**

### 2.2.1 Standard deviation (SD) test

Based on a fixed control limit and a fixed tolerance limit, the standard deviation (SD) test calculates a confidence level, which is compared to the specified confidence level to determine whether the lot passes the acceptance criterion. The SD test is widely used, simple, and straightforward.<sup>1</sup> Note from Figure 2-2 that, for a given number of defective items, SD tests often have only slight errors ( $< 0.01$ ) at high confidence levels ( $\geq 0.95$ ) and relatively large errors (up to 0.04-0.08 too high) for low confidence levels. Often, the confidence level error for the SD test (as compared to the BD test) will not significantly influence whether a lot passes the acceptance criterion, but the error can make comparisons between confidence levels of different lots difficult.

#### SD test method:

First, it must be determined if the sample size is sufficiently large. Calculate the attribute standard deviation ( $\sigma_a$ ) using:<sup>1</sup>

$$\sigma_a = \sqrt{\frac{L_t(1-L_t)}{n}} \quad (2-5)$$

where  $L_t$  is the tolerance limit and  $n$  is the total number of items in the sample. Check to see if the tolerance limit  $\pm 3$  attribute standard deviations is contained within the range of 0 to 1. If not, then the sample is not sufficiently large to perform the SD test and the sample should be analyzed with the traditional binomial distribution method.

Next, calculate the “z statistic” by the following equation.<sup>1</sup>

$$z = \frac{|L_t - p_m|}{\sigma_a} \quad (2-6)$$

where  $p_m$  is the measured fraction of defective items. Use the z statistic to look up a value on the z table in Appendix E. If  $p_m < L_t$ , then the z table value is the confidence level. If  $p_m > L_t$ , then the confidence level equals 1 minus the z table value. There is no standard format for z tables, so take care to read the z table properly. If the confidence level from a z table is greater than the confidence level of the acceptance criterion, then the lot passes the acceptance criterion.

#### SD test example:

This example is based on the acceptance criterion from Table 2-1. Assume a sample of 2825 kernels. A kernel is declared “defective” if it has a sphericity greater than 1.05. Assume that 12 kernels have a sphericity greater than 1.05.

First, determine if the sample size is sufficiently large. The attribute standard deviation is:

$$\sigma_a = \sqrt{\frac{(0.01)(0.99)}{2825}} = 0.001872$$

The tolerance limit  $\pm 3$  attribute standard deviations is 0.004384 to 0.01562. Since it is within the range of 0 to 1, the sample is considered large enough to use the SD test.

Next, calculate the z statistic. The measured fraction of defective kernels ( $p_m$ ) is:  $12/2825 = 0.004248$ . Thus, the z statistic is:

$$z = \frac{|(0.01) - (0.004248)|}{0.001872} = 3.073$$

Table 2-6 shows some data off of a z table (for full z table see Appendix E). The pertinent value from the z table is 0.9989. Since  $p_m < L_t$ , then the z table value is the confidence level.

**Table 2-6: Data taken from a z table with associated confidence level (one-tailed). The z statistic is given to three decimal places for key confidence levels.**

z statistic	z table value	Confidence level for $p_m < L_t$	Confidence level for $p_m > L_t$
0.674	0.750	0.750	0.250
1.282	0.900	0.900	0.100
1.645	0.950	0.950	0.050
2.326	0.990	0.990	0.010
3.07	0.9989	0.9989	0.0011
3.090	0.9990	0.999	0.0010

The lot passes the acceptance criterion, because the calculated confidence level (0.9989) is greater than the confidence level in the acceptance criterion (0.95). In SD tests, the calculated confidence level is usually reported.

### 2.2.2 Standard deviation tolerance limit (SDTL) approximation

The standard deviation tolerance limit (SDTL) approximation can be best used to estimate the defective content of different lots when the pertinent definition of “defective” is established. The SDTL approximation can provide two measures of lot quality. First, the SDTL approximation can calculate the greatest tolerance limit that the lot would pass for the specified control limit and confidence level, based on the measured sample. Second, the SDTL approximation can provide a confidence level for the defect fraction of the lot, based on the control limit.

Standard deviation methods inherently have associated error (compared to binomial distribution methods), but error is often small when the confidence level is high (see Figure 2-2). For most common acceptance criteria, the specified confidence level is high, so the SDTL approximation often is not significantly different from the BDTL approximation. The confidence interval for the defect fraction relies on a t table, which is derived based on a normal distribution. Slight errors in confidence interval would be caused by using a t table on a binomial distribution.

#### SDTL approximation method:

There are two separate results that can be obtained from the SDTL approximation: (a) the greatest tolerance limit that the lot would pass for the given control limit and confidence level according to the measured sample and (b) the confidence interval of the defect fraction of the lot.

(A) Greatest tolerance limit that would be passed

From a z table, look up the z statistic (z) that corresponds to the confidence level in the acceptance criterion. There is no standard format for z tables, so take care to read a z table properly. See Appendix E for a full z table.

Determine the fraction of defective items ( $p_m$ ). Appendix C derives the following equation, which is used to calculate the tolerance limit ( $l_T$ ) for a fixed control limit and confidence level.

$$l_T = \frac{2np_m + z^2 + z\sqrt{z^2 + 4np_m - 4np_m^2}}{2z^2 + 2n} \quad (2-7)$$

where n is the total number of items in the sample. *Then, the test must be proved valid.* A test is necessary to prove that the sample size was large enough for SDTL approximation to be valid.

In order to determine if the sample was sufficiently large, perform the following test. Calculate the attribute standard deviation using Eq. 2-5:

$$\sigma_a = \sqrt{\frac{l_T(1-l_T)}{n}}$$

Check to see if the calculated tolerance limit  $\pm 3$  attribute standard deviations is entirely contained between 0 and 1. If not, then the sample is not sufficiently large to have a valid result from the test, so the binomial distribution tolerance limit approximation should be used.

(B) Confidence interval for the defect fraction

The standard deviation of the defect fraction ( $\sigma_p$ ) is calculated using:

$$\sigma_p = \sqrt{\frac{p_m(1-p_m)}{n}} \quad (2-8)$$

Choose a confidence level for the confidence interval that is no greater than 99%. Use a t table to determine the t value associated with the sample size and chosen confidence level (see Appendix F for the t table). Often, values on a t table are found using v (the degrees of freedom) and  $\alpha$ . The v equals the sample size minus one. To get  $\alpha$ , use the following equation:

$$\alpha = \frac{1-L_c}{2} \quad (2-9)$$

where  $L_c$  is the confidence level of the confidence interval. Calculate the confidence interval for the fraction of defective items ( $p_p$ ) in the lot using:<sup>1</sup>

$$p_m - t\sigma_p \leq p_p \leq p_m + t\sigma_p \quad (2-10)$$

Check to see if the sample is sufficiently large to specify a valid confidence interval. Using the same procedure as for your chosen confidence level, calculate the 99.8% confidence interval for the fraction of defective items. If the 99.8% confidence interval is between 0 and 1, then the sample is sufficiently large to specify a valid confidence interval.

SDTL approximation example:

This example is based on the acceptance criterion from Table 2-1. Assume a sample of 2825 kernels. A kernel is declared “defective” if it has a sphericity greater than 1.05. Assume that 12 kernels have a sphericity greater than 1.05.

(A) Greatest tolerance limit that would be passed

From a z table (see Table 2-6), the fixed confidence level of 0.95 (i.e., 95%) can be found to correspond to a z statistic of 1.645. The fraction of defective items ( $p_m$ ) is:  $12/2825 = 0.004248$ . Using the assorted constants, Eq. 2-7 becomes:

$$l_T = \frac{(26.71) + (1.645)\sqrt{(50.50)}}{5655}$$

Thus, the solution for the tolerance limit ( $l_T$ ) is: 0.006790.

The SDTL approximation must be tested for validity. The attribute standard deviation is computed to be:

$$\sigma_a = \sqrt{\frac{(0.006790)(0.99321)}{2825}} = 0.001545$$

The tolerance limit  $\pm 3$  attribute standard deviations gives the range: 0.002155 - 0.01143. This range is entirely contained between 0 and 1, so the SDTL approximation is proven valid for this sample.

Therefore, the sample would pass a tolerance limit of 0.00679 to a 95% confidence level with the specified control limit (1.05). Although the SDTL approximation is intended for estimating tolerance limits, it also shows that the lot passes the acceptance criterion, because the calculated tolerance level (0.00679) is lower than the specified tolerance level (0.01).

(B) Confidence interval for the defect fraction

The standard deviation of the defect fraction is calculated to be:

$$\sigma_p = \sqrt{\frac{0.004248(1 - 0.004248)}{2825}} = 0.001224$$

Confidence intervals are often reported for a 95% confidence, so a 95% confidence interval will be determined for this example. From a t table (see Appendix F), the appropriate t value is found to be 1.961 ( $\nu = 2824$ ;  $\alpha = 0.025$ ). Using Eq. 2-10, the confidence interval for the defect fraction ( $p_p$ ) is determined to be:

$$(0.004248) - (1.961)(0.001224) \leq p_p \leq (0.004248) + (1.961)(0.001224)$$
$$0.00185 \leq p_p \leq 0.00665$$

The validity of the confidence interval must be tested. The t value for a 99.8% confidence interval is 3.094 ( $\nu = 2824$ ;  $\alpha = 0.001$ ). Using Eq. 2-10, the 99.8% confidence interval is found to be:

$$(0.004248) - (3.094)(0.001224) \leq p_p \leq (0.004248) + (3.094)(0.001224)$$

$$0.000461 \leq p_p \leq 0.00804$$

Since the 99.8% confidence interval is between 0 and 1, the calculated 95% interval is shown to be valid. Thus, the 95% confidence interval of the defect fraction (of the lot) is 0.00185-0.00665.

### 2.2.3 Standard deviation control limit (SDCL) approximation

It is often useful to determine what an acceptable control limit for a process would be. The standard deviation control limit (SDCL) approximation can provide two measures of the relative quality of a lot. First, the SDCL approximation can calculate the extreme control limit under which the lot would barely pass at the specified tolerance limit and confidence level, based on the measured sample. Second, the SDCL approximation can estimate a confidence interval for the attribute property cut-off (APC). The attribute property cut-off is the attribute property value that would define an amount of defective material *in the lot* that is equal to the tolerance limit.

Standard deviation methods inherently have associated error (compared to binomial distribution methods), but error is often small when the confidence level is high (see Figure 2-2). For most common acceptance criteria, the specified confidence level is high, so SDCL approximation often is not significantly different from BDCL approximation. The confidence interval for the APC relies on a t table, which is derived based on a normal distribution. Slight errors in confidence interval would be caused by using a t table on a binomial distribution. Note that it is impossible to apply SDCL approximation to qualitative measurements.

#### SDCL approximation method:

There are two separate results that can be obtained from the SDCL approximation: (a) the control limit that the lot would pass at the given tolerance limit and confidence level according to the measured sample and (b) the confidence interval of the attribute property cut-off for the given tolerance limit.

#### (A) Control limit that would be passed

First, it must be determined if the sample size is sufficiently large. Calculate the attribute standard deviation ( $\sigma_a$ ) using Eq. 2-5:

$$\sigma_a = \sqrt{\frac{L_t(1-L_t)}{n}}$$

Check to see if the tolerance limit  $\pm 3$  binomial standard deviations is entirely contained between 0 and 1. If not, then the sample is not sufficiently large to perform the test and the control limit binomial distribution method should be used.

From a z table, look up the z statistic that corresponds to the fixed confidence level. There is no standard format for z tables, so take care to read a z table properly. A full z table is given in Appendix E.

The number of allowable “defective” items in the sample ( $n_d$ ) is determined by:

$$n_d = nL_t - nz\sigma_a \quad (2-11)$$

where  $n$  is the total number of items in the sample and  $L_t$  is the tolerance limit. See Appendix D for a derivation. Round  $n_d$  down to the nearest integer. By looking at an ordered list of the attribute property for each item in the sample, determine the control limit that would declare no more than  $n_d$  items as “defective.” Be sure not to round numbers or choose numbers that would declare additional items as “defective.”

(B) Confidence interval for the attribute property cut-off

Choose a confidence level for the APC confidence interval that is no greater than 99% (a 95% confidence interval is common). Use a  $t$  table to determine the  $t$  value associated with the sample size and chosen confidence level (See Appendix F for the  $t$  table). Often, values on a  $t$  table are found using  $\nu$  (the degrees of freedom) and  $\alpha$ . The  $\nu$  equals the sample size minus one. To get  $\alpha$ , use Eq. 2-9. Calculate the confidence interval for the average number of defects which would occur in hypothetical random samples of  $n$  items ( $N_d$ ) using:

$$nL_t - nt\sigma_a \leq N_d \leq nL_t + nt\sigma_a \quad (2-12)$$

where  $L_t$  is the tolerance limit. Round the lower limit of  $N_d$  down to the nearest integer. Round the upper limit of  $N_d$  up to the nearest integer. By looking at an ordered list of the attribute property for each item in the sample, determine the attribute property value that would declare no more items than the lower limit of  $N_d$  as “defective.” This is one bound of the APC confidence interval. Also, determine the attribute property value that would declare no fewer items than the upper limit of  $N_d$  as “defective.” This is the other bound of the APC confidence interval.

Check to see if the sample is sufficiently large to specify a valid confidence interval for the attribute property cut-off (APC). Using the same procedure as for your chosen confidence level, calculate the 99.8% confidence interval for  $N_d$ . If the 99.8% confidence interval is greater than zero and less than the sample size, then the sample is sufficiently large to specify a valid estimate for the APC confidence interval.

SDCL approximation example:

This example is based on the acceptance criterion from Table 2-1. Assume a sample of 2825 kernels. Assume that Table 2-7 contains an ordered list of the 42 highest sphericities from the 2825 sphericity measurements.

**Table 2-7: A table of actual sphericity data for the SDCL approximation example. An ordered list of the 42 highest sphericities are listed for a sample of 2825 kernels. (same table as BDCL example)**

1.217	1.059	1.049	1.040	1.037	1.033	1.031
1.148	1.057	1.047	1.040	1.037	1.033	1.030
1.092	1.057	1.046	1.039	1.036	1.033	1.030
1.075	1.057	1.044	1.039	1.036	1.032	1.029
1.072	1.056	1.043	1.039	1.035	1.031	1.029
1.060	1.053	1.043	1.038	1.035	1.031	1.028

(A) Control limit that could be passed

First, determine if the sample size is sufficiently large. The attribute standard deviation is:

$$\sigma_a = \sqrt{\frac{(0.01)(0.99)}{2825}} = 0.001872$$

The tolerance limit  $\pm 3$  attribute standard deviations is 0.004384 to 0.01562. Since this is between 0 and 1, the sample is considered large enough to use the SDCL approximation.

From a z table (see Table 2-6), the fixed confidence level of 0.95 (i.e., 95%) can be found to correspond to a z statistic of 1.645. Using the assorted constants, the equation solving for  $n_d$  becomes:

$$n_d = (2825)(0.01) - (2825)(1.645)(0.001872) = 19.55$$

Rounding down to the nearest integer,  $n_d$  equals 19. By looking at Table 2-5, the control limit would need to be set at 1.041 for no more than 19 kernels to be declared “defective.”

According to the measured sample, the lot would pass a control limit of 1.041 at a tolerance limit of 1% (i.e., 0.01) to a 95% confidence level. Although the SDCL approximation is intended for estimating control limits, it also shows that the lot passes the acceptance criterion, because the calculated control limit (1.041) is lower than the specified control limit (1.05).

(B) Confidence interval for the attribute property cut-off

The confidence interval for the attribute property cut-off is determined from the attribute standard deviation (see above) and Eq. 2-12. A 95% confidence interval will be estimated for this example. From a t table (See Appendix F), the appropriate t value is found to be 1.961 ( $\nu = 2824$ ;  $\alpha = 0.025$ ).

$$\begin{aligned} (2825)(0.01) - (2825)(1.961)(0.001872) &\leq N_d \\ &\leq (2825)(0.01) + (2825)(1.961)(0.001872) \end{aligned}$$

$$17.88 \leq N_d \leq 38.62$$

Once the lower and upper limits are appropriately rounded, the lower and upper limits become 17 and 39, respectively. One bound of the APC confidence interval is identified as 1.044 (Table 2-5), so that no more than 17 items are declared as “defective”. The other bound of the APC confidence interval is identified as 1.030 (Table 2-5), so that no fewer than 39 items are declared as “defective”.

Finally, the validity of the estimated 95% confidence interval must be determined. Eq. 2-12 is used with the t value (3.094) associated with 99.8% confidence (for  $v = 2824$  &  $\alpha = 0.001$ ).

$$\begin{aligned} (2825)(0.01) - (2825)(3.094)(0.001872) &\leq N_d \\ &\leq (2825)(0.01) + (2825)(3.094)(0.001872) \\ 11.89 &\leq N_d \leq 44.61 \end{aligned}$$

Since the 99.8% confidence interval falls between 0 and the sample size, the estimated 95% confidence interval for the control limit is valid.

The approximate 95% confidence interval for the attribute property cut-off is 1.030-1.044. In other words: only 1% of the sample (the tolerance limit) is above the attribute property cut-off, and the attribute property cut-off can be stated to be somewhere in the range of 1.030-1.044 to approximately 95% confidence (the confidence level of the APC confidence interval).

### 2.3 Comparison of attribute sampling methods

The examples for all six attribute sampling methods have used the same experimental data. Table 2-8 allows a direct comparison of the various statistical perspectives of the same experimental data. In standard deviation methods, approximation of a binomial distribution with a normal distribution did cause a small error for this sample. In this case, the standard deviation methods have significant advantages over the binomial distribution methods for ease of calculation and for the ability to compare quantitative results of different lots of material. The error associated with the standard deviation test did not affect the comparison to acceptance criterion (see Table 2-1) in this case. To relate these results to Figure 2-2, the corresponding “ $nL_i$ ” values are approximately 19 for BDTL & SDTL approximation and approximately 29 for the other methods.

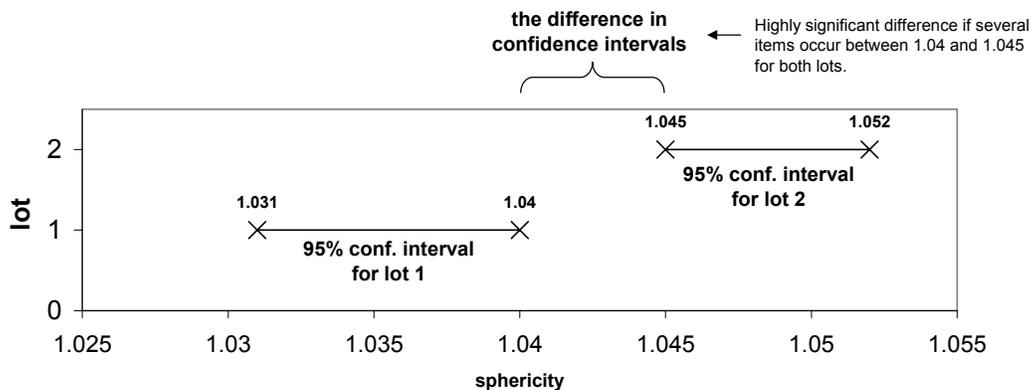
The statistical approximations provide interesting insights into what criteria the sample would pass, but they do not provide specifications that the lot would necessarily pass. This subtle distinction is crucial to avoid erroneous conclusions from the results of these approximations.

For the statistical approximations, special caution must be used in comparing the calculated criteria that would be passed by the samples of different lots. The sample size affects the criteria that a sample would pass, so a variation in sample size would cause differences in the calculated passable criteria.

**Table 2-8: Summary of the six methods for analyzing attribute sampling data. Calculated values are in bold type. The percent error caused by using standard deviation methods (instead of binomial distribution methods) is shown for all calculated values.**

Method	Control Limit	Tolerance Limit	Confidence Level
BD test	1.05	0.01	<b>0.9995</b>   --
SD test	1.05	0.01	<b>0.9989</b>   0.06% error
BDTL approx.	1.05	<b>0.00687</b>   --	0.95
SDTL approx.	1.05	<b>0.00679</b>   1.2% error	0.95
BDCL approx.	<b>1.041</b>   --	0.01	0.95
SDCL approx.	<b>1.041</b>   no resultant final error	0.01	0.95

The most statistically sound method to compare attributes of different lots is the comparison of the confidence intervals of the defect fraction or control limits. SDTL and SDCL approximations provide these confidence intervals using techniques based on widely used statistical methods. The confidence intervals of the defect fraction for multiple lots can be graphed to provide a highly intuitive comparison of lot quality. The confidence intervals of the control limits for multiple lots provide valuable information on lot quality, but differences in confidence intervals of control limits are somewhat deceptive. The true significance of differences in the confidence intervals of control limits is based on the number of defective items associated with the difference in the confidence intervals (see the example in Figure 2-3).



**Figure 2-3: For control limits, the true significance in differences between lots is not graphically intuitive. The significance comes from the number of defective items that occur inside the difference in confidence intervals. (Results for two hypothetical lots are shown.)**

### 3 Variable sampling

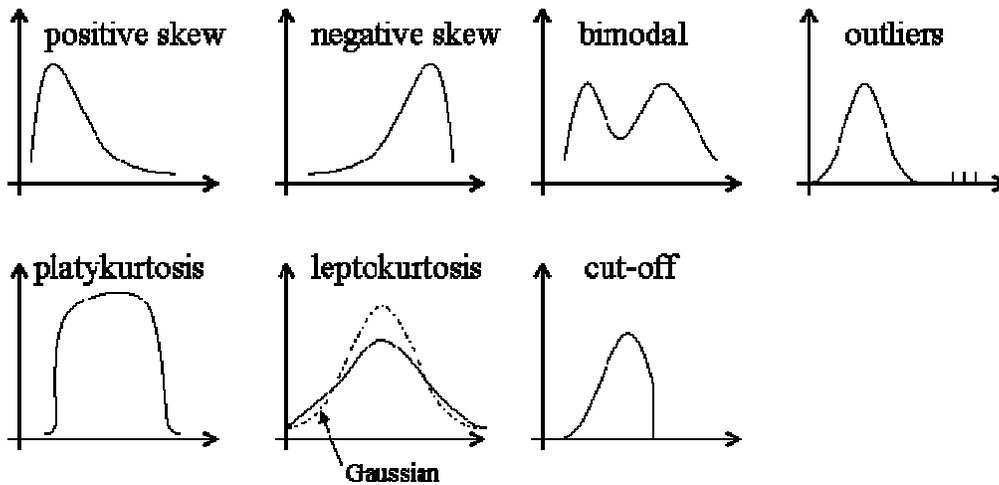
In variable sampling, property measurements of individual items are used to calculate numbers that describe the entire lot. Variable sampling can be based on the mean property measurement of a sample or the distribution of property measurements in a sample. An acceptance criterion of variable sampling for a mean defines acceptance limit(s) that the measured mean must satisfy to a specified confidence level. In variable sampling for a distribution, a tolerance limit defines the maximum percent of an ideal distribution (which is almost always a normal distribution) that can go outside the critical limit(s) based on a specified confidence level.

Generally, variable sampling methods are based on the assumption that property measurements are normally distributed. The assumption of a normal distribution makes it straightforward to calculate variable sampling results that describe the lot to a specified confidence level. Unfortunately, experimental data can be significantly different from a normal distribution. Experimentally generated distributions often have skewness, kurtosis,<sup>3</sup> bimodal distribution, multimodal distribution, processing-induced cut-offs, or outliers.<sup>4</sup> Methods to determine the adequateness of the “normal distribution” assumption are discussed in section 3.1.

One source of confusion and debate is whether a statistical test in variable sampling should be “one-tailed” or “two-tailed.” If the confidence level is based on satisfying only one critical limit, then the test is “one-tailed.” If the confidence limit is based on simultaneously satisfying two critical limits, then the test is “two-tailed.” Therefore, if an acceptance criterion for variable sampling defines only one critical limit, then the statistical test is always “one-tailed.” If the acceptance criteria define two critical limits, then the acceptance criteria should specify whether the two critical limits apply to one “two-tailed” statistical test or the two critical limits apply to two separate “one-tailed” statistical tests. Unfortunately, acceptance criteria with two critical limits usually don’t specify whether the appropriate statistical test is “one-tailed” or “two-tailed.” The distinction between “one-tailed” and “two-tailed” is crucial for properly reading the statistical tables needed for variable sampling methods. Special attention to unambiguous language for writing acceptance criteria and reporting statistical results are included in sections 3.2 and 3.3.

#### *3.1 Adequacy of the normal distribution assumption*

Common features of non-normality are skewness, kurtosis, bimodal (or multimodal) distributions, cut-offs, or outliers (Figure 3-1). Only severe features of non-normality will affect variable sampling based on a mean (VSM) for reasonably large sample sizes, but even moderate features of non-normality can significantly affect variable sampling based on a distribution (VSD). “Overconservative” VSD reports a lower confidence level than actual and is more likely to reject good material. “Underconservative” VSD reports a higher confidence level than actual and is more likely to accept bad material. The potential effects of various non-normal distribution features are discussed below.



**Figure 3-1: Examples of various sources of non-normality.**

Skewness is a measure of the asymmetry in the distribution, which can appear as an imbalance in the tails. Significant skewness can cause significant errors for variable sampling based on a distribution, because a normal distribution overestimates the content in one tail and underestimates the content in the other tail.

Kurtosis is a measure of the content of the center of the distribution relative to the tails. A distribution with platykurtosis has too much content in the center of the distribution relative to the tails. For statistical tests for a distribution, the assumption of normality usually causes overconservative confidence limits for distributions with platykurtosis. A distribution with leptokurtosis has too much content in the tails relative to the center of the distribution. For a distribution with leptokurtosis, statistical tests for a distribution that assume normality claim confidence levels that usually are underconservative.

Bimodal distributions are the sum of two distinguishable distributions (i.e., two peaks discernable in the histogram). Multimodal distributions are the sum of multiple distinguishable distributions (multiple peaks discernable in the histogram). Composite lots (lots made from combining multiple batches) almost always have distributions that are the sum of multiple distributions, but depending on the mean and deviation of each batch, the composite lot can have a nearly normal distribution, a multimodal distribution, or various non-normal distributions. Depending on the relative intensity and location of the modes, bimodal and multimodal distributions can cause significant errors in reported means (depending on sample size) and severe errors in confidence levels reported by statistical tests for a distribution that assumes normality.

Cut-offs are instances where the distribution has a sudden abrupt change from having content to having no content. Sometimes, cut-offs are caused by some inherent limitation in each item, which can be caused by processing, initial selection, or separations. For example, particles may be screened to eliminate overly large and/or overly small items. For statistical tests for a distribution, a cut-off can cause

overconservative or underconservative results. Introduction of cut-offs to a sample through processing, initial selection, or separations is a valuable quality control tool, but the effect of a cut-off on statistical tests should be considered before introducing cut-offs to items of a sample.

Outliers are items with aberrant variable properties that don't fit within the distribution of the remaining sample. It is always possible that an outlier is simply an unlikely occurrence from an ordinary probability distribution, but an outlier may not be caused by the same probability distribution as the majority of the sample. If a sample has a significant percentage of outliers that are beyond the critical limit, statistical tests for a distribution will provide results that are underconservative.

Several analytical techniques have been developed over the years to provide a measure of the normality of a distribution (i.e., its similarity to a normal distribution). Shapiro-Wilk and Lilliefors techniques are commonly used techniques to measure normality. As can be imagined, it is difficult to adequately describe the normality of a distribution with one or two clear, simple indicators, so such techniques are difficult to implement practically. Graphical techniques can be used to compare a histogram of the sample data with a normal distribution. There are two common options for graphical examination of a histogram: (1) directly compare the histogram to a Gaussian (normal) curve or (2) a linearized probability plot. Graphical methods can identify potential problems. Since essentially all sample data are not "normally distributed", it is difficult to determine the potential effect of deviations from normality.

Rigorous analytical and graphical techniques can identify non-normal features in almost all measured distributions, but the presence of non-normal features often does not cause significantly flawed results from the statistical method. The histogram should be examined for non-normal features in order to identify flawed results or questionable results before employing rigorous analytical and graphical techniques.

### 3.2 *Variable sampling based on a mean (VSM)*

For variable sampling based on a mean (VSM), the acceptance criterion has acceptance limit(s) that must be met to specified confidence level. VSM methods are usually based on sampling producing a normal distribution of means, and only VSM methods based on normal distribution will be discussed. The Central Limit Theorem states that means based on sufficiently large samples always will have an approximately normal distribution, even if the distribution of variable properties in the sample is not normal. For large sample sizes, the VSM method works well for distributions of variable properties that aren't very Gaussian (normal).<sup>1</sup>

Small sample sizes with non-Gaussian distributions of variable properties can cause error in the VSM method. For small samples ( $n < 30$ ), non-normality that is clearly apparent in the histogram can cause a significantly non-normal distribution for a sample mean.<sup>1</sup> Unfortunately, any error in the reported mean due to a non-normal distribution is unpredictable in direction.

### VSM Method:

The sample mean is calculated as a simple average of variable properties of items in the sample. The interval for the specified confidence level is called the confidence interval and is calculated with a t-test. Some texts will mention using a “z-test” for this type of test, but a “z-test” is only valid for large samples while a t-test is valid for all sample sizes. To avoid confusion, always use a t-test. Compute the confidence interval with Eq. 3-1 for two-tailed tests, Eq. 3-2 for a one-tailed test with an upper acceptance limit, or Eq. 3-3 for a one-tailed test with a lower acceptance limit.<sup>1</sup>

$$\bar{x} - \frac{t_2 \sigma_x}{\sqrt{n}} \leq \mu_p \leq \bar{x} + \frac{t_2 \sigma_x}{\sqrt{n}} \quad (3-1)$$

$$\mu_p \leq \bar{x} + \frac{t_1 \sigma_x}{\sqrt{n}} \quad (3-2)$$

$$\mu_p \geq \bar{x} - \frac{t_1 \sigma_x}{\sqrt{n}} \quad (3-3)$$

where  $\bar{x}$  is the measured mean of the sample,  $\sigma_x$  is the measured standard deviation of the sample,  $n$  is the number of items in the sample,  $\mu_p$  is the predicted mean of the lot, and  $t_1$  (one-sided) or  $t_2$  (two-sided) come from a t table (see Appendix F). The proper t value is chosen from the t table based on (a) the number of items in the sample, (b) the specified confidence level, and (c) whether the test is one-tailed or two-tailed. Often, values on a t table are found using  $\nu$  (the degrees of freedom) and  $\alpha$ . The  $\nu$  equals the sample size minus 1. To get  $\alpha$ , use:

$$\alpha = \frac{1 - L_c}{T} \quad (3-4)$$

where  $L_c$  is the confidence level and  $T$  is the number of tails of the test (1 or 2).

The lot passes the acceptance criterion if the entire confidence interval falls within the acceptance limit(s) of the acceptance criterion. For one-tailed tests, the confidence interval should be reported with the appropriate inequality sign ( $\geq$  or  $\leq$ ). For two-tailed tests, the confidence interval should be reported as a range (i.e., 453-527  $\mu\text{m}$ ); do not use nomenclature with a “ $\pm$ ” (or “+/-”), because the “ $\pm$ ” terminology can have several different interpretations.

### VSM Example:

This example is based on the acceptance criterion from Table 3-1. Assume a sample of 2924 kernels. Assume that the upper and lower acceptance limits are simultaneously imposed, so that this is a two-tailed test. The mean of the sample is measured to be 503  $\mu\text{m}$ . The standard deviation of the sample is measured to be 208  $\mu\text{m}$ .

**Table 3-1: Example of acceptance criterion for variable sampling based on a mean.**

Property	Acceptance Limits	Confidence Level
Mean Kernel Diameter [ $\mu\text{m}$ ]	480 – 520	0.95 (two-tailed)

To perform the t-test, a t value must be obtained from a t table based on the specified confidence level, the sample size, and the number of “tails” to the test. From Eq. 3-4,  $\alpha$  equals 0.025. From the small t table in Table 3-2, the pertinent t value can be found to be 1.961 (full t table in Appendix F).

**Table 3-2: Data taken from a t table**

v	t values for various $\alpha$	
	$\alpha = 0.05$	$\alpha = 0.025$
1	6.314	12.706
10	1.812	2.228
30	1.697	2.042
60	1.671	2.000
200	1.653	1.972
1000	1.646	1.962
2000	1.646	1.961
3000	1.645	1.961

Eq. 3-1 is solved to obtain the confidence interval for the lot mean.

$$503 - \frac{(1.961)(208)}{\sqrt{2924}} \leq \mu_p \leq 503 + \frac{(1.961)(208)}{\sqrt{2924}}$$

$$495.5 \leq \mu_p \leq 510.5$$

Thus, the 95% confidence interval is 495.5 - 510.5  $\mu\text{m}$ . The confidence interval falls inside of the acceptance limits, so the sample passes the acceptance criterion.

Because the VSM method is often misinterpreted, the following point should be noted. Unless the sample distribution is extremely peculiar, the majority of this sample would be smaller than 480  $\mu\text{m}$  and larger than 520  $\mu\text{m}$  (note the standard deviation of the sample was exceptionally large: 208  $\mu\text{m}$ ). The VSM method provides no information on how broad the distribution of the sample is. The VSM method only specifies a confidence interval for the mean. In this example, the VSM method predicts that the mean of the sample’s distribution is within the range of 495.5 - 510.5  $\mu\text{m}$  to a 95% confidence level.

### 3.3 Variable sampling based on a distribution (VSD)

For variable sampling based on a distribution (VSD), the acceptance criterion specifies a tolerance limit for how many items can be beyond the critical limit(s) to a specified confidence level. VSD methods are usually based on the items of the sample having a normal distribution of the variable property. The assumption of normal distribution often is a limitation of the VSD method. The Central Limit Theorem only applies to means; for distributions, significant deviations from a normal distribution can cause significant errors in the VSD method. See section 3.1 for more detailed discussion on the errors associated with a normality assumption.

#### VSD Method:

The lot passes the acceptance criterion if less than the tolerance limit of a normal distribution based on the sample's mean and standard deviation is outside the critical limit(s) to the specified confidence level.

The appropriate statistical test is a "tolerance factor" test. To determine the range containing all but the tolerance limit of the sample ( $R_t$ ), use Eq. 3-5 for a "one-tailed" test with an upper critical limit, use Eq. 3-6 for a "one-tailed" test with a lower critical limit, or use Eq. 3-7 for a two-tailed test.<sup>3</sup>

$$R_t \leq \bar{x} + K_1 \sigma_x \quad (3-5)$$

$$R_t \geq \bar{x} - K_1 \sigma_x \quad (3-6)$$

$$\bar{x} - K_2 \sigma_x \leq R_t \leq \bar{x} + K_2 \sigma_x \quad (3-7)$$

where  $\bar{x}$  is the mean of the sample,  $\sigma_x$  is the standard deviation of the sample,  $K_1$  is a one-tailed tolerance factor,  $K_2$  is a two-tailed tolerance factor, and  $R_t$  is the range containing all but the tolerance limit of sample (to the specified confidence level). The tolerance factor is obtained from a table using the confidence level (often written as  $1-\alpha$ ), the tolerance limit ( $L_t$ ), the sample size, and the number of tails of the test. Separate tolerance factor tables are available for a one-tailed test and a two-tailed test. A full table of one-tailed tolerance factors ( $K_1$ ) are given in Appendix G. A full table of two-tailed tolerance factors can be found in Appendix H.

If the range defined above ( $R_t$ ) is contained within the critical limit(s), then the sample passes the acceptance criterion.

#### VSD Example:

This example is based on the acceptance criteria from Table 3-3. Assume a sample of 2924 kernels. Assume that the upper and lower critical limits are separately imposed, so that there are two separate one-tailed tests. The mean of the sample is measured to be 503  $\mu\text{m}$ . The standard deviation of the sample is measured to be 42  $\mu\text{m}$ .

**Table 3-3: Example of acceptance criteria for variable sampling based on a distribution**

Property	Critical Limits	Tolerance Limit	Confidence Level
Kernel Diameter [ $\mu\text{m}$ ]	upper limit 600 lower limit 400	0.01	0.95 (one-tailed)

The pertinent tolerance factor must be obtained from a tolerance factor ( $K_t$ ) table. For convenience, Table 3-4 contains a small tolerance factor table for one-tailed tests (full table in Appendix G). On this table,  $L_t$  equals the tolerance limit. Note that a two-tailed test would use a completely different table.<sup>3</sup>

**Table 3-4: Data from a tolerance factor table for one-tailed tests (Appendix G)**

n	Confidence level = 0.95		Confidence level = 0.99	
	$L_t = 0.05$	$L_t = 0.01$	$L_t = 0.05$	$L_t = 0.01$
10	2.911	3.981	3.739	5.075
100	1.92	2.68	2.06	2.85
500	1.76	2.47	1.81	2.54
1000	1.73	2.43	1.76	2.47
2000	1.70	2.40	1.73	2.43
3000	1.69	2.39	1.71	2.41

The two ranges are calculated using Eq. 3-5 and Eq. 3-6.

$$R_t \leq 503 + (2.40)(42)$$

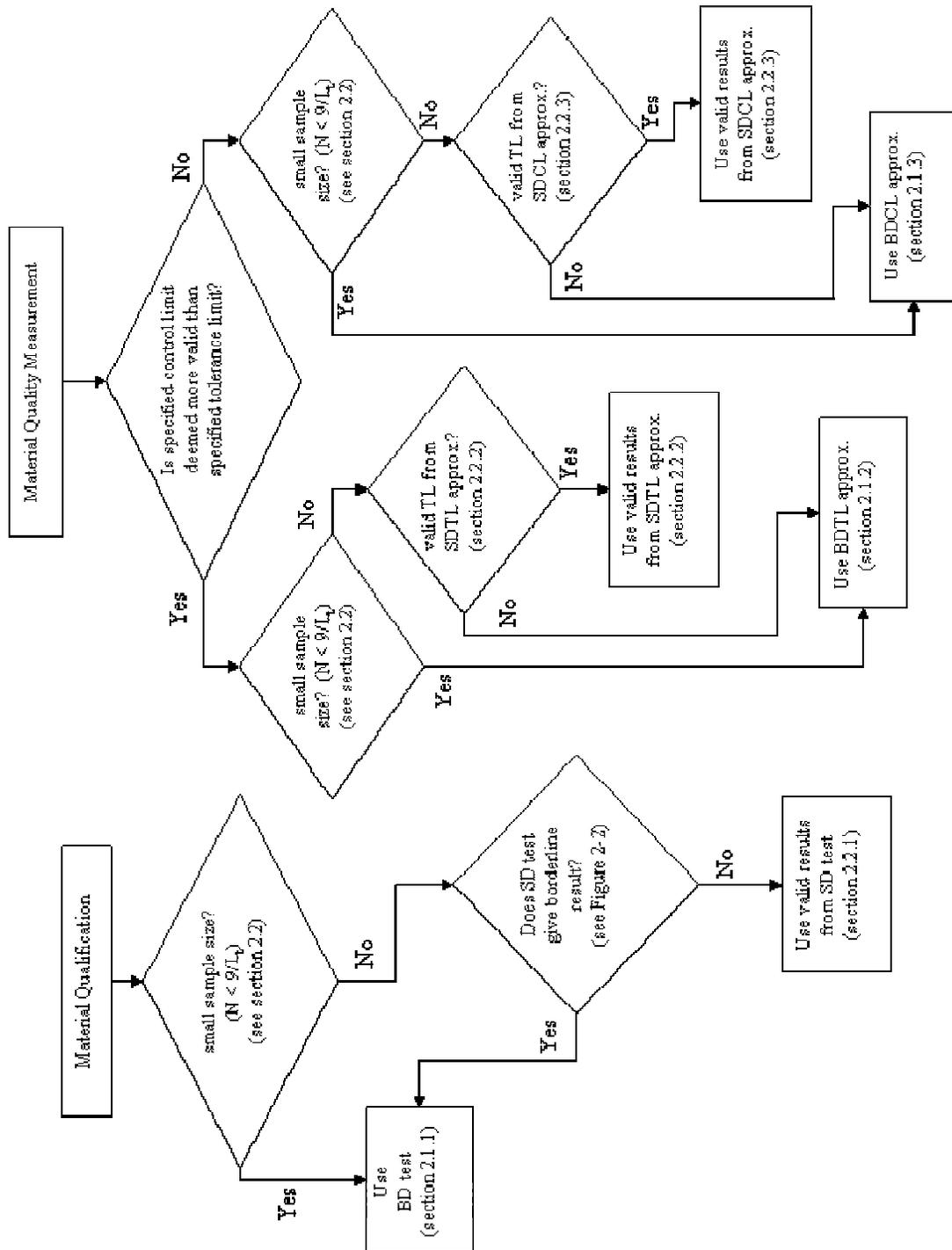
$$R_t \leq 604 \mu\text{m}$$

$$R_t \geq 503 - (2.40)(42)$$

$$R_t \geq 402 \mu\text{m}$$

The lot fails the acceptance criteria, because the one-tailed test for the upper limit does not calculate a range ( $R_t \leq 604 \mu\text{m}$ ) that is entirely below the upper critical limit. Note that the sample did pass the one-tailed test for the lower limit ( $R_t \geq 402 \mu\text{m}$  is above  $400 \mu\text{m}$ ), but both one-tailed tests must be passed to meet the acceptance criteria.

## Appendix A Guide for suggested use of statistical methods for attribute sampling



## Appendix B Derivation of the rule of thumb for “acceptable” use of SD methods

For results to be considered “acceptable”, standard deviation methods require that the tolerance limit is at least three attribute standard deviations away from zero in order to have a distribution that is reasonably Gaussian. The attribute standard deviation is written as:

$$\sigma_a = \sqrt{\frac{L_t(1-L_t)}{n}}$$

The “acceptability” requirement can be expressed as:

$$L_t \geq 3\sqrt{\frac{L_t(1-L_t)}{n}}$$

This expression can be solved for n to obtain:

$$n \geq \frac{9(1-L_t)}{L_t}$$

Since  $(1-L_t) < 1$ , then the conservative simple rule of thumb can be written as:

$$n \geq \frac{9}{L_t}$$

Grant and Leavenworth<sup>2</sup> recommended a slightly different rule of thumb. They considered a normal curve as an unacceptable approximation of a binomial distribution when  $n < 5/L_t$ .

## Appendix C Derivation of the standard deviation tolerance limit (SDTL) approximation

The standard deviation tolerance limit (SDTL) approximation is derived by solving the two fundamental equations of the standard deviation (SD) test for tolerance limit (Eq. 2-5 and Eq. 2.6).

$$\sigma_a = \sqrt{\frac{l_T(1-l_T)}{n}} \quad z = \frac{|l_T - p_m|}{\sigma_a}$$

By substitution, the following equation is obtained:

$$\sqrt{\frac{l_T(1-l_T)}{n}} = \frac{|l_T - p_m|}{z}$$

The equation can be solved for  $p_o$ .

$$\begin{aligned} \frac{l_T(1-l_T)}{n} &= \frac{(l_T - p_m)^2}{z^2} \\ z^2(l_T(1-l_T)) &= n(l_T - p_m)^2 \\ z^2l_T - z^2l_T^2 &= nl_T^2 - 2np_m l_T + np_m^2 \end{aligned}$$

Thus, the final equation is obtained:

$$(z^2 + n)l_T^2 + (-2np_m - z^2)l_T + np_m^2 = 0$$

This quadratic equation (with the form  $ax^2 + bx + c = 0$ ) can be solved with the quadratic formula and will have two solutions.

$$l_T = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

By inserting the appropriate constants into the quadratic formula, the following two solutions are obtained, which correspond to positive and negative  $z$  on the normal distribution.

$$l_T = \frac{2np_m + z^2 \pm z\sqrt{z^2 + 4np_m - 4np_m^2}}{2z^2 + 2n}$$

One solution will apply to the fixed confidence level of the acceptance criterion, while the other solution would apply to a confidence level of 1 minus the fixed confidence level. Since tolerance limits are the maximum allowable fraction of defective items and all practical acceptance criteria have confidence levels of greater than 50%, then the applicable solution is:

$$l_T = \frac{2np_m + z^2 + z\sqrt{z^2 + 4np_m - 4np_m^2}}{2z^2 + 2n} \quad (\text{see Eq. 2-7})$$

## Appendix D Derivation of the standard deviation control limit (SDCL) approximation

The standard deviation control limit (SDCL) approximation is derived by solving a fundamental equation (Eq. 2.6) of the standard deviation (SD) test for the allowable number of defective items ( $n_d$ ).

$$z = \frac{|L_t - p_m|}{\sigma_a}$$

First, solve for  $p_m$ . For confidence levels in practical acceptance criteria,  $p_m$  is less than  $L_t$ , so the absolute value signs can be removed.

$$z\sigma_a = L_t - p_m$$

$$p_m = L_t - z\sigma_a$$

In order to obtain an expression for the allowable number of defective items ( $n_d$ ),  $p_m$  needs to be expressed as  $n_d/n$ , where  $n$  is the number of items in the sample.

$$\frac{n_d}{n} = L_t - z\sigma_a$$

$$n_d = nL_t - nz\sigma_a \quad (\text{see Eq. 2-11})$$

## Appendix E z table (for one-tailed testing)

Confidence levels are given for different z values. For example, the confidence level associated with  $z = 1.28$  can be found in the “1.2” row and the “0.08” column, which is a confidence level of 0.8997.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

This z table was calculated using Microsoft Excel 2003.

**Appendix F t table**

v	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$	$\alpha = 0.001$
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
10	1.372	1.812	2.228	2.764	3.169	4.144
20	1.325	1.725	2.086	2.528	2.845	3.552
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.364	2.626	3.174
120	1.289	1.658	1.980	2.358	2.617	3.160
140	1.288	1.656	1.977	2.353	2.611	3.149
160	1.287	1.654	1.975	2.350	2.607	3.142
180	1.286	1.653	1.973	2.347	2.603	3.136
200	1.286	1.653	1.972	2.345	2.601	3.131
250	1.285	1.651	1.969	2.341	2.596	3.123
300	1.284	1.650	1.968	2.339	2.592	3.118
350	1.284	1.649	1.967	2.337	2.590	3.114
400	1.284	1.649	1.966	2.336	2.588	3.111
450	1.283	1.648	1.965	2.335	2.587	3.108
500	1.283	1.648	1.965	2.334	2.586	3.107
600	1.283	1.647	1.964	2.333	2.584	3.104
700	1.283	1.647	1.963	2.332	2.583	3.102
800	1.283	1.647	1.963	2.331	2.582	3.100
900	1.282	1.647	1.963	2.330	2.581	3.099
1000	1.282	1.646	1.962	2.330	2.581	3.098
2000	1.282	1.646	1.961	2.328	2.578	3.094
3000	1.282	1.645	1.961	2.328	2.577	3.093
5000	1.282	1.645	1.960	2.327	2.577	3.092
10000	1.282	1.645	1.960	2.327	2.576	3.091
30000	1.282	1.645	1.960	2.326	2.576	3.091
infinity	1.282	1.645	1.960	2.326	2.576	3.090

This t table was calculated using Microsoft Excel 2003.

**Appendix G K<sub>1</sub> table (one-tailed tolerance factor)**

n	Confidence level = 0.95		Confidence Level = 0.99	
	L <sub>t</sub> = 0.05	L <sub>t</sub> = 0.01	L <sub>t</sub> = 0.05	L <sub>t</sub> = 0.01
5	4.202	5.741		
10	2.911	3.981	3.739	5.075
20	2.396	3.295	2.807	3.832
30	2.220	3.064	2.516	3.446
40	2.126	2.941	2.365	3.250
50 actual	2.065	2.863	2.269	3.124
50 estimated	2.06	2.85	2.27	3.13
60	2.02	2.80	2.20	3.04
70	1.99	2.76	2.15	2.98
80	1.96	2.73	2.11	2.93
90	1.94	2.70	2.08	2.89
100	1.92	2.68	2.06	2.85
200	1.84	2.57	1.92	2.68
300	1.80	2.52	1.87	2.61
400	1.78	2.49	1.84	2.57
500	1.76	2.47	1.81	2.54
600	1.75	2.46	1.80	2.52
700	1.74	2.45	1.79	2.51
800	1.74	2.44	1.78	2.49
900	1.73	2.44	1.77	2.48
1000	1.73	2.43	1.76	2.47
2000	1.70	2.40	1.73	2.43
3000	1.69	2.39	1.71	2.41
4000	1.69	2.38	1.70	2.40
5000	1.68	2.37	1.70	2.39
10000	1.67	2.36	1.68	2.37

K<sub>1</sub> values below  $n \leq 50$  were taken from McCuen.<sup>3</sup> K<sub>1</sub> values for  $n \geq 50$  were approximated with an equation from Zwillinger & Kokoska.<sup>5</sup> Although not mentioned by Zwillinger & Kokoska,<sup>5</sup> their equation is an approximation that is inappropriate for low n, but is accurate within about 1% or less for  $n > 40$ . The actual and estimated K<sub>1</sub> values are given for  $n = 50$  to demonstrate the accuracy of the values from the approximation equation.

**Appendix H K<sub>2</sub> table (two-tailed tolerance factor)**

n	Confidence level = 0.95		Confidence Level = 0.99	
	L <sub>t</sub> = 0.05	L <sub>t</sub> = 0.01	L <sub>t</sub> = 0.05	L <sub>t</sub> = 0.01
5	5.079	6.634	7.855	10.260
10	3.379	4.433	4.265	5.594
20	2.752	3.615	3.168	4.161
30	2.549	3.350	2.841	3.733
40	2.445	3.213	2.677	3.518
50	2.379	3.126	2.576	3.385
60	2.333	3.066	2.506	3.293
70	2.299	3.021	2.454	3.225
80	2.272	2.986	2.414	3.173
90	2.251	2.958	2.382	3.130
100	2.233	2.934	2.355	3.096
200	2.143	2.816	2.222	2.921
300	2.106	2.767	2.169	2.850
400	2.084	2.739	2.138	2.809
500	2.070	2.721	2.117	2.783
600	2.060	2.707	2.102	2.763
700	2.052	2.697	2.091	2.748
800	2.046	2.688	2.082	2.736
900	2.040	2.682	2.075	2.726
1000	2.036	2.676	2.068	2.718
infinite	1.960	2.576	1.960	2.576

Table values were taken from Natrella.<sup>6</sup>

---

<sup>1</sup> McClave JT, Dietrich FH. *Statistics* 4<sup>th</sup> edition. San Francisco: Dellen Publishing, 1988, pp. 187-191, 194, 292-298, 363-370, 375-379.

<sup>2</sup> Grant EL, Leavenworth RS. *Statistical Quality Control* 6<sup>th</sup> edition. New York: McGraw-Hill, 1988, pp. 209-214, 220.

<sup>3</sup> McCuen RH. *Statistical Methods for Engineers*. Englewood Cliffs NJ: Prentice-Hall, 1985, pp. 22-26, 158-163, 403-406, 407-408.

<sup>4</sup> Freedman D, Pisani R, Purves R, Adhikari A. *Statistics* 2<sup>nd</sup> edition. New York: WW Norton, 1991, pp. 95-96.

<sup>5</sup> Zwillinger D, Kokoska S. *CRC Standard Probability and Statistics Tables and Formulae*. Boca Raton FL: Chapman & Hall/CRC, 2000, pp. 175-177.

<sup>6</sup> Natrella, MG. *Experimental Statistics*. National Bureau of Standards Handbook 91. Washington: U.S. Government Printing Office, 1963, pp. T11-T13.

ADVANCED GAS REACTOR PROGRAM  
OAK RIDGE NATIONAL LABORATORY

ORNL DOCUMENT CLEARANCE / REGISTRATION FORM

PERSON PREPARING FORM: Jan Z. Palmer	PHONE NO.: 576-7054	DATE SUBMITTED: January 6, 2005
DOCUMENT NO.: ORNL/CF-05/01	SPONSOR: DOE-NE / Dr. Madeline Feltus	
TITLE: Statistical Methods for Material Characterization and Qualification		
AUTHOR(s): Andrew K Kercher and John D. Hunn		
SIGNATURES:		
AUTHOR:		
		1-19-05
Name: Andrew K. Kercher and John D. Hunn		Date
TECHNICAL REVIEWER:		
		1-20-05
Name: John J. Saurwein		Date
APPROVER:		
		1-24-05
Name: G. L. Bell		Date