Designation: D7720 − 11 (Reapproved 2017)

# Standard Guide for Statistically Evaluating Measurand Alarm Limits when Using Oil Analysis to Monitor Equipment and Oil for Fitness and Contamination[1]

This standard is issued under the fixed designation D7720; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

## 1. Scope

1.1 This guide provides specific requirements to statistically evaluate measurand alarm thresholds, which are called alarm limits, as they are applied to data collected from in-service oil analysis. These alarm limits are typically used for condition monitoring to produce severity indications relating to states of machinery wear, oil quality, and system contamination. Alarm limits distinguish or separate various levels of alarm. Four levels are common and will be used in this guide, though three levels or five levels can also be used.

1.2 A basic statistical process control technique described herein is recommended to evaluate alarm limits when measurand data sets may be characterized as both parametric and in control. A frequency distribution for this kind of parametric data set fits a well-behaved two-tail normal distribution having a "bell" curve appearance. Statistical control limits are calculated using this technique. These control limits distinguish, at a chosen level of confidence, signal-to-noise ratio for an in-control data set from variation that has significant, assignable causes. The operator can use them to objectively create, evaluate, and adjust alarm limits.

1.3 A statistical cumulative distribution technique described herein is also recommended to create, evaluate, and adjust alarm limits. This particular technique employs a percent cumulative distribution of sorted data set values. The technique is based on an actual data set distribution and therefore is not dependent on a presumed statistical profile. The technique may be used when the data set is either parametric or nonparametric, and it may be used if a frequency distribution appears skewed or has only a single tail. Also, this technique may be used when the data set includes special cause variation in addition to common cause variation, although the technique should be repeated when a special cause changes significantly or is eliminated. Outputs of this technique are specific measurand values corresponding to selected percentage levels in a cumulative distribution plot of the sorted data set. These percent-based measurand values are used to create, evaluate and adjust alarm limits.

1.4 This guide may be applied to sample data from testing of in-service lubricating oil samples collected from machinery (for example, diesel, pumps, gas turbines, industrial turbines, hydraulics) whether from large fleets or individual industrial applications.

1.5 This guide may also be applied to sample data from testing in-service oil samples collected from other equipment applications where monitoring for wear, oil condition, or system contamination are important. For example, it may be applied to data sets from oil filled transformer and circuit breaker applications.

1.6 Alarm limit evaluating techniques, which are not statistically based are not covered by this guide. Also, the techniques of this standard may be inconsistent with the following alarm limit selection techniques: "rate-of-change," absolute alarming, multi-parameter alarming, and empirically derived alarm limits.

1.7 The techniques in this guide deliver outputs that may be compared with other alarm limit selection techniques. The techniques in this guide do not preclude or supersede limits that have been established and validated by an Original Equipment Manufacturer (OEM) or another responsible party.

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

1.9 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

[1] This guide is under the jurisdiction of ASTM Committee D02 on Petroleum Products, Liquid Fuels, and Lubricants and is the direct responsibility of Subcommittee D02.96.04 on Guidelines for In-Services Lubricants Analysis.

Current edition approved May 1, 2017. Published July 2017. Originally approved in 2011. Last previous edition approved in 2011 as D7720 – 11. DOI:10.1520/D7720-11R17.

## 2. Referenced Documents

2.1 *ASTM Standards:*[2]

D445 Test Method for Kinematic Viscosity of Transparent and Opaque Liquids (and Calculation of Dynamic Viscosity)

D664 Test Method for Acid Number of Petroleum Products by Potentiometric Titration

D974 Test Method for Acid and Base Number by Color-Indicator Titration

D2896 Test Method for Base Number of Petroleum Products by Potentiometric Perchloric Acid Titration

D4378 Practice for In-Service Monitoring of Mineral Turbine Oils for Steam, Gas, and Combined Cycle Turbines

D4928 Test Method for Water in Crude Oils by Coulometric Karl Fischer Titration

D5185 Test Method for Multielement Determination of Used and Unused Lubricating Oils and Base Oils by Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES)

D6224 Practice for In-Service Monitoring of Lubricating Oil for Auxiliary Power Plant Equipment

D6299 Practice for Applying Statistical Quality Assurance and Control Charting Techniques to Evaluate Analytical Measurement System Performance

D6304 Test Method for Determination of Water in Petroleum Products, Lubricating Oils, and Additives by Coulometric Karl Fischer Titration

D6439 Guide for Cleaning, Flushing, and Purification of Steam, Gas, and Hydroelectric Turbine Lubrication Systems

D6595 Test Method for Determination of Wear Metals and Contaminants in Used Lubricating Oils or Used Hydraulic Fluids by Rotating Disc Electrode Atomic Emission Spectrometry

D6786 Test Method for Particle Count in Mineral Insulating Oil Using Automatic Optical Particle Counters

D7042 Test Method for Dynamic Viscosity and Density of Liquids by Stabinger Viscometer (and the Calculation of Kinematic Viscosity)

D7279 Test Method for Kinematic Viscosity of Transparent and Opaque Liquids by Automated Houillon Viscometer

D7414 Test Method for Condition Monitoring of Oxidation in In-Service Petroleum and Hydrocarbon Based Lubricants by Trend Analysis Using Fourier Transform Infrared (FT-IR) Spectrometry

D7416 Practice for Analysis of In-Service Lubricants Using a Particular Five-Part (Dielectric Permittivity, Time-Resolved Dielectric Permittivity with Switching Magnetic Fields, Laser Particle Counter, Microscopic Debris Analysis, and Orbital Viscometer) Integrated Tester

D7483 Test Method for Determination of Dynamic Viscosity and Derived Kinematic Viscosity of Liquids by Oscillating Piston Viscometer

D7484 Test Method for Evaluation of Automotive Engine Oils for Valve-Train Wear Performance in Cummins ISB Medium-Duty Diesel Engine

D7596 Test Method for Automatic Particle Counting and Particle Shape Classification of Oils Using a Direct Imaging Integrated Tester

D7647 Test Method for Automatic Particle Counting of Lubricating and Hydraulic Fluids Using Dilution Techniques to Eliminate the Contribution of Water and Interfering Soft Particles by Light Extinction

D7670 Practice for Processing In-service Fluid Samples for Particulate Contamination Analysis Using Membrane Filters

D7684 Guide for Microscopic Characterization of Particles from In-Service Lubricants

D7685 Practice for In-Line, Full Flow, Inductive Sensor for Ferromagnetic and Non-ferromagnetic Wear Debris Determination and Diagnostics for Aero-Derivative and Aircraft Gas Turbine Engine Bearings

D7690 Practice for Microscopic Characterization of Particles from In-Service Lubricants by Analytical Ferrography

E2412 Practice for Condition Monitoring of In-Service Lubricants by Trend Analysis Using Fourier Transform Infrared (FT-IR) Spectrometry

## 3. Terminology

3.1 *Definitions:*

3.1.1 *alarm, n*—means of alerting the operator that a particular condition exists.

3.1.2 *assignable cause, n*—factor that contributes to variation in a process or product output that is feasible to detect and identify; also called *special cause*.

3.1.3 *boundary lubrication, n*—condition in which the friction and wear between two surfaces in relative motion are determined by the properties of the surfaces and the properties of the contacting fluid, other than bulk viscosity.

3.1.3.1 *Discussion*—Metal to metal contact occurs and the chemistry of the system is involved. Physically adsorbed or chemically reacted soft films (usually very thin) support contact loads. Consequently, some wear is inevitable.

3.1.4 *chance cause, n*—source of inherent random variation in a process which is predictable within statistical limits; also called *common cause*.

3.1.5 *characteristic, n*—property of items in a sample or population which, when measured, counted or otherwise observed, helps to distinguish between the items.

3.1.6 *data set, n*—logical collection of data that supports a user function and could include one or more data tables, files, or sources.

3.1.6.1 *Discussion*—Herein a data set is a population of values for a measurand from within a particular measurand set and covering an equipment population.

3.1.7 *distribution, n— as used in statistics*, a set of all the various values that individual observations may have and the frequency of their occurrence in the sample or population.

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

3.1.8 *measurand, n*—particular quantity subject to measurement.

3.1.8.1 *Discussion*—In industrial maintenance a measurand is sometimes called an *analysis parameter*.

3.1.8.2 *Discussion*—Each measurand has a unit of measure and has a designation related to its characteristic measurement.

3.1.9 *nonparametric, n*—term referring to a statistical technique in which the probability distribution of the constituent in the population is unknown or is not restricted to be of a specified form.

3.1.10 *normal distribution, n*—frequency distribution characterized by a bell shaped curve and defined by two parameters: mean and standard deviation.

3.1.11 *outlying observation, n*—observation that appears to deviate markedly in value from other members of the sample set in which it appears, also called *outlier*.

3.1.12 *parametric, n*—term referring to a statistical technique that assumes the nature of the underlying frequency distribution is known.

3.1.13 *population, n*—well defined set (either finite or infinite) of elements.

**Statistical Process Control Technique Terms**

3.1.14 *statistical process control (SPC), n*—set of techniques for improving the quality of process output by reducing variability through the use of one or more control charts and a corrective action strategy used to bring the process back into a state of statistical control.

3.1.15 *state of statistical control, n*—process condition when only common causes are operating on the process.

3.1.16 *center line, n*—line on a control chart depicting the average level of the statistic being monitored.

3.1.17 *control limits, n*—limits on a control chart that are used as criteria for signaling the need for action or judging whether a set of data does or does not indicate a state of statistical control based on a prescribed degree of risk.

3.1.17.1 *Discussion*—For example, typical three-sigma limits carry a risk of 0.135 % of being out of control (on one side of the center line) when the process is actually in control and the statistic has a normal distribution.

3.1.18 *warning limits, n*—limits on a control chart that are two standard errors below and above the center line.

3.1.19 *upper control limit, n*—maximum value of the control chart statistic that indicates statistical control.

3.1.20 *lower control limit, n*—minimum value of the control chart statistic that indicates statistical control.

**Cumulative Distribution Technique Terms**

3.1.21 *cumulative distribution, n*—representation of the total fraction of the population, expressed as either mass-, volume-, area-, or number-based, that is greater than or less than discrete size values.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *alarm limit, n*—alarm condition values that delineate one alarm level from another within a measurand set; also called *alarm threshold*.

3.2.1.1 *Discussion*—When several alarm levels are designated, then a first alarm limit separates the normal level from the alert level, and a second alarm limit separates the alert level from action level. In other words, measurand data values greater than the first alarm limit and less-than-or-equal-to the second alarm limit are in the state of the second level alarm.

3.2.1.2 *Discussion*—An alarm limit, "X", may be single-sided such as "greater than X" or "less than –X"; or it may be double-sided such as "greater than X and less than –X". Alarm limit values may represent the same units and scale as the corresponding measurand data set, or they may be represented as a proportion such as a percent. Alarm limit values may be zero-based, or they may be relative to a non-zero reference or other baseline value.

3.2.1.3 *Discussion*—Statistical process control is used to evaluate alarm limits comparing a control limit value with an alarm limit value. Statistical cumulative distribution is used to evaluate alarm limits by identifying a cumulative percent values corresponding with each alarm limit value and comparing those results, for example, percentages of a data set in each alarm level, with expected percentages of the data set typically associated with each alarm level.

3.2.2 *alarm limit set, n*—collection of all the alarm limits (alarm condition threshold values) that are needed for an alarm-based analysis of measurands within a measurand set.

3.2.3 *critical equipment, n*—category for important production assets that are not redundant or high value or highly sensitivity or otherwise essential, also called *critical assets* or *critical machines*.

3.2.4 *equipment population, n*—well defined set of like equipment operating under similar conditions, selected and grouped for condition monitoring purposes; also called *machine population*, *asset population*, and *fleet*.

3.2.4.1 *Discussion*—Like equipment may refer to equipment of a particular type that may include make, model, lubricant in use, and lubrication system. Similar conditions may include environment, duty-cycle, loading conditions.

3.2.5 *measurand set, n*—meaningful assemblage of measurands collectively representing characteristic measurements that reveal modes and causes of failure within an equipment population.

3.2.5.1 *Discussion*—In industry, a measurand set is sometimes called an *analysis parameter set*.

3.2.6 *noncritical equipment, n*—category for production assets that are not critical equipment; also called *balance of plant*.

3.2.7 *optimum sample interval, n*—optimum (standard) sample interval is derived from failure profile data. It is a fraction of the time between initiation of a critical failure mode and equipment failure. In general, sample intervals should be short enough to provide at least two samples prior to failure. The interval is established for the shortest critical failure mode.

**Alarm Level Terms (in order of severity)**

3.2.8 *WHITE, adj*—favorable level alarm designation showing undamaged or as-new condition having reasonable wear or expected operational condition.

3.2.8.1 *Discussion*—Some other terms used for this level of alarm may include but are not limited to normal, satisfactory, acceptable, level 1, level A, suitable for continued use and good.

3.2.8.2 *Discussion*—WHITE level alarm condition is not usually accentuated by any special color indication on displays or reports.

3.2.9 *GREEN, adj*—favorable alarm level designation showing acceptable condition and showing a measurable change in a measurand value compared with WHITE alarm level.

3.2.9.1 *Discussion*—Some other terms used for this level of alarm may include but are not limited to fair, watch list, monitor, acceptable, level 2, level B and moderate.

3.2.9.2 *Discussion*—GREEN level alarm condition is commonly accentuated by green letters or green highlight or green background in displays or reports.

3.2.10 *YELLOW, adj*—intermediate level alarm designation warning a fault condition is present and will likely need attention in the future.

3.2.10.1 *Discussion*—Some other terms used for this level of alarm may include but are not limited to amber, alert, level 3, level C, low action priority, caution, warning, and abnormal.

3.2.10.2 *Discussion*—YELLOW level alarm condition is commonly accentuated by yellow letters or yellow highlight or yellow background in displays or reports.

3.2.11 *RED, adj*—high level alarm designation showing significant deterioration, review other condition information and consider a possible intervention.

3.2.11.1 *Discussion*—Some other terms used for this level of alarm may include but are not limited to extreme, danger, level 4, level D, unsuitable, actionable, alarm and fault.

3.2.11.2 *Discussion*—RED alarm condition is commonly accentuated by red letters or red highlight or red background in displays or reports.

## 4. Summary of Guide

4.1 This guide is used to statistically evaluate and adjust alarm limits for condition monitoring based on representative measurand data sets from in-service oil sample testing and analysis. This statistical analysis should be performed periodically to update alarm levels using historical data available to the user.

4.2 The user defines an equipment population. The user then selects an appropriate measurand set representing characteristic measurements that reveal likely modes and causes of degradation or failure for the lubricated machinery and for the lubricants for that equipment population.

4.3 For each alarm based measurand the user must have a statistically representative data set covering the equipment population. If the data set follows a parametric statistical distribution, then the user may apply statistical process control (SPC) and cumulative distribution techniques to statistically evaluate alarm limit values. If the data set is nonparametric or if it includes special cause variation, then the user may apply cumulative distribution technique to statistically evaluate and make practical adjustments to existing alarm limit values.

## 5. Significance and Use

5.1 Alarm limits are used extensively for condition monitoring using data from in-service lubricant sample test results. There are many bases for initially choosing values for these alarm limits. There are many questions that should be addressed. These include:

Are those limits right or wrong?

Are there too many false positive or false negative results?

Are they practical?

5.2 This guide teaches statistical techniques for evaluating whether alarm limits are meaningful and if they are reasonable for flagging problems requiring immediate or future action.

5.3 This guide is intended to increase the consistency, usefulness, and dependability of condition based action recommendations by providing machinery maintenance and monitoring personnel with a meaningful and practical way to evaluate alarm limits to aid the interpretation of monitoring machinery and oil condition as well as lubricant system contamination data.

## 6. Assumptions and Limitations

6.1 The assumptions below define the ideal conditions and limitations for alarm limits from a data set representing an equipment population. It is understood that ideal conditions are not often met and that actual conditions may impact the accuracy or sensitivity of the alarm limits. Assumption and conditions include:

6.1.1 Caution should be used for data sets with too few members.

6.1.1.1 For SPC techniques using a normal distribution, caution should be used for data sets with fewer than 30 members. Tentative limits can be set from as little as 10 samples although the quality of the limits will improve with larger populations. Larger populations (for example, in the hundreds) can provide best alarm limits. However, the data needs to be representative of the equipment population.

6.1.1.2 For cumulative distribution techniques regardless of the form of distribution, caution should be used for data sets with fewer than 100 members. Tentative limits can be set from as little as 50 samples although the quality of the limits will improve with larger populations. Larger populations (for example, 1000 plus) can provide best alarm limits. However the data needs to be representative of the equipment population.

6.1.2 The machinery process is a closed loop system whereby test measurements are only affected by operations, maintenance or the onset of a failure mode.

6.1.3 An equipment population or fleet is a population of like machines that would be expected to be maintained according to the same protocol. The machines in the equipment population operated in a similar environment, under a similar duty cycle and load conditions to include use of similar fluids and capacities. Where machinery is maintained as such, it remains part of the same population, regardless of age.

6.1.4 An optimum sample interval has been established accounting for the likely or expected failure modes and at least two samples will be available between failure mode initiation and its terminal phase.

6.1.5 The data set should represent historical measurements covering at least one overhaul interval or in the case of a large fleet, should cover all operational phases from new to overhaul.

6.1.6 Each established measurand is free from interference.

6.2 The following comments only apply to parametric data for which the data set fits a normal distribution:

6.2.1 The population satisfies a normal distribution in accordance with Practice D6299 Anderson-Darling (A-D) statistic which is used to objectively test for normality as described in Subsection A1.4 of Practice D6299, or in accordance with an equivalent test for normality.

6.2.2 Most WHITE and GREEN level alarm data are expected to fall within two standard deviations of the mean or represent about 94 % of all samples taken.

6.2.3 Abnormal sample data are expected to fall outside two standard deviations of the mean and represent about 6 % of all samples taken. These data are expected to exceed a YELLOW level alarm and unacceptable performance or an indication of a degrading condition is expected.

6.3 When using cumulative distribution technique for parametric data, alarm limits may be set at points that do not coincide with standard deviations.

6.4 Careful consideration should be given to the grouping of a population. Improved accuracy to the alarm values and limits being generated can be obtained by dividing a larger group of less similar equipment/machinery into smaller more similar ones.

6.5 Alarm limits that are deemed to be practical must be tested at a minimum using the data set from which they were derived to demonstrate that the functional conclusions are verifiably correct.

6.6 Other statistical methods beyond those stated within this guide may also provide reliable and useful alarm limits. This guide is limited to those discussed in Section 7 as they can be readily applied without extensive statistical training. This guide does not intend to preclude the use of other statistical models.

6.7 Alarm limits may be or may have been developed by OEMs based upon experience, or in house data, or both. These recommendations may be based upon current information or they may have been generated by a company that no longer manufactures the equipment.

6.7.1 For the case of limits based upon current data, these limits can have great value for product support and maintenance. This guide should be considered when variations in usage and maintenance may occur. The user who wishes to depart from OEM suggested alarm limits should consider contact and discussions with the OEM when deviations from their defined limits are made.

6.7.2 For the case of limits based upon old data or from a company that no longer produces or supports the product, changes in lubricants or maintenance practices may have an effect on the OEMs limits provided. These limits may be used as a starting point for limits as discussed in 7.2.2. The techniques stated within this guide would be expected to aid the quality and accuracy of these limits.

## 7. Procedure

7.1 In-service lubricant sample analysis is commonly used for condition monitoring of lubricant characteristics, lubricating system contamination, and equipment wear. Samples are periodically and consistently collected from designated sample points on equipment and are analyzed either by an off-site laboratory, by an on-site laboratory, by on-site test kits or by in-line sensors.

7.1.1 Analyses typically involves multiple tests that produce several measurands (also called analysis parameters) which have been intentionally selected to report and measure characteristics covering the intended range of conditions to be monitored. The group of tests (for example, test profile) is intended to target selected characteristics associated with the asset or equipment type being monitored and produce a list of measurands called a measurand set (also called analysis parameter set). It is common to have three alarm limits between four alarm levels associated with each alarm-based measurand. Alarm limits may be upper or lower or upper and lower depending on the nature of each measurand. The combination of all the alarm limits for a complete measurand set is called an alarm limit set.

7.1.2 It is not necessary for every measurand to have alarm limits. Measurand and data values that are not alarm-based have other uses such as supporting, correlating, or validity checking.

7.1.3 Measurand based alarm limits serve as an intermediate contribution in a process for condition monitoring. Work orders and maintenance actions are based on a review of all data from a measurand set, on historical data and on other information for a measurement point.

7.1.4 This procedure outlines two techniques to statistically evaluate alarm limits applied to data from in-service lubricant analysis condition monitoring: a statistical process control technique and a cumulative distribution technique. Both of these techniques depend on statistical information from multiple data sets where each data set corresponds to a measurand. And the combination of multiple data sets covers all the alarm-based measurands within a measurand set.

7.2 *Equipment Population*—There are many types of equipment in a condition monitoring database. A particular type of equipment is selected for an equipment population that includes a large number of similar equipment items having the same lubricant and operating under similar conditions. A list of all measurands from a lubricant sample test profile selected for an equipment population results in a measurand set representing characteristic measurements selected to reveal likely modes and causes of degradation or failure for the lubricated machinery and for the lubricants.

**TABLE 1 Generic Example**

| How to create a measurand set for an equipment population? | |
|---|---|
| First, choose test of modes and causes. Here are examples: | Then, a measurand set will list measurands specified by your preferred methods, guides, and practices: |
| Particle counting | D6786, D7647, D7416, D7596 |
| Ferrous density | D7416, other |
| Water-in-oil | D4928, D6304, D6439, D7416, E2412 |
| Lubricant chemistry | D664, D974, D2896, D7414, D7416, D7484 |
| Elemental Fe, Pb, Si, Ba, and Na | D5185, D6595 |
| Lubricant viscosity | D445, D7042, D7279, D7416, D7483 |
| Wear debris analysis | D7670, D7416, D7684, D7690, D7685 |

7.2.1 For each measurand for which the user wishes to evaluate alarm limits, the user produces a data set covering the equipment population. If the data set follows a parametric statistical distribution, then the user may apply statistical process control and cumulative distribution techniques to statistically evaluate alarm limit values. If the data set is nonparametric or if it includes special cause variation, then the user may apply cumulative distribution technique to statistically evaluate and make practical adjustments to existing alarm limit values.

7.2.2 Statistical analysis suggested in this guide is most effective using large sets of measured data values (≥30) for each alarm-based measurand. Historical data is necessary for statistical analysis. Thus, this guide is typically used to evaluate and adjust alarm limits. If sufficient historical data is not available, alarm limits for similar or related equipment can be used as a starting point until limits can be generated for the specific equipment. Other get-started alarm limits may be based on sources such as original equipment manufacturers, lubricant suppliers, industry expert consultants. However, these alarm limits should be migrated toward statistically based alarm limits as data becomes available.

7.2.3 The user of this guide will need access to a historical database containing the following information:

7.2.3.1 Equipment information for lubricated machinery and other equipment,

7.2.3.2 Lubricant identity for the lubricant used in each lubricant compartment,

7.2.3.3 An equipment population listing a set of like equipment based on similarity of equipment information and lubricant identity,

7.2.3.4 Failure modes and causes for common problems within the equipment population,

7.2.3.5 A measurand set listing in-service sample test measurands which can identify modes and causes, and

7.2.3.6 Preferably not less than 30 historical measurand data values within the data set used with each measurand for statistical evaluation of alarm limits.

7.2.4 As mentioned earlier, the user studies detail equipment information to designate an equipment population made up of similar equipment, using same lubricant, and operating under similar load, operation, and environmental conditions. Within an equipment population it is possible to have the same type of equipment that is both critical equipment, such as important production assets that are not redundant or high value or highly sensitivity or otherwise essential, as well as noncritical equipment, such as balance of plant equipment and redundant assets. Statistical analysis of data from an equipment population including both critical and noncritical equipment is likely to be applied more conservatively when statistical results are used for evaluating or adjusting alarm limits for critical equipment as compared with noncritical equipment.

7.2.5 This guide works best with a segregated, well-behaved population of identical equipment so that common cause variation in measurand data values will be small and failure modes will be readily observable in the data.

7.2.6 It is recognized that for a laboratory with hundreds or even thousands of equipment variations, it may be time consuming to create, use, and manage dozens or hundreds of different alarm limit sets. Therefore practical application of this guide may require compromise trade-offs when selecting specific equipment for inclusion in a statistical equipment population. The limitation in the quality of alarm limits generated in this fashion should be recognized.

7.2.7 It is common practice for users to designate equipment in categories such as pumps, motors, compressors, gearboxes, steam turbines, gas turbines, diesel engines, etc. Further dividing these down by make and model is desirable, particularly for fleets. Still further dividing groupings down by speed, duty cycle, and load will yield the best alarms.

7.2.8 If a user groups equipment from too many different equipment types or operational functions, then it becomes difficult to assure that resulting alarm limits are relevant and statistically accurate. If there are too few pieces of equipment in the population then variation of measured data within each population becomes too broad causing some problems not to get alarmed, while others are prematurely flagged.

7.3 *Modes and Causes*—For each equipment population, the user creates or selects an appropriate measurand set, which, includes a list of measurands covering many of the commonly experienced failure modes and root causes (“modes and causes”) of component and lubricant. To demonstrate this point, an exemplary discussion is provided in the following paragraphs about how to create a measurand set that covers a few commonly experienced modes and causes of undesirable wear conditions such as abrasive wear, premature fatigue wear, corrosive wear, and abnormal wear resulting from boundary lubrication. The discussion further suggests measurands to identify the condition of the equipment and lubricant used in the associated equipment population. Table 1 is provided as a generic example of how to create a measurand set based on the discussion in the following paragraphs. Measurand sets and the logic behind them vary depending on equipment and lubricant modes and causes of failure. There is more discussion on measurand data set distribution forms in 7.9 – 7.12 for typical measurand data sets referenced in Table 1. Additional examples for selecting tests and therefore measurands to make up a measurand set may be found in Practice D4378 and D6224 where various test options are specified for turbine oils and auxiliary power equipment.

7.3.1 *Abrasive Wear*[3]—Hard particle contamination of a lubrication system is a cause of abnormal abrasive wear in mechanical systems. Excessive abrasive wear is frequently caused by elevated dust contamination. Particle counting and measurement of the element, silicon, are two techniques to reveal a root cause for abrasive wear. Ferrous density measurement and wear debris analysis (atomic emission spectroscopy, ferrography) are two of many techniques to monitor wear such as abrasive wear. There are many other in-service lubricant sample tests one may perform to monitor cause and presence of abrasive wear.

7.3.2 *Fatigue Wear*[3]—Long-term dynamic loading of load bearing surfaces is a principal cause of fatigue wear in tribology. Excessive or premature fatigue wear is frequently caused by elevated dynamic loading. Other condition monitoring means such as vibration analysis are often used to ascertain information about root causes relating to elevated dynamic loading such as imbalance, resonance, or misalignment. Three of many different in-service lubricant sample measurements capable of detecting the onset and progression of fatigue wear include ferrous density, wear debris analysis, and elemental metals analysis.

7.3.3 *Corrosive Wear*[3]—Corrosive fluid contamination of lubrication systems is a principal cause of corrosive wear in tribology. Excessive corrosive wear is frequently caused by elevated water-in-oil or by contamination with another corrosive liquid or gas such as water based coolant or a corrosive process material. Three of many in-service lubricant sample measurements capable of detecting corrosive fluids contaminating a lubricating system are these: water-in-oil, elemental analysis for detecting a coolant additive such as sodium (Na) and for detecting corrosive wear such as iron (Fe) and lead (Pb), and relative permittivity or acid number or base number indicating the lubricant has either gained acidity or lost alkalinity. For this discussion we will include one or more measurands appropriate for measuring water-in-oil, acid number, base number, or elemental Na, Fe, and Pb.

7.3.4 *Wear Due to Boundary Lubrication Failure*[3]—Boundary lubrication wear is the result of transferring at least a portion of bearing load between moving surfaces though metal to metal contact when the fluid film is not fully supporting load. Anti-wear (AW) and extreme pressure (EP) additives are often used to combat effects of friction and wear under boundary lubrication conditions. AW and EP additives may be physically adsorbed or chemically reacted soft films (usually very thin) that help support contact loads. Nonetheless, under boundary lubrication some wear is inevitable. Excessive wear in boundary lubrication regimes often results from inadequate lubrication conditions such as (A) lubricant supply is insufficient, (B) mechanical loading is too high, (C) machine speed is too slow, or (D) lubricant viscosity too low. Measurands from particle counting, microscopic wear debris analysis, elemental analysis, and viscosity are a few examples of tests capable of detecting this failure.

7.3.5 *Lubricant Degradation*[3]—Lubricants are susceptible to chemical degradation from prolonged elevated temperature and exposure to oxygen sources such as moisture or aeration. There are various in-service lubricant sample tests one may use to verify identity and state of degradation for a lubricant. Often these will include tests for viscosity changes (such as a kinematic viscosity testing) and tests for significant chemical changes in lubricant chemistry (such as Fourier Transform Infrared, relative permittivity, acid number, or base number). Measurands from selected tests are included in the measurand set.

7.3.6 *Lubricant Mixing*—Mixing or misapplication is another common problem that should be detected and corrected.

7.4 *Alarm Levels*—An alarm is a means of alerting the operator that a particular condition exists. An example of four distinct levels of alarm states is provided with this guide. They are WHITE, GREEN, YELLOW, and RED. Fewer or more distinct levels may be used. Additional nomenclature for these levels is provided in Section 3.

7.4.1 Alarm limits are the values representing greater-than or less-than thresholds between these levels of alarm.

7.5 *Statistical Process Control (Spc) Techniques for Evaluating Alarm Limits:*

7.5.1 SPC is used to evaluate alarm limits for a measurand data set population that fit a statistically normal distribution. Calculate the standard deviation for the data population. Fig. 1 graphically represents a parametric, normal distribution.

7.5.1.1 *SPC "One-Sigma Limits"*—68.27 % of data set values will fall within one standard deviation of the population mean. Sample data within less than or equal to one standard deviation of the center line (for example, approximately equal to statistical mean) may be comparable to a WHITE alarm level.

7.5.1.2 *SPC "Two-Sigma Limits" or "Warning Limits"*—94.45 % of samples will fall within two standard deviation of the population mean. Sample data greater than one standard deviation away from the center line and less than or equal to two standard deviations away may be comparable with a GREEN alarm level.

7.5.1.3 *SPC "Three-Sigma Limits" or "Control Limits"*—99.73 % of samples will fall within three standard deviation of the population mean. Abnormal to failure conditions are suggested by measurand data greater than a second and less than or equal to a third standard deviation away from the center line may be comparable with a RED alarm level.

7.5.1.4 *SPC "Four-Sigma Variance"*—99.99 % of samples will fall within four standard deviation of the population mean. Assignable or special cause variance is often suggested for data exceeding three standard deviations from the population mean.

7.5.1.5 Outlier values identified using accepted statistical techniques should be removed from the data set populations as outliers. Once removed, the statistics should be reevaluated.

[3] Toms, Larry A., and Allison M. Toms, *Machinery Oil Analysis - Methods, Automation and Benefits*, 3rd edition, STLE, Park Ridge, IL, 2008.

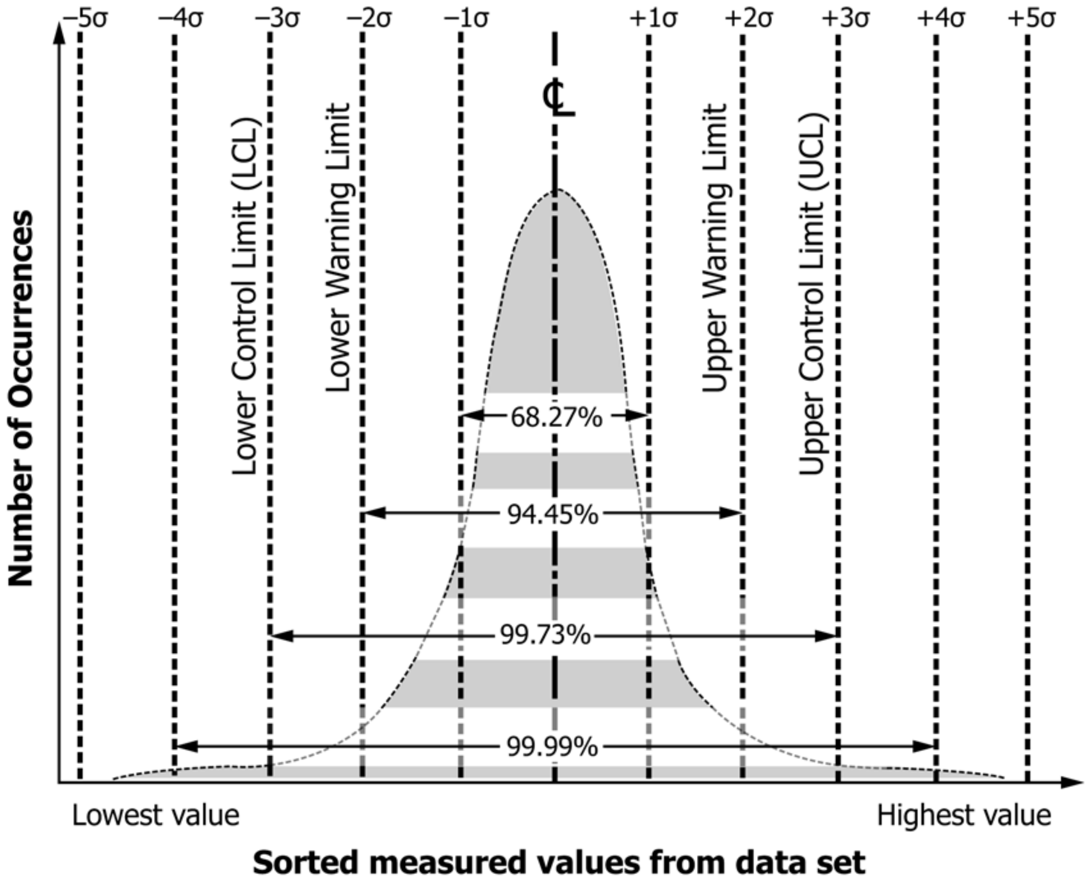


FIG. 1 SPC Parametric, Normal Distribution

7.5.2 The appearance of a low-flattened pattern in the frequency distribution generally indicates that the measurand is not sensitive to the failure mode. The appearance of a multi-modal pattern in the frequency distribution generally indicates that the measurand is not unique to the failure mode, that is, there is more than one failure mode identified by the measurand. In the above situations, the measurand cannot be used alone to identify the fault.

7.5.3 When the sample population has abnormally high incidences of data indicating significant failures, then these results will raise the mean and standard deviation values which will generate higher alarm limits. These higher limits are evident by a frequency distribution that has a distinct curve at the high end. To overcome this effect, sample data with greater than the average plus-or-minus 6 standard deviations should be culled from the population. The recalculated mean and standard deviation on the remaining data set should provide improved alarm limits. If there are no values over 6 standard deviations, the equipment, maintenance practice, etc. should be reviewed to determine the cause of the failures.

7.5.4 Using the average and standard deviation data for each measurand data set, calculate a series of tentative alarm limits based on 7.5.1 – 7.5.3.

7.5.5 The alarm limits listed herein are illustrative examples of how the standard deviation may be used. Some users may want earlier warning for difficult to maintain, critical or high cost components. Keep in mind that policy for setting alarm levels is typically a high level management decision.

7.5.6 *Background Information on SPC:*

7.5.6.1 SPC uses various statistical methodologies to improve the quality of a process by reducing the variability of one or more of its outputs, for example, a quality characteristic of a product or service. A certain amount of variability will exist in all process outputs regardless of how well the process is designed or maintained. A process operating with only this inherent variability is said to be in a state of statistical control, with its output variability subject only to chance, or common causes.

7.5.6.2 Process upsets, said to be due to assignable, or special causes, are manifested by changes in the output level, such as a spike, shift, trend, or by changes in the variability of an output. When special cause variation is eliminated, process variability is reduced to its inherent variability, and control charts then function as a process monitor. Further reduction in variation would require modification of the process itself.

7.5.6.3 The use of three standard errors (for example, standard deviations) for control limits (so-called "three-sigma limits") was chosen by Shewhart,[4] and therefore are also

[4] Shewhart, W. A., *Economic Control of Quality of Manufactured Product*, D. Van Nostrand Company, Inc., 1931.

known as Shewhart Limits. Shewhart chose these limits to balance the two risks of: (1) failing to signal the presence of a special cause when one occurs, and (2) occurrence of an out-of-control signal when the process is actually in a state of statistical control (a false alarm).

7.5.6.4 Special cause variation may also be indicated by certain nonrandom patterns of the plotted subgroup statistic, as detected by using the so-called Western Electric Rules.[5] Herein a subgroup statistic is a subset of a data set such as data related to a particular piece of equipment or to a specific lubricant or a particular root cause affecting a portion of the data set population. To implement these rules, additional limits are shown on the chart at ± two standard errors ("two-sigma limits" or "warning limits") and at ± one standard error ("one-sigma limits").

7.5.6.5 *Western Electric Rules*—The following points are occasionally used to reveal potential indications of a statistical shift away from control:

*(1)* one value falls outside either control limit,

*(2)* two out of three consecutive values fall outside the warning limits on the same side,

*(3)* four out of five consecutive values fall outside the ±one-sigma limits on the same side, and

*(4)* eight consecutive values either fall above or fall below the center line.

7.5.6.6 Other Western Electric rules indicate less common situations of nonrandom behavior:

*(1)* Six consecutive values in a row are steadily increasing or decreasing (trend),

*(2)* Fifteen consecutive values are all within the ±one-sigma limits on either side of the center line,

*(3)* Fourteen consecutive values are alternating up and down, and

*(4)* Eight consecutive values are outside the ±one-sigma limits.

7.5.6.7 These rules should be used judiciously since they will increase the risk of a false alarm, in which the control chart indicates lack of statistical control when only common causes are operating. The effect of using each of the rules, and groups of these rules, on false alarm incidence is discussed by Champ and Woodall.[6]

7.5.7 From the candidate data sets representing machine sample data histories, assemble a population containing the relevant condition data. For best performance select all available data. For large machinery fleets, select data from ten (10) or more consecutive samples from each machine in the equipment population. More consecutive samples would be expected to improve the initial quality of the alarm limits being set and a higher number should be considered when the alarm limits are applied to critical equipment rather than noncritical equipment.

7.5.8 Do not selectively pick through the data or randomly pick samples from the data for a data set. Selectively picked sample data values may introduce a user bias. Randomly picked data from a measurand data set for a large fleet do not always reflect the failure modes under consideration and may generate unreliable alarm limits.

7.5.9 Tentative alarm limits can be set with a data population as small as 10 as stated. In cases where all of the identified failure modes and causes are not represented, lower quality alarm limits will result. These alarm limits can be manually adjusted upward through experience with similar machinery. In cases where failure modes and causes are over represented, higher alarm limits will result. These alarm limits can be manually adjusted downward through experience with similar machinery. In all cases, alarm limits should be continually reviewed and updated either through statistics or experience as more sample data and more failure modes and causes are encountered.

7.5.10 Practical alarm limits may be determined by statistical analysis of monitored machinery and lubricant. Historical and trendable condition monitoring data populations can be described in parametric and nonparametric forms. These data set populations are not time dependant although best results are obtained when consecutive sample data sets are used.

7.5.11 Parametric data can be described as data which satisfies a normal distribution. Three tests can be used to determine if the data is parametric: (1) mean and median values for the data population are approximately the same, (2) a frequency distribution plot of the data population is bell shaped, and (3) the distribution has two tails. When a clear skew or bi-modal shape exists, the data distribution is nonparametric, and it is not a normal distribution.

7.5.12 Degraded conditions that are not yet at a failure condition may also exist. For this case, remediating action is not always required and continued operation may still be reasonable. These conditions however generally still progress to a failure. Alarm levels are used to document these occurrences and produce organizational sensitivity for these conditions.

7.6 *Cumulative Distribution Technique*—A statistical cumulative distribution technique is also recommended to evaluate alarm limits. This particular technique employs a percent cumulative distribution of sorted data set values. The technique is based on an actual data set distribution and therefore is not dependent on a presumed statistical profile. The technique may be used when the data set is either parametric or nonparametric, and it may be used if a frequency distribution appears skewed or has only a single tail. Also, this technique may be used when the data set includes special cause variation in addition to common cause variation, although the technique should be repeated when the identified special cause variation either changes significantly or is eliminated. Outputs of the cumulative distribution technique include specific measurand data values corresponding to selected percentage levels in cumulative distribution plot of the sorted data set. These percent-based measurand data values are used to evaluate and adjust alarm limits. Users should study available information and discern if any data should be removed from the data set population as outlying observations (also called outliers).

[5] Western Electric Company, Inc., *Statistical Quality Control Handbook*, The Mack Printing Company, Easton, PA, 1956.

[6] Champ, C. W., and Woodall, W. H., "Exact Results for Shewhart Control Charts with Supplementary Runs Rules," *Technometrics*, Vol. 29, No. 4, pp. 393–399.

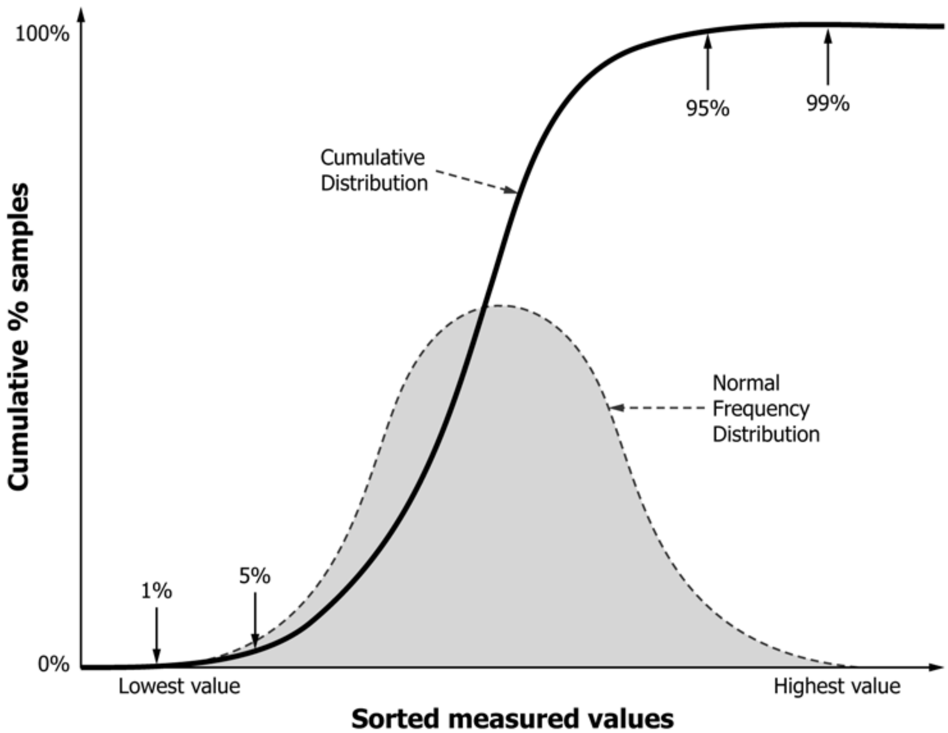


**FIG. 2 Double-Tail Cumulative Distribution with Continuous Data**

7.7 Users of this guide sometimes prefer an option to view sorted data represented on a logarithmic abscissa axis, depending on the nature of measurand data values in response to progressive degradation of a measured characteristic.

7.8 Cumulative distribution users will begin to recognize aspects of a data set distribution such as those summarized in Figs. 2-5.

7.9 An example cumulative distribution for a data set having normal frequency distribution with bell shape and two tails is shown in Fig. 2. Sample viscosity measurements often fit this form. Shearing of a viscosity improver within the oil or fuel dilution contaminating the oil may cause the viscosity to decrease. Oxidation or soot accumulation may cause a viscosity increase. A review of the sample data population being considered will indicate if nonparametric data set is present.

7.9.1 Outer-bound arrows in Fig. 2, mark off segments of the data set population for comparison with measurand alarm levels. These arrows mark off 99 % and 1 % measurand values which have been selected for example, for a user for comparison with YELLOW-to-RED alarm limits for a user who desires to mark off limits confining 98 % of the data set. This approach focuses attention on 2 % of the measurements, it is consistent with capabilities of available limited maintenance resources, it draws attention to data that is frequently associated with assignable or special cause variation, and when this approach is repeated season-after-season together with root-cause elimination it is likely to reduce variance over time.

7.9.2 Inner-bound arrows in Fig. 2 mark off segments of the data set population for comparison with GREEN-to-YELLOW alarm limits. Actual percentages are chosen based on experience. For this example, limits corresponding to measurand values at 95 % and 5 % of the data set have been chosen.

7.9.3 Cumulative distribution techniques applied to actual measurand data values may generate preferred percentages to denote WHITE-to-GREEN alarm limits. The operator selects these limits considering personal experience, published statistical repeatability and reproducibility precision associated with the test method, and other variances associated with good-as-new or WHITE alarm level. These comments regarding use of cumulative distribution techniques to evaluate alarm limits around a WHITE alarm level apply to all four of the following examples discussed below and represented in Figs. 2-5.

7.10 An example cumulative distribution for a data set having a zero-based reference, and a skewed continuous frequency distribution is shown in Fig. 3. Examples of this type of measurand data sets are commonly observed when measuring characteristics that arise when a lubricant is in-service, but are not actually parts of the lubricant such as measurand data sets monitoring contamination and wear characteristics. Initial measured value are often low (often zero) and then increase as the condition progresses. This characteristic measurand and others intended to assist a user when deciding to change oil are well suited for cumulative distribution technique because the data is progressively changing (for example, increasing) from oil change to oil change and between oil top-offs. For these measurands data values do not consistently hover about a statistical mean value. This diagram also represents other measurands such as acid number (AN), oxidation, and relative permittivity (for example, dielectric constant) for equipment

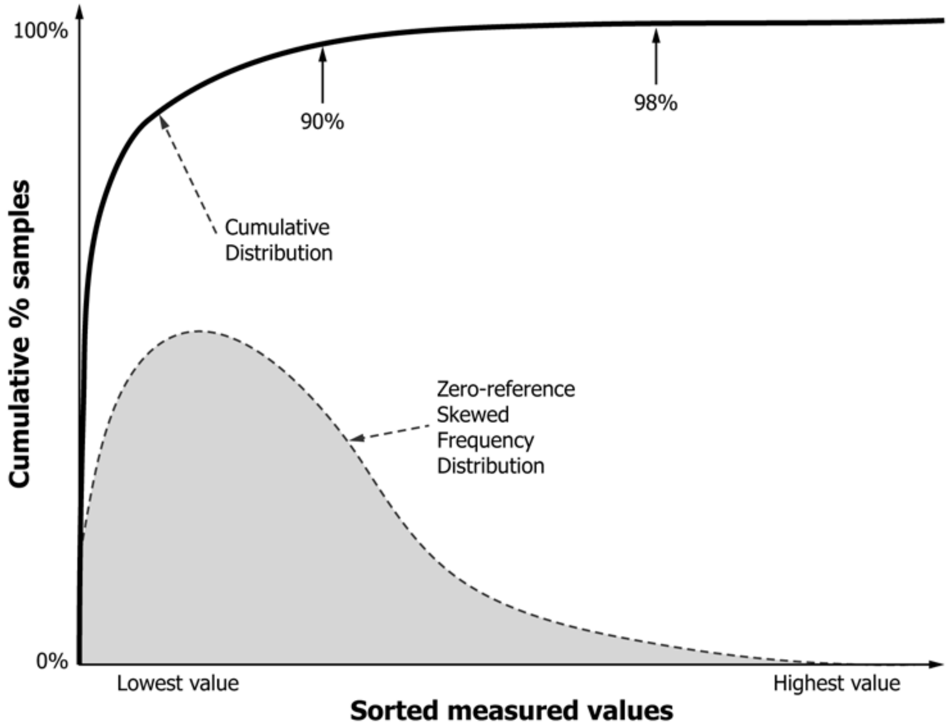


**FIG. 3 Zero-Based Reference, Single-Tail Cumulative Distribution with Continuous Data**

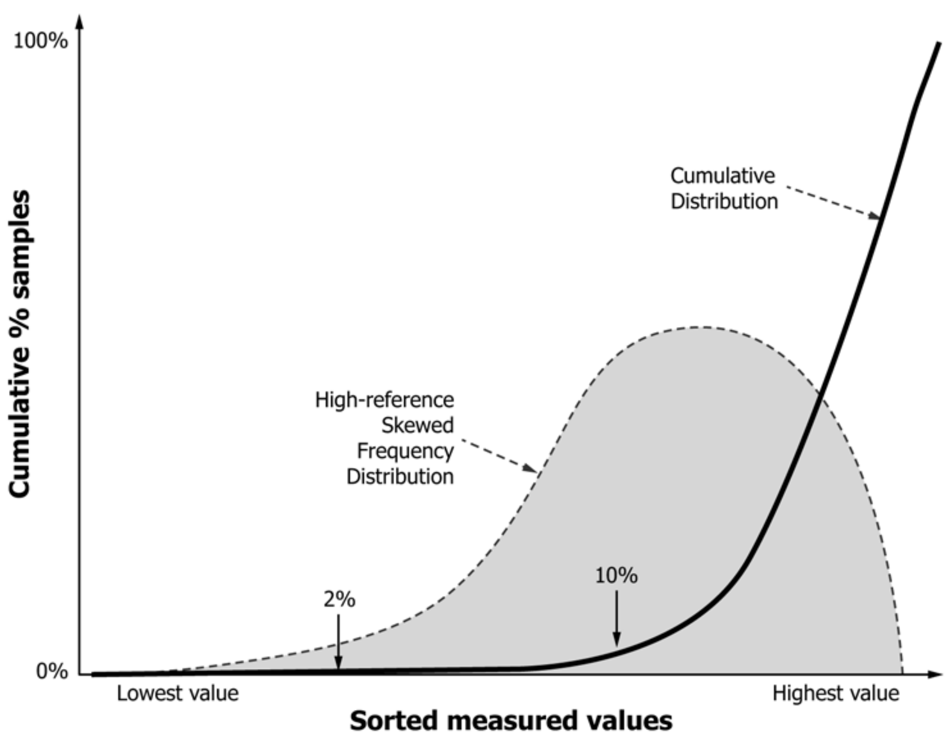


**FIG. 4 High-Reference, Single-Tail Cumulative Distribution with Continuous Data**

applications where data follows a sawtooth trend, sharply dropping to near-zero after each oil change and then increasing over time.

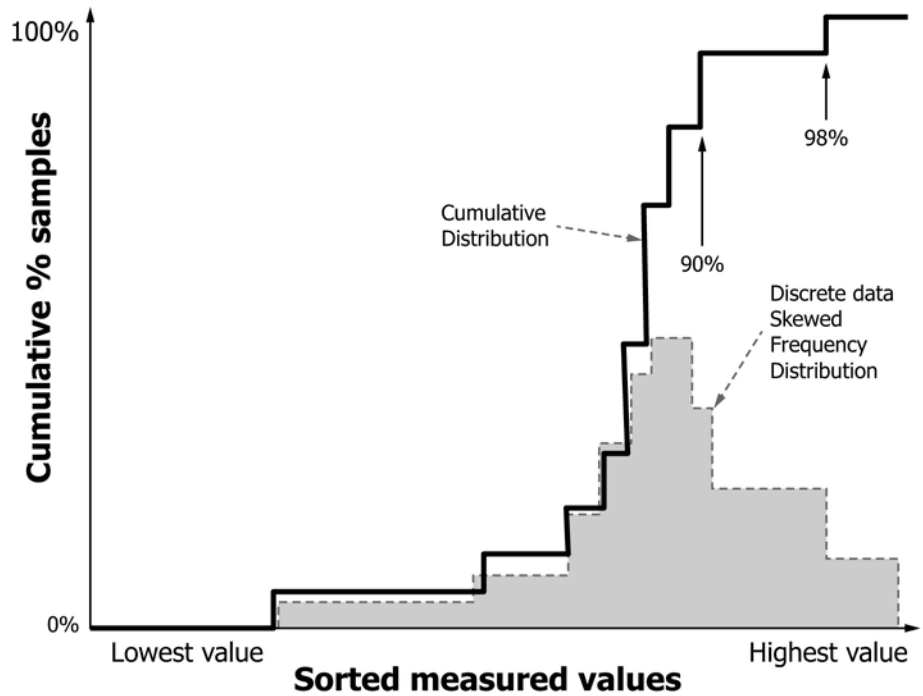


**FIG. 5 Single Tail, Skewed Cumulative Distribution with Discrete Data**

7.10.1 In the Fig. 3 example, an arrow marks off 98 % selected by an operator for comparison with a YELLOW-to-RED alarm limit and another arrow marks off 90 % selected by an operator for comparison with a GREEN-to-YELLOW alarm limit.

7.11 An example cumulative distribution for a data set having a high-reference skewed continuous frequency distribution is shown in Fig. 4. An example of this type of measurand data set is base number (BN). BN values start high and decrease due to oxidation and acid by-products that deplete the additives. This characteristic measurand and others used for determining when to change oil are well suited for cumulative distribution technique because the data are progressively changing (for example, decreasing) from oil change to oil change, and between oil top-offs, so data values do not consistently hover about a statistical mean value. This type of data population is usually nonparametric. A review of the data population being considered will indicate if this type of nonparametric data set is present.

7.11.1 In the Fig. 4 example, an arrow marks off 98 % selected by an operator for comparison with a YELLOW-to-RED alarm limit and another arrow marks off 90 % selected by an operator for comparison with a GREEN-to-YELLOW alarm limit.

7.12 An example cumulative distribution for a data set having a skewed discrete frequency distribution is shown in Fig. 5. Examples of this type include measurand data sets a commonly associated with quantized coded measurement values commonly used for particle counting, as well as severity levels commonly associated with wear particle analysis. Integer based code values are discrete and are often configured to represent exponential, or other nonlinear, data in an easy-to-interpret quantified structure. This type of data population is usually nonparametric. A review of the data population being considered will indicate if this type of nonparametric data set is present.

7.12.1 In the Fig. 5 example, an arrow marks off 98 % selected by an operator for comparison with a YELLOW-to-RED alarm limit and another arrow marks off 90 % selected by an operator for comparison with a GREEN-to-YELLOW alarm limit.

7.13 *Data Validation:*

7.13.1 Perform a reasonableness review based upon experience and knowledge. Repeat the procedure steps as necessary to confirm the alarms if the review does not pass this test.

7.13.2 Validate the tentative limits set by either the SPC standard deviations technique or by the cumulative distribution technique, and compare these statistics with historical data for each equipment problem represented in the test population. Run historical data samples against the newly calculated alarms and compare results against the following:

7.13.2.1 Compare with original recommendations, if any, for the sample and look for correlation, agreement and disagreement.

7.13.2.2 Compare with machinery overhaul and teardown documentation to determine viability and accuracy of the new alarm limits for specific machinery or fluid problems identified.

7.13.2.3 Compare with documentation of instances where “no problems were found” to determine if the new alarm correlates with a false positive or if the new alarm is too sensitive.

7.13.3 Avoid the trap of requiring a 100 % correlation between the new alarm performance and previous alarms or recommendations. Alarm limit selections are based on data interpretation that includes much more than exceeded alarm limits. The purpose of this step is to find gross discrepancies.

The validation evaluation should show a minimum of false positives and each one should be readily explainable.

7.13.4 Compare data validation discrepancies to alarms that can be adjusted using the following additional information.

7.13.4.1 See if there are too many false positive alarms indicating the alarm values are set too low. This is also a problem if the alarm values were calculated from a population where the critical failure modes were not adequately represented. This can be resolved by increasing the alarm limits, ensuring all failure modes are represented.

7.13.4.2 Look for failure of equipment in the absence of an alarm indicating that the alarm limits are set too high, the sampling interval is too long or the measurement used to capture the failure mode is inadequate. This can be resolved by reducing the alarm limit, decreasing the sampling interval or by using a different measurand to represent the failure mode.

7.13.4.3 Look for failure of equipment in the absence of an alarm. This situation is a problem if the alarm limits were calculated from a population where the failure modes were over represented in comparison to the population used to set the alarm values. This can be resolved by reducing the alarm limits or removing some of the most significant failure measurements from the population.

## 8. Report

8.1 A written report shall be recorded and retained for reference, comparison, and continuous improvement processes. The report must include relevant and meaningful user notes about the operation or condition of the equipment, or both. The notes should also include any maintenance preformed on the equipment. The remainder of this section provides suggestions that may be included in or omitted from the report as deemed appropriate by the user.

8.2 *Equipment Population:*

8.2.1 Equipment description. What are defining characteristics for the items are included in this population?

8.2.2 Number of units in population.

8.2.3 Lubricant identity.

8.3 *Measurand Set*—(also called analysis parameter set).

8.3.1 List of tests and method or practice (if available) defining the tests used to produce measurands.

8.3.2 List names identifying each measurand and describe the physical or functional condition monitoring characteristic that is measured by the measurand. Users may find it helpful to organize measurands within a measurand set logically and describe them clearly according to machine wear characteristics, system contamination characteristics, and fluid chemistry characteristics.

8.3.3 Unit of measure for each measurand.

8.3.4 Indicate each measurand that is alarm-based.

8.4 *Alarm Limit Sets:*

8.4.1 Identify and name each alarm limit set for an equipment population. Keep in mind that one measurand set may have several alarm limit sets. That way some equipment, such as critical equipment or frequently sampled equipment may have relatively higher or lower alarms than others in the same equipment population.

8.4.2 For each alarm limit set list the measurands from the measurand set that are alarm-based measurands.

8.4.3 Describe behavior for each alarm-based measurand in the alarm limit set.

8.4.3.1 What is the alarm limit base type? In other words, is it zero-base (meaning it is an absolute scalar value starting from zero when "as new") or reference-base (meaning it is compared with a reference value typically from an "as new" measurement) or point base (meaning it is compared with a statistical mean or median based on prior measurand data from this same sample point); or is the alarm limit value compared to some other baseline?

8.4.3.2 What is the alarm limit delta type? In other words, is the alarm limit expressed in absolute value or in percentage?

8.4.3.3 Are the alarm limits one-sided, one-sided low, or two-sided? In other words, are the alarm limits up (only high sided), down (only low sided), or up and down (both high and low sided)?

8.4.4 What are final alarm limits for each measurand in the alarm limit set? Alarm limits are typically reported as "greater than" values. That means that the alarm limit for YELLOW alarm level is GREEN-to-YELLOW threshold value.

8.5 *Notes*—Regarding statistical analysis used to evaluate and adjust alarm limits.

8.5.1 Number of samples included in each data set.

8.5.2 If available and convenient provide graphical plots for data set populations for each measurand.

8.5.3 When SPC technique is used to evaluate and adjust alarm limits, report meaningful SPC results such as the following:

8.5.3.1 Three-sigma limits, also called control limits or upper and lower control limits.

8.5.3.2 Two-sigma limits, also called warning limits.

8.5.3.3 One-sigma limits.

8.5.3.4 Other SPC findings or possible indications.

8.5.3.5 User notes.

8.5.4 When cumulative distribution technique is used to evaluate and adjust alarm limits report the following:

8.5.4.1 Measurand data value and corresponding cumulative distribution percentage value for each alarm limit on each measurand.

8.5.4.2 Other cumulative distribution findings or possible indications.

8.5.4.3 User notes.

## 9. Keywords

9.1 alarm; alarm level; alarm limit; condition monitoring; control limit; cumulative distribution; equipment population; measurand; measurand set; nonparametric; oil analysis; parametric; SPC; statistical process control