Tutorial on the Statistical Basis of ACE-PT Inc.'s Proficiency Testing Schemes

Note: For the benefit of those who are not familiar with details of ISO 13528:2015 and with the underlying statistical principles upon which it relies, the following very brief and highly simplified tutorial has been prepared as an aid to understanding the process involved.

Part 1: Determining the Assigned Values

Suppose we assign an PT Test Artifact to be measured by a "Perfect Test Operator" who works in a "Perfect Test Lab". This "Perfect Test Operator" then makes emissions measurements (i.e., collects measured data) that are error free, and those "error free" measurements are made using an instrumentation chain (that includes a non-weather-protected OATS equipped with an infinite, perfectly conducted Ground Plane and located in a completely ambient-free environment) that is "error free" as well. After reducing the measured data, the resulting final measurements will be "perfectly accurate" (i.e., will be un-biased, and will have a measurement uncertainty equal to 0.00...0 dB). Thus, this set of "perfect" measurements would be the "true" values of the emissions from that PT Test Artifact.

If these "Perfect Results" could then be compared to actual PT results obtained by real (i.e., imperfect) Test Operators using real (i.e., imperfect) test instrumentation on a real OATS or in a real RF Anechoic Chamber, it would be a relatively simple matter to determine the "quality" of each set of actual PT results.

Of course, there are no "Perfect Test Operators" and there are no "Perfect Test Labs". So, the first thing that needs to be done in a Proficiency Test is to use the measurements collected by each participating Test Lab for each of the types of measurements made (e.g., Antenna Port Conducted Emissions, Radiated Average Emissions at 3 m in Horizontal Polarization, etc.) at each RF Frequency (e.g., 1 GHz, 2 GHz, etc.) to determine a set of estimates of the "true" values, and, to determine the amount of error inherent (i.e., the uncertainty) in each such estimate.

Without going into a lot of detail, the estimates of the "true" values are best determined using some kind of averaging technique to determine the central tendency of the data. The use of a simple arithmetic mean (of the data) readily suggests itself. Unfortunately, the results of simple arithmetic mean calculations are severely affected by "outliers". (Note: an "outlier" is a measurement that has a value that is significantly larger or smaller than most of the other measurements in the same PT Data Set). The results of simple arithmetic mean calculations are even more significantly affected by measurements that are bi-modal or tri-modal. (All three of these conditions are common in Antenna Conducted Emissions and Radiated Emissions PT Data Sets).

What is actually needed to determine the central tendency of the PT data is some kind of weighted averaging technique that is not very sensitive to "outliers", and to bi-modal or tri-modal data distributions. Statisticians call such techniques "Robust". Appendix C.3 of the ISO 13528:2015 Standard contains an iterative algorithm (the so-called "Algorithm A") that produces a robust estimate of central tendency called the "ISO 13528 Algorithm A Robust Average". Hereafter we refer to the result (i.e., the statistic) obtained by inputting measured PT Data (collected at a specific RF Frequency from a specific PT Data Set) into the "ISO 13528 Algorithm A" as the Robust Average, and it will hereafter be denoted as X*). As might be expected, the use of the "ISO 13528 Algorithm A" also produces a statistic that is a measure of the dispersion (i.e., "spread") of the PT data. This statistic, denoted as S*, is the Standard Deviation of the Robust Average. *[Note: S* is often called the Robust Standard Deviation]*. Additionally, and very importantly, the "ISO 13528 Algorithm A Robust Average" technique also produces a statistic called u_x, which is the (derived) Uncertainty of the Robust Average (i.e., of X*). This uncertainty value (i.e., u_x) depends on both S* and the number of participants (denoted as "p").

Part 2: Procedure for the computation of the Performance Statistics (i.e., X*, S* & u_x)

As stated above in Part 1 of this tutorial, in all ACE-PT Inc.'s PT Schemes, the Assigned Values (X*s) are to be determined in accordance with Clause 7.7 "Consensus value from participants" of ISO 13528:2015, using the iterative algorithm (the so-called "Algorithm A") detailed in Appendix C.3 of the ISO 13528:2015 Standard.

Using the "ISO 13528 Algorithm A", X*, S* and, u_x, are calculated for each RF Frequency in each PT Data Set. Thus, if a particular PT Data Set in a given "Round" of a given PT Scheme (say, the Direct Peak Antenna Conducted Emissions PT Data Set from an Initial Round using a particular PT Test Artifact) had 10 RF Frequencies at which measurements were made, then 10 sets of X*, S*, and, u_x values will be calculated by the algorithm - one set of X*, S*, and, u_x values for each RF Frequency in that PT Data Set.

In ISO 13528:2015, the term "Assigned Value" is defined as a: "value attributed to a particular property of a proficiency test item". In a more general sense, the term "Assigned Value" can also be understood as a value attributed to a particular quantity and accepted, sometimes by convention, as having an uncertainty appropriate for a given purpose. In accordance with Clause 7.7 "Consensus value from participants" of ISO 13528:2015, for the purposes of **all** of ACE-PT Inc.'s PT Schemes, we have chosen to denote the Robust Average, X*, as the Assigned Value. This means that we use X* as our best estimate of the "true value" (which is unknown and is unknowable), and we use u_x as the derived Uncertainty of the Assigned Value (i.e., of X*). What this means in practice is that the idea of a "true value" (i.e., a "perfect measurement") has been replaced with the idea of an "estimate of the true value" that has a non-trivial Uncertainty associated with it (and, that we cannot ignore this Uncertainty of the estimate of the "true value").

The calculations of X* and S*, using the "ISO 13528 Algorithm A" is an iterative process. Typically between three (3) and six (6) separate iterations of the algorithms are required in order to achieve a properly converged solution (i.e., a solution where the X*, and, S* values are identical between iterations to at least the second decimal place, and preferably to the third decimal place). *[Note: it is ACE-PT Inc.'s standard practice to carry out six iterations of the algorithms in all cases. This conservative practice serves to simplify the setup of the spreadsheets, because the number of iterations is always constant].*

Consequently, a set of custom-designed spreadsheets and plots were developed in order to mechanize the extremely large number of calculations that result from the use of the "ISO 13528 Algorithm A". All of these custom-designed spreadsheets and plots were developed by ACE-PT Inc.'s Senior Statistician in close consultation with the ACE-PT Inc. PT Technical Manager. The spreadsheets and plots were produced using Microsoft® Excel 2007®. The computations contained in the spreadsheets have been validated by using the PT data set "d1", the contents of which are listed in the leftmost column of Table 2 of the previous version of ISO 13528 (i.e., ISO 13528:2005). The results of this validation exercise matched those provided at the bottom of Table 2 of ISO 13528:2005 to within two decimal places for the value of X* and within one decimal place for the value of S*. The differences, which were not numerically significant, were due to the fact that the calculations were performed using Microsoft Excel 2007®, which employs IEEE Standard 754 floating point arithmetic with 15 digits of precision, whereas the calculations used to produce the values at the bottom of Table 2 of the previous version of ISO 13528 (i.e., ISO 13528:2005) were performed using fixed point arithmetic. *[Note: When iterative calculations are performed in floating point arithmetic, the representational and rounding errors propagate through each iteration, whereas calculations performed in fixed point arithmetic (e.g. hand calculations) do not suffer such errors].* The numerous other calculations (viz. the Z-Scores, the Control Limits, etc.) were validated by as series of spot-checked hand calculations.

Part 3: Procedure for the Computation of the Performance Evaluation Parameters (i.e., the Control Limits) in dB Units

It is a practical necessity to analyze each PT Data Set in such a manner that the results are expressed in the original dB units of measure (i.e., in dBμV units or in dBμV/m units, as applicable) and the Control Limits are also expressed in the original dB units of measure (i.e., dBμV units or dBμV/m units, as applicable). This procedure is highly desirable because most EMC and Wireless/RF Test personnel will find it easier to understand their PT results if those results are expressed in the "natural" units (i.e., dBμV or dBμV/m, as applicable) used to collect the data. For all of ACE-PT Inc.'s current PT Schemes, when using dB units (i.e., either dBμV or dBμV/m as applicable), each Upper Control Limit [UCL] and each Lower Control Limit [LCL] will be calculated as follows:

$UCL = X^* + (U_p + u_x)$

$LCL = X^* - (U_p + u_x)$

where,

X^* *is the Robust Average derived using Algorithm A from ISO 13528; and,*

U_p *is a* ***prescribed*** *measurement uncertainty* ***that is treated as having a fixed value*** *(i.e., is a constant) for a specific PT Test Process in a given round of Proficiency Testing; and,*

u_x *is the standard measurement uncertainty for the Assigned Value (i.e., X*) for a specific PT Test Process being examined in a given round of Proficiency Testing.*

Note 1: For PT Schemes based on normative commercial EMC and Wireless/RF Test Method Standards such as ANSI C63.10-2013 and/or ANSI C63.4-2014, the prescribed Up values are obtained from the applicable expanded measurement uncertainties for a k=2 coverage factor (corresponding to approximately 95% confidence) that are to be found in normative standards such as CISPR 16-4-2 or ETSI TR 100 028. Regardless of the source of the prescribed U_p values, they are always treated as fixed values (constants) for a specific PT Test Process being examined in a given round of Proficiency Testing.

Note 2: In the computation procedures described above, the computed UCL and LCL values are fixed thresholds for performance evaluation for a specific PT Test Process being examined in a given round of Proficiency Testing. Each UCL and LCL consists of two component parts – one part from the prescribed test methods (i.e., U_p - which is treated as a fixed value [constant] for each measured frequency in a given data set from a given round), relaxed by adding the second part, u_x [i.e., the uncertainty value of the assigned value (X*)]. Note: u_x is treated as a *constant* within each measured frequency and as a *variable* for the different measured RF Frequencies in a given Data Set from a given round of Proficiency Testing.

It is important to note here that the MU (i.e., U_p) values given in CISPR 16-4-2 are for a coverage factor k=2, corresponding approximately to a 95% Confidence Interval, whilst the MU (i.e., U_p) values given in ETSI TR 100 028 V1.4.1 are stated for a coverage factor of k=1.96. ACE-PT Inc. has chosen to follow the now universal practice in the commercial EMC and Wireless/RF Testing Community - i.e., to employ the MU (i.e., U_p) values for a k=2 coverage factor in all cases. *[Note: The MU* (i.e., U_p) *values in ETSI TR 100 028 V1.4.1 are easily transformed from their k=1.96 to a k=2 coverage factor]*.

In Clause 9.2 "Limiting the uncertainty of the assigned value" of ISO 13528:2015, it is noted that if $u_x > 0.3S^*$, then the uncertainty of the assigned value is NOT negligible. Because of this, and in accordance with Clause 9.2.1 of ISO 13528:2015, we have chosen ***in all cases*** to use the uncertainty of the assigned value in the interpretation of the proficiency testing.

Part 4: Applicable U_p Values for ANSI C63.10-based and ANSI C63.4-based PT Schemes

CISPR 16-4-2:2003-11, CISPR 16-4-2:2011-06 + A1:2014-02, ETSI TR 100 028-1 V1.4.1, and ETSI TR 100 028-2 V1.4.1 are all extremely lengthy documents that address the Measurement Uncertainty MU (U_p) values for a great many kinds of EMC and Wireless/RFR measurements. However, the measurements addressed by ACE-PT Inc.'s PT Schemes have the following MU values (i.e., $U_p = U_{CISPR\ (k=2)}$ or $U_p = U_{ETSI\ TR\ 100\ 028\ (k=2)}$):

- for Antenna Port Conducted Emissions from 150 kHz to 30 MHz: $U_p = U_{ETSI\ TR\ 100\ 028\text{-}1(k=2)} = 4.10$ dB.
- for Antenna Port Conducted Emissions from 30 MHz to 1000 MHz: $U_p = U_{ETSI\ TR\ 100\ 028\text{-}1(k=2)} = 4.10$ dB.
- for Antenna Port Conducted Emissions from 1 GHz to 18: $U_p = U_{ETSI\ TR\ 100\ 028\text{-}1(k=2)} = 4.10$ dB.

- for Radiated Emissions in all ANSI C63.4 PT Schemes at 3 m in Horizontal Polarization from 30 MHz to 1000 MHz in either a Semi-Anechoic Chamber or on an OATS: $U_p = U_{CISPR\ (k=2)}$ = +/- 5.06 dB from 30 MHz to 200 MHz, and, $U_p = U_{CISPR\ (k=2)}$ = 5.24 dB from 200 MHz to 1000 MHz.
- for Radiated Emissions in all ANSI C63.4 PT Schemes at 3 m in Vertical Polarization from 30 MHz to 1000 M Hz in either a Semi-Anechoic Chamber or on an OATS: $U_p = U_{CISPR\ (k=2)}$ = 5.07 dB from 30 MHz to 200 MHz, and $U_p = U_{CISPR\ (k=2)}$) = 5.26 dB from 200 MHz to 1000 MHz.
- for Radiated Emissions in all ANSI C63.4 PT Schemes at 10 m in Horizontal Polarization from 30 MHz to 1000 M Hz in either a Semi-Anechoic Chamber or on an OATS: $U_p = U_{CISPR\ (k=2)}$ = 5.05 dB from 30 MHz to 200 MHz, and $U_p = U_{CISPR\ (k=2)}$ = 5.21 dB from 200 MHz to 1000 MHz.
- for Radiated Emissions in all ANSI C63.4 PT Schemes at 10 m in Vertical Polarization from 30 MHz to 1000 M Hz in either a Semi-Anechoic Chamber or on an OATS: $U_p = U_{CISPR\ (k=2)}$ = 5.03 dB from 30 MHz to 200 MHz, and, $U_p = U_{CISPR\ (k=2)}$ = 5.22 dB from 200 MHz to 1000 MHz.
- for Radiated Emissions in all ANSI C63.4 PT Schemes from 1 GHz to 18 GHz at 3 m in an Fully Anechoic Room [FAR] or on a Free-Space OATS: $U_p = U_{CISPR\ (k=2)}$ = 5.18 dB from 1 GHz to 6 GHz; $U_p = U_{CISPR\ (k=2)}$ = 5.48 dB from 6 GHz to 18 GHz.

Part 5: Performance Evaluation (Pass/Fail) using UCLs & LCLs computed in dB Units

The participant's results are evaluated by comparing each participating Test Laboratory's measurements (results) with the computed UCL and LCL values. Any measurements which are between the Upper Control Limit [UCL] and the Lower Control Limit [LCL] are defined as being "acceptable" measurements and are therefore deemed to have "Passed". Any measurements which are either above the Upper Control Limit [UCL] or are below the Lower Control Limit [LCL] are defined as being ***"Unacceptable" Measurements. "Unacceptable" Measurements are also defined to be "outliers" and are therefore deemed to have "Failed".*** In other words, the Pass/Fail performance criteria are as follows:

$LCL \leq X_i \leq UCL \rightarrow$ PASS

$X_i > UCL \rightarrow$ FAIL on the "hot side" (i.e., the measurement had a higher value than the UCL)

$X_i < LCL \rightarrow$ FAIL on the "cold side" (i.e., the measurement had a lower value than the LCL)

where,

X_i *is the measurement made at the* i^{th} *RF Frequency in a given PT test in a given "round" of a given PT Scheme.*

Part 6: Performance Evaluation of Laboratory Measurement Bias using D-Statistics

The participant's results are also evaluated by determining an estimate of each participating Test Laboratory's measurement bias. For all of ACE-PT Inc.'s PT Schemes, the Bias (denoted as "D") of the measurement made at the i^{th} RF Frequency in a given PT Data Set from a given "round" of a given PT Scheme shall be determined as follows:

$D = X_i - X^*$

where,

D *is the Bias;*

X_i *is the measurement made at the* i^{th} *RF Frequency in a given PT test in a given "round" of a given PT Scheme;*

X^* *is the Robust Average derived using Algorithm A from ISO 13528.*

Bias is most easily visualized by plotting the computed Bias values (for a given PT Data Set) as a function of RF Frequency, with Bias = 0.00 being in the middle of the Y-Axis of each plot. (In other words, the plots will employ a format similar to those used in plotting the Deviations from Zero Normalized Site Attenuation). We **suggest** (but do not require) that those Test Laboratories whose PT Data exhibit consistently "large" biases (i.e., biases that are consistently > 3 dB above or below the applicable values of X*) should attempt to identify the cause of such measurement biases.

Each participant's Bias values are then evaluated against the Test Method/Procedure Uncertainty (U_p) relaxed by adding the second part, u_x [i.e., the uncertainty value of the assigned value (X*)] as follows:

$|D| = |X_i – X^*| \leq (U_p + u_x) \rightarrow$ PASS

$|D| = |X_i – X^*| > (U_p + u_x) \rightarrow$ FAIL

where,

D is the Measurement Bias.

Part 7: Evaluation of results using Z-Scores and Standardized Control Limits (Z_{LCL} & Z_{UCL})

The participant's results are also evaluated by employing an alternative (equivalent) method that standardizes each participating Test Laboratory's Bias (i.e., D = Xi – X*) and the corresponding computed LCL and UCL values. The performance evaluation of the participants is then made by computing their Z-Score values and comparing them with the standardized LCL (Z_{LCL}) and UCL (Z_{UCL}) values, as follows:

Z-Score = (Xi – X*) / S*

where:

Xi *is the measurement made at the i^th RF Frequency in a given PT test in a given "round" of a given PT Scheme.*

X* *is the Robust Average derived using Algorithm A from ISO 13528.*

S* *is the standard deviation of the assigned value (the Robust Average) for this round of Proficiency Testing.*

and,

Z_{UCL} (Z-Score for UCL) = UCL/S*

Z_{LCL} (Z-Score for LCL) = LCL/S*

Note: *Care should be taken to avoid confusing the term "**Z-Score**" with the term "**Z-Value**".* ***A Z-Score has no underlying assumption of normality (either implicit or explicit), whereas a Z-Value refers to the standardized deviations of a normal distribution.*** *We cannot, on an a priori basis, simply assume normality on collected PT measurement data. In other words, the assumption of normality in the real world EMC and Wireless/RF PT data is erroneous since we can obtain a Z-Score < 2 (or even < 1) and it would be a "FAIL" or a Z-Score >>3 and it would be a "PASS". In this instance, equating (or approximating) the Z-Scores with the Z-Values would lead to giving a "PASS" to participants that should have got a "FAIL" and vice versa.*

Z-Scores that fall between the applicable Z_{UCL} and the applicable Z_{LCL} are defined as being ***"acceptable" measurements***, and are deemed to have "Passed". Z-Scores that fall above the applicable Z_{UCL} or below the applicable Z_{LCL} are defined as being ***"unacceptable" measurements. "Unacceptable" measurements*** are defined as "outliers" and therefore deemed to have "Failed"**.**

Thus,

$Z_{LCL} \leq$ Z-Score $(X_i) \leq Z_{UCL} \rightarrow$ PASS

and,

Z-Score $(X_i) > Z_{UCL} \rightarrow$ FAIL on the "hot side" [i.e., the Z-Score (X_i) had a higher value than Z_{UCL}]

Z-Score $(X_i) < Z_{LCL} \rightarrow$ FAIL on the "cold side" [i.e., the Z-Score (X_i) had a lower value than Z_{LCL}]