E_n Number (101)

Speaker/Author
Michael A.V. Cruz
Metrology Consultant
Science Technology & Research, Inc.
1845 Citadel Ct.
Chula Vista, Ca 91913
619.421.5717
mcruz30@cox.net

Abstract

Have you ever asked the question "what is an E_n Number?" If you have you are not alone. I had the opportunity to ask several senior metrologists recently and the only answer I received was "I'll get back to you on that one". E_n Number (denoted as E_n in ISO Guide 43-1) is a widely used method in analyzing inter-laboratory comparisons. Well, I had discussions with metrology statisticians and did additional researched on E_n and finally came to a basic understanding on its use and application. Understanding the meaning and use of En can be a very useful tool to the working metrologist. The intent of this paper is to describe the basic understanding and application of E_n to your inter-laboratory measurement data.

1. Introduction: What peaks my interest in using E_n was that the Navy's NAVAIR METCAL Program was initiating a Proficiency Test Round Robin (PTRR) effort. The Navy Primary Standards Laboratory (NPSL) had used the E_n Method in prior Proficiency Test efforts, but fully comprehending its use was still limited. The NPSL was the reference laboratory that was part taking in this new PTRR effort. So, I was tasked to review and provide a resolve to the understanding and utilization of E_n for this new effort.

2. E_n is described in ISO/IEC Guide 43-1:1997 (E) [1] as follows: E_n numbers (typically used in measurement comparison schemes, where

$$E_n = \frac{x - X}{\sqrt{{U_{lab}}^2 + {U_{ref}}^2}}$$

and U_{lab} is the Expanded Uncertainty of a participant's result and U_{ref} is the Expanded Uncertainty of the reference laboratory's assigned value.

The evaluation of performance of E_n is also described in ISO/IEC 43-1:1997 (E) as: for E_n numbers:

$| E_n | \leq 1$ = satisfactory

$| E_n | > 1$ = unsatisfactory.

3. These E_n definitions in ISO/IEC Guide 43-1 lack derivation and "how to use guidance". Putting E_n to work needs additional understanding of how to apply to inter-laboratory comparison data. Ref [2] has a very good discussion on E_n as used in comparison measurements. The following section intends to describe my interpretation of the intended application of E_n to inter-laboratory comparison data.

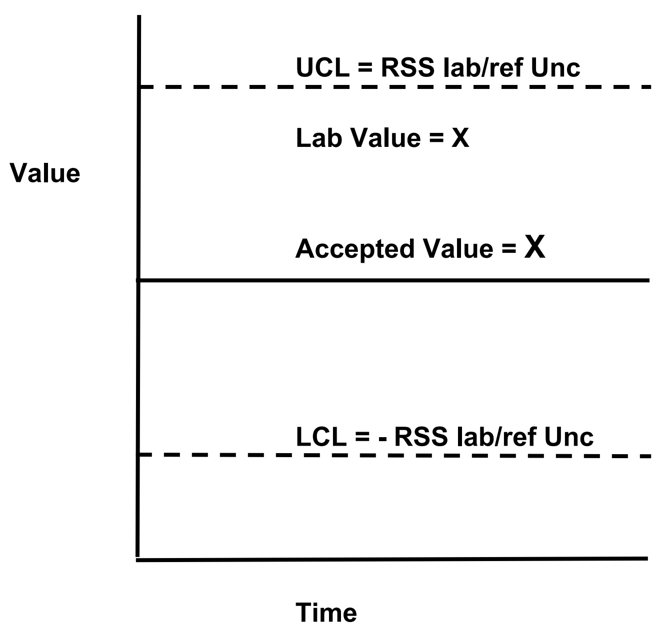
4. E_n has two basic parts:

   Part 1: The numerator, which contains (x-X) also known as "estimate of laboratory bias" as per ISO 5725-4, displays closeness of measurements. This difference value accounts for the closeness of the laboratory's measured value (x) to the reference laboratory's assigned value (X) This comparison data has merit but it is purely subjective! Without additional information you can not reference the difference to any level of performance, such as manufactures specifications or tolerance limits set up by your metrology program. My problem is really trying to use the En Number with some level of performance criteria. So I may say the intercomparison is this "Good". Good meaning some level of acceptance toward our program requirements. This is problematic for me in using this method to describe how good our proficiency test will describe this acceptance level. In E_n the value in providing an estimate of performance levels is in the denominator.

   Part 2: The denominator contains the "root sum square" of the Expanded Uncertainty, where K=2), of the laboratory (U_{lab}) and the reference laboratory (U_{ref}). Each uncertainty contains Type A and B Errors. Type A Error component is derived from the inter-laboratory comparison data. The Type B Error can be derived from uncertainty analysis done on the measurement system as per NIST Tech Note 1297 [3], which may include from manufactures specifications, calibration services provided by higher level laboratory, or tolerance limits/requirements as applied by a metrology program. Each factor sited above can contribute significant meaning to the E_n number. In my recent experience with the Proficiency Test performed by the NAVAIR METCAL Program by far the major factor was the Type B Error from the participating laboratory. The measurement data from the Proficiency Test showed that the Type A Error from both participating and reference laboratories was small as compared to the Type B Error from the participating laboratory, meaning that the observed precision of both measurement systems in the inter-laboratory comparison were quite good. There is an issue here that I will describe later. The participating laboratory Type B Error is significant since it relates the measurements to a level of performance criteria. [Our goal in the proficiency test for NAVAIR was to determine if the participating laboratory could provide measurements within the uncertainty associated with the uncertainty of its measurement system]. **The notion here is that you must have knowledge of both measurement systems involved in the proficiency test! This is essential in using E_n to its fullest**. The intent now is to describe some level of performance criteria in the E_n analysis. As you establish a value for the Type B Error, this can be your point of reference; actually this can be your acceptance or rejection criteria for your inter-laboratory comparison. Our participating laboratory's Type B Error was our point of reference for acceptance/rejection for this proficiency test. This is critical if you are required to compare the measurement data to some metrology program requirement. The real value of E_n comes from the fact that you can now reference your difference data (x-X) to some level of expected performance.

5. $|E_n| \leq 1$, this is such an interesting relationship. What it tells the user is that the quantity (x-X) is equivalent to the RSS of the Total Expanded Uncertainty of the laboratory and reference laboratory. In our case K=2, (95 % confidence level) reiterated in Ref [4]. Now, think of it in terms of the traditional control chart. See (Chart 1). The assigned value (X) from the reference laboratory is the baseline; the measured value (x) from the participating laboratory is displayed on the chart. The Upper/Lower Control Limits (Ucl/Lcl) are equal to the magnitude of the RSS of the uncertainties of the both the reference and participating laboratories. As long as (x) is less than or equal to the Ucl/Lcl values the participating laboratory's measurements are considered acceptable. In other words, when (x) is ≤ the Ucl/Lcl the measurements made by the participating laboratory are acceptable. If we equate the value of (x-X) to the UCL/LCL and take their ratio (as in E_n calculations) we get a value of 1, which is our criterion of acceptance. The acceptance criteria level is at the K=2 level of confidence (95%). How neat is that! So lets return to the E_n calculation, we compute the value for (x-X) and RSS of Expanded Uncertainty (lab/ref) and take their ratio. If the ratio is less than or equal to 1, we understand that the participating laboratory's difference from the assigned value is less than or equal to the RSS Expanded Uncertainty (lab/ref), which happens to be our level of expected performance and we consider the measurements to be acceptable. If the E_n number is greater than 1, (just like in a control chart scenario, the value (x-X) lies outside the control limits and we reject the measurement data) then the value (x-X) is greater than our performance criteria and we reject the data.

Chart 1:
Control Chart Related to E_n



6. Affects of K factors on the value of the RSS of uncertainty values (lab/ref) will enhance or diminish the sensitivity the E_n Number. If you expand the uncertainty to K factors of 2 or 3 you essentially allow the value (x-X) to become larger. So be sensitive to the application of consistent K factors to uncertainty value determination in evaluating E_n numbers.

7. Data from the Navy's proficiency test indicated that the Type A Error was much smaller than Type B (see section 4, part 2). The Type B Error in this scenario would overwhelm the actual measurement data (Type A Error) derived from the laboratory inter-comparison. See Table 1. What we can conclude is that our definition of the Type B Error is correct or that we may not have total understanding of its derivation (knowledge) and may have a certain amount of risk associated with its definition.

Table 1: Small Type A Errors & Large Type B Error

	Type A Error	Type B Error
Reference Laboratory	Small	Small
Participating Laboratory	Small	**Large**

8. We can see from the chart above we can have many different outcomes, see Table 2. Therefore; we must recognize that a simple proficiency test between two laboratories may have many different outcomes and we must recognize the impacts of each error contribution and definition.

Table 2: Type A &B Error w/Laboratories

	Type A Error	Type B Error
Reference Laboratory	Small/large	Small/large
Participating Laboratory	Small/large	Small/large

9. One last comment on E_n Number, the E_n Number is a measure of "Consistency" this means how well one laboratory is compared to another based on knowledge of the measurement systems used in the inter-comparison. As I described the relationship of the E_n Number to a control chart and its Upper/Lower Control Limits we can gleam a relationship of acceptance/rejection criteria. But, we must remember that the uncertainty factor used in the E_n Number is the combination of both laboratories. In the case of the Navy Proficiency Test since the combined uncertainty of the reference laboratory was much smaller than the participating laboratory our acceptance/rejection criteria is mainly based on the combined uncertainty of the participating laboratory. Which in our particular case is what we wanted to compare our findings with.

10. Also, my intent of this paper was to find the origin of E_n Number (AKA Normalized Error). At the time of the writing of this paper, I have not found the originating paper or author. I will continue my search for the origin of this concept and its application. If any reader of this paper can assist in my search, please contact me.

11. Acknowledgements:

I would like to thank: 1) Mr. Ron Nassar, NAVAIR METCAL Program Manager for allowing me to work on this project, which will insure a high level of integrity in his program, 2) Ms. Ding Huang, NAVAIR Statistician, for her invaluable knowledge of statistics and long discussions on E_n and also for allowing me to communicate with Wolfgang Woeger, (PTB) and through other communications with Mr. Raghu Kacker (NIST) and Mr. Norm Belecki (retired NIST) on the subject of E_n, 3) Mr. Steven Dwyer, and 4) Mr. Dennis Jackson, Metrology Science Directorate-NAVSEA Corona, for their insight to the application of E_n to Navy Metrology. And to Mr. James Wheeler, Navy Primary Standards Laboratory, who brought to my attention the concept of E_n.

References:

1. ISO/IEC Guide 43-1:1997 (E), Annex A, section A.2.1.4 (e) and A.3.1.1 (c) (ii)

2. Woeger, Wolfgang, "Remarks on the En-Criterion Used in Measurement Comparison", PTB-Mitteilungen 109, 1/1999, Internationale Zusammenarbeit, p24-27

3. Taylor, B.N. and Kuyatt, C.E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, January 1993. (Available from Superintendent of Documents, U.S. Government Printing Office, Washington DC 20402.)

4. NCSLI, Recommended Practice for Interlaboratory Comparisons, RP-15, August 2005, pp 28-29.