LA-UR-74-794

Conf-740622--1

TITLE: SOME FUNNY THINGS HAPPENED ON THE WAY TO THE LIMIT OF ERROR STANDARD

AUTHOR(S): Roger H. Moore

SUBMITTED TO: Proceedings of the Annual Meeting of the Institute of Nuclear Materials Management

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87544

Form No. 836
St. No. 2629
1/73

UNITED STATES
ATOMIC ENERGY COMMISSION
CONTRACT W-7405-ENG. 36

SOME FUNNY THINGS HAPPENED ON THE WAY TO THE LIMIT OF ERROR STANDARD

(A Memoir)

by

Roger H. Moore
Alternate Group Leader
Statistical Services
Los Alamos Scientific Laboratory
of the
University of California
Los Alamos, New Mexico 87544

Abstract. This paper contains a singular view of the concept of "limit of error" and its several interpretations within the nuclear materials industry. A description and discussion of an American National Standard on limit of error concepts are included.

I. A Disclaimer and Some Background.

A memoir can be defined as "a narrative composed from personal experience." That is the approach I wish to take in this presentation. Although a great many people participated in the creation and the criticism of the forthcoming American National Standard N15.16 on Limit of Error Concepts and Principals of Calculation in Nuclear Materials Control, the comments and views expressed herein are mine alone. I solicit and will be happy to respond to remarks and correspondence generated by this presentation.

I first encountered the creature known as "limit of error" (sometimes "limits of error") in the late 1950s when various members of the Los Alamos Scientific Laboratory's staff and of the Atomic Energy Commission's Albuquerque Operations Office inquired about the quantities required for certain reporting forms, the latest of which I understand to be Form AEC-741 Nuclear Material Transfer Report. That form contains columns headed by "Limit of Error" for both "Element" and "Isotope" for both "Shipper's Data" and "Receiver's Data." The form, designed to simplify data card punching, implies that five characters are available to record each entry.

As we shall see, various governmental and nuclear industrial practices are not consistent with the five-character limitation. Therein lies my long-running puzzlement and concern that in those days a great many of us -- both directly and indirectly involved in the management of nuclear materials -- were not sure what we were doing. Even more clearly, we didn't know what the other fellows (both Ms. and Mr.) were doing either.

Now, however, well headed into the 1970s, some glimmers of improvement are discernable. The jargon of the early days of nuclear material management is being scrutinized, shaken down, and made consistent with concepts and procedures used elsewhere. "Limit of error" is one such jargonal element, and it is the one upon which I will concentrate in this discussion.

Following that flurry of puzzlement a decade ago and some disappointment at not being able to resolve the differences and considerable frustration at being told "you'll never get it straightened out" by some of the pioneers in the game, several years of no contact passed. Then, in September, 1970, I was asked to become a member of "Subcommittee N15-3 (Statistics) of the American National Standards Institute Standards Committee N15 (Methods of Nuclear Materials Control)." In common with many literary characters of the 19th century, little did I know what was in store for me. The first few months of Subcommittee activity, especially in the development of ANSI Standard N15.5 on Statistical Terminology and Notation for Nuclear Materials Management, renewed my interest in helping to bring statistical methods to the needs of nuclear materials managers. Moreover, I gradually became aware that committees such as ours really could bring about needed adjustments. Along with the ability to adjust comes the responsibility to promote those adjustments. And that is the spirit in which these remarks are offered.

II. "Limit of Error" in Statistical Textbook Literature.

Prompted by those early requests for assistance, I examined a number of basic mathematical and statistical texts to try to determine the background of "limit of error." My primary recollection is that I came up empty-handed. There the matter lay for several years.

As I considered this presentation, I decided to retrace my search. I examined a large number (but by no means a random sample) of both old and new books. Neither James and James' Mathematics Dictionary (D. Van Nostrand Company, Inc., 1959) nor Kendall and Buckland's A Dictionary of Statistical Terms (Hafner Publishing Company, Inc., 1971) mention "limit of error." Only three of the many books selected use the phrase in their indexes, and none of those really develop and continue to use the concept.

From Cochran and Cox's Experimental Designs (John Wiley and Sons, Inc., 1957, p. 27):

"2.22. *Number of Replications for Prescribed Limits of Error.* The confidence limits for δ, the true difference between the effects of two treatments, are

$$\delta = d \pm \sqrt{\frac{2}{r}}\, st$$

where t is the value in the t-table corresponding to the confidence probability chosen and to the number of degrees of freedom n in the estimate of error. The quantity $\sqrt{2}\, st/\sqrt{r}$ may be called the limit of error, since it measures the maximum distance between d and δ for the confidence probability chosen. Since the limit of error depends on s, it is not known until the experiment is completed. If s is replaced by its average or expected value, we obtain the expected limit of error

$$L = \sqrt{\frac{2}{r}}\, f\sigma t$$

The factor f enters because the average value of s is slightly less than σ; to a close approximation, f equals

$\left(1 - \frac{1}{4n}\right)$."

From Ku's Precision Measurement and Calibration (NBS Special Publication 300 -- Volume 1, 1969, p. 311-25):

"It is recommended that an index of accuracy be expressed as a pair of numbers, one of the credible bounds for bias, and the other an index of precision, usually in the form of a multiple of the standard deviation (or estimated standard deviation). The terms 'uncertainty' and 'limits of error' are sometimes used to express the sum of these two components, and their meanings are ambiguous unless the components are spelled out in detail."

The third book indexing "limits of error" is Cochran's Sampling Techniques (John Wiley and Sons, Inc., 1963) in which the indexed phrase is "Limits of error (tolerable), 72, 74, 75." Frankly, upon reading those pages, I was unable to choose a quotation or distill a paraphrase for inclusion here. At best, I must leave that as an exercise for the reader.

However, a clear distinction can be made between Cochran and Cox's statement that a "limit of error" is the half-width of a confidence interval and Ku's implication that "limits of error" requires the summing of two quantities (although he does not say how to go about it; i.e., an algebraic sum or a sum of squares?). Furthermore, Ku seems to be promoting the construction of something which, at best, must be considered a mixed bag; random behavior can be characterized by a standard deviation but bias most assuredly is not random behavior, as evidenced by Ku's (op. cit.) remark: "Bias, defined as the difference between the limiting mean and the true value, is a constant, and does not behave in the same way as the index of precision, the standard deviation."

I submit that the path to understanding "limit of error" does not lie among the tomes comprising statistical textbook literature.

III. "Limit of Error" in Nuclear Materials Management Literature.

Firmly ensconced in and committed to the work of the INMM-sponsored ANSI Subcommittee on Statistics, I was bewildered by the vagaries of statistical expression which I encountered at (West) Palm Beach (Shores) in 1971 and again at Boston in 1972. Several inquiries resulted in my being directed to "Ralph's book." And it was here that I began my quest, a search that has uncovered a number of other souces of information.

From Lumb's Management of Nuclear Materials (D. Van Nostrand Company, Inc., 1960, p. 505):

"*limit of error* A value derived from the standard deviation and expressed as a confidence interval; by convention, a 95 percent confidence interval."

Eight other page references to "limit of error" appear in that book's index. And each of the eight appears slightly out of focus when compared with the others. For example, on page 403, we find: "The ... total variance can be used to calculate the limit of error (by convention the limit of error is equal to twice the square root of the variance)." On page 35, we have: "Perhaps the most important factor in determining the number of samples necessary is the amount of error that can be tolerated in the sample estimates, simply stated as a limit of error or margin of error."

From Smith's "Limit of Error Calculations Used in Integrated Safeguards Experiment" (NBS Report 10 491. Also appearing in the Proceedings of the 12th Annual Meeting of the INMM. p. 3 (a footnote)):

"LE defined as boundaries within which the true value of the parameter being measured lies with a probability of 95 percent. Measurement error and sampling error are assumed to be normally distributed."

From Safeguards Glossary (WASH-1162, 1970, p. 8):

"LIMITS OF ERROR The boundaries within which the value of an attribute being determined lies with a specified probability, often, by convention, 95%. [TSO]."

From ANSI N15.5-1972, "Statistical Terminology and Notation for Nuclear Materials Control" (American National Standards Institute, Inc., 1972, p. 7):

"2.3.5.2 *Limits of Error.* A term used by the U.S. Atomic Energy Commission in accountability applications to denote the end points of a 95% confidence interval."

[A personal interjection: When N15.5 was prepared, this definition of "limits of error" seemed to be most consistent with general usage understood by members of the Statistical Subcommittee. I believe subsequent research, bolstered by the requests of those active in nuclear materials management to pin down a definition to a single quantity, has shown that we made an appropriate choice at the time. A modification is in the works.]

From Federal Register (Vol. 38, No. 21 -- Thursday, February 1, 1973, Section 70.51, p. 3078): "(5) 'Limit error' means the uncertainty component used in constructing a 95 percent confidence interval associated with a quantity after any recognized bias has been eliminated or accounted for."

[A personal interjection: I objected to this phrasing, particularly in the light of the forthcoming ANSI standard to deal with the definition. On November 6, 1973, the Federal Register (Vol. 38, No. 213, p. 30544) showed a modification: "Limit error" was changed to "Limit of error." No other change has been discerned.]

From Regulatory Guide 5.18, "Limit of Error Concepts and Principles of Calculation in Nuclear Materials Control" (U.S. Atomic Energy Commission, Directorate of Regulatory Standards, January, 1974):

"The concepts, principles, and referenced methods for calculating limits of error contained in the final draft of ANSI N15.16 ... are generally acceptable to the Regulatory staff for use in nuclear material control and accounting procedures, subject to the following: ... The calculated limits of error ... should be based on 95% confidence intervals for the estimator, which must consider the effective degrees of freedom associated with the estimated variance..."

As a final example, consider this passage from WASH-1282, "Methods for the Accountability of Plutonium Nitrate Solutions" (Lawrence Livermore Laboratory, February, 1974, pp. 3-5):

"The term 'limit of error' is defined[2] to mean the endpoints of a 95% confidence interval on a population mean and is usually calculated from:

$$LE = \pm 1.96\sigma_{\bar{x}} = \pm 1.96\sigma/n^{1/2}$$

where $\sigma_{\bar{x}}$ is the standard error of the mean for a single variable...

"Sometimes there is no prior information about the magnitude of the population standard deviation. The standard deviation is then determined from a limited number ($n<30$) of current measurements. In such instances the sample standard deviation is an estimator of the population standard deviation and the symbol is s.... The limit of error is then given by:

$$LE = t_{0.05,f} s_{\bar{x}}$$

where: $t_{0.05,f}$ is the factor obtained from the t-distribution for f degrees of freedom and total probability of 0.05 (0.025 from each tail of the distribution)...." (Note: The superscripted "2" points to a reference: "Section 70.51, 'Material balance, inventory, and records requirements' of Part 70, 'Special Nuclear Material' of Title 10 of the Code of Federal Regulations.")

By now, I trust the problem is clear. Neither statistical or nuclear materials management literature are of any real assistance in helping us find out what a "limit of error" is.

However, before describing the Statistical Subcommittee's resolution of the question, I am compelled to make some remarks about confidence intervals. I am not concerned about the technical definition. A confidence interval can be constructed with any desired confidence coefficient (as long as the coefficient lies between 0 and 1). It need not be symmetric; indeed, by some criteria of "best," best confidence intervals in some cases cannot be symmetic. But, if we confine our attention to symmetric confidence intervals, they can be expressed in the general form $a \pm b$. The value a is the midpoint of the interval; it is not the quantity being estimated; it is the estimate of that quantity in most applications of confidence interval concepts. The interval consists of all points lying between the quantities (a-b) and (a+b). The endpoints of the interval are given by (a-b) and (a+b). The half-width of the interval is given by the quantity b; the half-width is not designated by $\pm b$. The width of the interval is given by 2b; the width is not given by $\pm 2b$. The magnitude of the quantity b is not necessarily given by doubling the standard deviation of the quantity a, nor is the magnitude of the quantity b necessarily given by multiplying the standard deviation of the quantity a by a value from a t-table "which must consider the effective degrees of freedom."

Enough of sermonizing. The Statistical Subcommittee was faced with the task of resolving the many differences and nuances of use of "limit of error" in nuclear materials management. I believe the foregoing should demonstrate that the task was not an easy one and it was not entered into lightly. I turn now to the outcome of the subcommittee's deliberations.

IV. "Limit of Error" Defined in ANSI N15.16 -- and Why.

Following is the definition appearing in ANSI Standard N15.16 on Limit of Error Concepts and Principles of Calculation in Nuclear Materials Control:

> "The limit of error of an estimator T is twice the standard deviation of T; i.e., twice the square root of the variance of T."

As a technical statistical concept, an estimator is a rule or method of estimating the value of quantity of interest. It is a function of random variables, and it therefore is itself a random variable. Since one of the characterizations of a random variable is in terms of its variance, it follows then an estimator can be characterized in terms of its variance. Once a set of observations is subjected to manipulation by that "rule or method of estimating," the resulting value is called an estimate.

Now, why did the Subcommittee opt for this definition of limit of error?

First, it was apparent that no particular definition had arisen from the general usage of the term which had caught on and had been understood by any sizeable sector of the nuclear industry.

Second, the Subcommittee felt that no real advantage would be gained by giving the endpoints or the half-widths of a confidence interval a special name.

Third, there were implications in some usages that the interval of interest was to be treated as a probability interval, either in terms of stating something about the fraction of elements having values lying within the interval or in terms of specifying something about future observations.

Fourth, there has been some precedence for a multiple of a standard deviation being given a special name; i.e., the probable error is 0.6745 times the standard error, where standard error is the positive square root of the variance of the sampling distribution of a statistic.

Fifth, since limit of error was considered too ingrained to be removed from the literature, the Subcommittee concurred with the proposal that -- whatever we came up with -- the definition in the ANSI standard should not be a mere renaming of some other well-established statistical concept.

[A personal interjection: My opinion at the time of the Subcommittee's debate about the preparation of a limit of error standard was that the whole concept of limit of error should be removed from nuclear industrial theory and practice. Failing in that stand, my next best choice was to work toward a standard that would, at the very least, not destroy and not rename other fundamental statistical concepts.]

Sixth, the many requirements for reporting a limit of error seemed to lead to a single number, thus eliminating the plural connotation of "limits of error" for a single reported value and the particularly troublesome question of what to do when non-symmetric intervals arise.

I believe that the definition contained in ANSI N15.16 meets the many conflicting views and requirements with which the Subcommittee had to grapple. Is the word "statesmanlike" too self-aggrandizing?

V. The "Other Part" of ANSI N15.16: Principles of Calculation.

Any set of "principles," whether they ultimately become Bill of Rights or instructions for creating an entry into a chili cook-off among United States senators from the Southwest, is subject to discussion, debate, and interpretation. In my opinion, one of the strongest virtues of ANSI N15.16 is that it contains a set of principles for the calculation of a limit of error. As we say in the standard: "Although the principles here are not intended to be exhaustive, a potential user of a proposed technique [for calculating a limit of error] should be circumspect if the technique fails any of the following [principles]."

Herewith, then, are six principles of calculation for a limit of error as stated in ANSI N15.16 (along with my parenthetical views):

"(1) The information and procedures used to determine a limit of error shall be relevant and timely. NOTE: 'Information' includes historical measurement system performance as well as data taken to observe the system's current status." (Don't say something is a limit of error unless it's based on facts that are valid when the statement is made, but don't forget what you've learned in the past.)

"(2) Measurement operations which affect a reported value shall be identified and evaluated to determine their contributions to a limit of error." (Various contributors to variability -- measured by variance -- must be sought out and understood.)

"(3) A limit of error shall reflect both random and systematic components. NOTE: Because the limit of error is a function of the variance of a process, it is imperative to recognize that any systematic components are themselves consequences of random behavior and are not biases...." (In my opinion, the nuclear materials industry is the only major industry which suggests that a bias (which I take as interchangeable with a systematic error) is also a random variable. This is an unfortunate state of affairs. It may well be that a bias is conveyed upon us by the gods of chance, but--unless we can get at least two assessments of that celestial probability mechanism--we are acting improperly when we talk about "systematic error variances." An estimate of bias is another matter; we may, indeed, be able to attach a variance to such an estimate. But, for the nonce and until fully explored by a cadre of industrial statisticians, I believe that we should continue to do everything possible to distinguish between systematic errors and random errors and that attempts to combine them into single statements should be scrupulously avoided.)

"(4) The procedures and information used in calculating a limit of error shall be documented." (Elementary advice, perhaps, but without documentation, such information and procedures can seldom be understood and appreciated by newcomers to the endeavor. Indeed, at some later time, the Browning effect* can engulf even the most careful artist or scientist.)

"(5) Any procedure for calculating a limit of error shall provide for the inclusion of whatever covariances might exist among the random variables involved." (Not all observations are statistically independent of each other. At the same time, statistical dependence is not necessarily the consequence of doing something in the same way or using the same equipment. When covariances are not zero, include them; when they are zero, no difficulty will arise from using formulations and procedures that allow them.)

"(6) A limit of error shall reflect only the contributions of random phenomena inherent in the measurement process." (This principle serves at least two purposes: First, pure mistakes and blunders are not desireable and should not pollute the reported characterizations of the process. Second, such non-random phenomena as biases are not to be included in the reported limit of error.)

VI. Parting Shots.

I realize that some of the subjects touched upon in this presentation are not consistent with the ideals and procedures of many practitioners of nuclear materials management. At the same time, true ecumenical attitudes can be arrived at only by exploring alternatives.

Some may dismiss these concerns with a mere "It's only a matter of semantics." To which I usually reply, "Exactly so!" For it is with semantics (the study of meanings) that understanding comes.

I believe that the standards work of all disciplines will result in a better understanding of the contributions that those disciplines can bring to any particular industry. At the same time, the professionals contributing to those standards must be ready both to defend their work against misunderstanding and misuse and to adjust those standards when improvements can be effected.

Finally and specifically, I wish to mention those on the Subcommittee who contributed so much to ANSI N15.16: L.T. Hagie, D.E. Heagerty, L.D.Y. Ong, R.I. Post, and R.D. Smith, and (for his diplomatic way of settling disagreements) Subcommittee Chairman J.L. Jaech.

----

*The poet Robert Browning, when queried about the meaning of some of his verses, is reported to have replied: "When I wrote that, only God and I knew what it meant. Now, only God knows."