Statistical Standards and ISO, Part 2—Terminology

Article *in* Quality Engineering · April 2012

DOI: 10.1080/08982112.2012.654437

2 authors:

Stephen N. Luko
Retired

Mark E. Johnson
University of Central Florida

Quality Engineering
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lqen20

Statistical Standards and ISO, Part 2—Terminology

Stephen N. Luko [a] & Mark E. Johnson [b]

[a] Hamilton Sundstrand Corporation, Windsor Locks, Connecticut

[b] Department of Statistics, University of Central Florida, Orlando, Florida

Available online: 26 Mar 2012

To cite this article: Stephen N. Luko & Mark E. Johnson (2012): Statistical Standards and ISO, Part 2—Terminology, Quality Engineering, 24:2, 346-353

To link to this article: http://dx.doi.org/10.1080/08982112.2012.654437

Quality Engineering, 24:346–353, 2012

ISSN: 0898-2112 print/1532-4222 online
DOI: 10.1080/08982112.2012.654437

Taylor & Francis
Taylor & Francis Group

Reviews of Standards and Related Material*

Statistical Standards and ISO, Part 2—Terminology

Stephen N. Luko[1], Mark E. Johnson[2]

[1]Hamilton Sundstrand Corporation, Windsor Locks, Connecticut
[2]Department of Statistics, University of Central Florida, Orlando, Florida

ABSTRACT This is the second article in a series reviewing statistical standards published under the auspices of the International Organization for Standardization Technical Committee 69 on Applications of Statistical Methods. In this second article the emphasis is on terminology and nomenclature documents in the ISO 3534 series. In addition to reviewing these documents, a general discussion of lexicography is provided that reveals the deceptive difficulty in producing a coherent and comprehensive vocabulary for probability and statistics for use in international statistical standards.

KEYWORDS Applied statistics, General statistical terms, ISO, Probability terms, Statistical standards

INTRODUCTION

Terminology and nomenclature are the raison d'être for ISO TC69/SC1. To decipher this opening sentence, a terminology document covering French phrases and International Standards acronyms would prove useful, yielding a definition of *raison d'être* (the reason for existence) and a translation of ISO TC69/SC1 (International Standards Organization Technical Committee 69 on applications of statistical methods/Subcommittee One). Terminology standards document the "official" definitions of terms and symbols to facilitate the production and understanding of the standards that include them. Because some international standards are noted in contractual relationships, explicit and precise definitions are essential to facilitate legal proceedings (avoiding contentious discussions involving terminology).

The creation and development of international statistical terminology standards follows the same elaborate, multistage approval process as other statistical standards that was recently reviewed in Part One of this series of review articles on ISO/TC 69 standards (Boulanger and Johnson 2011). In brief, standards can require years of development, voting in stages from a new work item, to sundry committee drafts, draft international standards, and final draft international standards. Having withstood the consensus-building and review process, published terminology documents are recognized as *normative references* by other statistical standards published under TC69.

*Edited by Stephen N. Luko.

Address correspondence to Stephen N. Luko, 1 Hamilton Road, Hamilton Sundstrand Corporation, Windsor Locks, CT 06096, USA. E-mail: stephen.luko@hs.utc.com

TC 69/SC1 presides over three published terminology standards having a common designation of ISO 3534. The three currently published parts with their latest publication dates are

- ISO 3534-1:2006—Statistics—Vocabulary and symbols—Part 1: General statistical terms and terms used in probability
- ISO 3534-2:2006—Statistics—Vocabulary and symbols—Part 2: Applied statistics
- ISO 3534-3:1999—Statistics—Vocabulary and symbols—Part 3: Design of experiments

This series of documents is dynamic in that revisions occur on an ongoing basis. Revisions to Parts 1 and 2 are not yet officially registered work items, although preliminary planning for these efforts is underway. Part 3 has been revised through multiple committee draft stages and is at the second draft international standard voting stage (closed on January 2, 2012). Pending accommodation of comments and corrections, the document could be published late in 2012. A new fourth standard on Survey Sampling to be designated ISO 3534-4 is at the initial committee draft stage, for which a planned publication date in 2013 is projected.

This article begins with the development of terminology and nomenclature documents, following the advice of ISO TC37, which is charged with standardizing terminology standards. TC37 champions the use of concept diagrams that can delineate the interrelationship of terms and concepts, with an emphasis on subordinate relationships. The next three sections concern the three published ISO 3534 parts in turn. The final technical section deals with the aforementioned survey sampling standard under development and describes other SC1 collaborations within TC69 and across other ISO technical committees.

TERMINOLOGY AS A TECHNICAL DISCIPLINE

Many excellent sources for terminology in Probability and Statistics exist. For example ASTM's E456 is a terminology standard developed using a similar process to ISO standards. The general ASTM overview article was published in this journal during 2010 (see Ulman and Luko). ASQ's Glossary and Tables for Statistical Quality Control is more targeted to *Quality Engineering* professionals. Other comprehensive terminology surveys include works by Dodge, Cox, Commenges and Davison (2006); Everitt (2006); Freund and Williams (1991); Marriot (1990); Tietjen (1986); and Upton and Cook (2008).

Statisticians in general and statistical standards writers in particular are not naturally lexicographers. Familiarity with a multitude of definitions used in statistics and probability is an asset from a subject matter perspective but does not automatically translate to the capability of producing viable definitions that will pass an international muster. In an analogous sense, individuals who have completed many surveys are not necessarily adept at constructing excellent survey instruments. Fortunately, ISO Technical Committee 37 on terminology and other language and content resources has published standards governing terminology documents that SC1 has fully implemented. These guidelines when followed produce clean definitions that do not suffer common difficulties. The following are such recommendations and warnings:

1. Use concept diagrams to identify the interrelationships among terms and to provide a logical framework for all the terms defined in the standard.
2. Abide by the *substitution principle*.
3. Restrict one concept per term.
4. Avoid definitions in the negative.
5. Use singular form for terms that are nouns.
6. Avoid peripheral qualifiers not critical to the definition.
7. Avoid circular definitions.

An example of a concept diagram is given in Figure 1, adapted from a more complicated one in the proposed ISO-3534-4. The term *error of estimation* is partitioned into two subordinate terms and

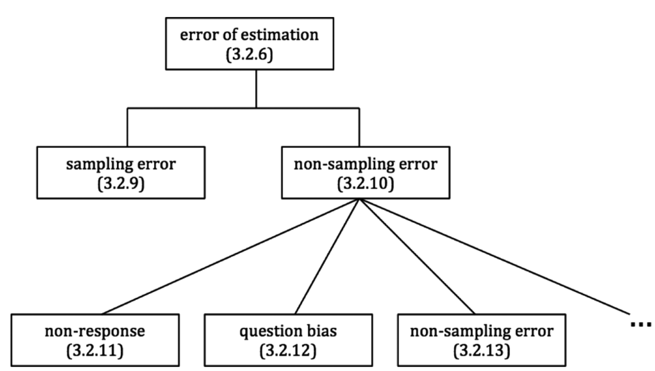

FIGURE 1 Sample concept diagram.

non-sampling error is illustrated with three specific examples and an indication with the ellipsis that there could be other possibilities.

The substitution principle requires that when a term is used within another definition and is substituted by its own definition there should be no change in the meaning of the text (up to minor syntax adjustments). This principle is particularly important because its violation indicates a fundamental problem with a definition that must be rectified.

The one concept per term restriction is the essence of eliminating ambiguity and confusion. The one term is the preferred term with the possibility of synonyms acceptable. In some circumstances there may be two terms that are used interchangeably for a single concept, with neither taking precedence. Definitions in the negative are not of direct use. For example, a definition of *interaction* that relates to nonparallel lines in a plot of the response variable versus a predictor variable with identified settings of a second predictor variable is a poor definition. Likewise, circular definitions are not helpful (e.g., random sample is a sample that is random). Circular definitions could be avoided by employing well-understood nontechnical words in some cases.

Implicit in this discussion is that any existing published statistical dictionaries are not themselves suitable to serve as normative references on terminology for international standards work. Many of the available statistical dictionaries are useful reference documents and are particularly helpful in covering rarely used terms or for consideration in determining whether a concept diagram is "correct" and whether the appropriate concept has indeed been captured. Seven such dictionaries are listed in the References. Terminology standards from member countries (e.g., the United States ASTM terminology document) can be consulted, although the general direction is to establish an international standard that is then provided as the candidate for a national standard. A further challenge is that the ultimate definition must be suitably clear that it can be translated into other languages. In this situation, translators from English to other languages greatly appreciate the concept diagrams given in the ISO 3534 series. The complete standards are translated (not merely the definitions), so care must also be given to the notes and examples so that they can be understood in multiple cultures. Bellos (2011) described many challenges in translating less technical material as found in the examples.

In terms of purely technical difficulties, the various published dictionaries do not provide panaceas for the standards writers. For example, *bias* is defined sometimes in the context of estimator bias and sometimes relative to slanted designs. Incorrect definitions

TABLE 1 **ISO 3534 Part 1, general statistical terms**

population	sample correlation coefficient	significance level
sampling unit	standard error	Type I error
sample	interval estimator	Type II error
observed value	statistical tolerance interval	statistical test
descriptive statistics	statistical tolerance limit p-value	
random sample	confidence interval	power of a test
simple random sample	one-sided confidence interval	power curve
statistic	prediction interval	test statistic
order statistic	estimate	graphical descriptive statisties
sample range	error of estimation	numerical descriptive statistics
mid-range	bias	class
estimator	unbiased estimator	class limits
sample median	maximum likelihood estimator	mid-point of class
sample moment of order k	estimation	class width
sample mean (average)	maximum likelihood estimation frequency	
sample variance	likelihood function	frequency distribution
sample standard deviation	profile likelihood function	histogram
sample coefficient of variation	hypothesis	bar chart
standardized sample random variable	null hypothesis	cumulative frequency
sample coefficient of skewness	alternative hypothesis	relative frequency
sample coefficient of kurtosis	simple hypothesis	cumulative relative frequency.
sample covariance	composite hypothesis	

are also sometimes encountered; for example, kurtosis is a measure of tail weight, though several dictionaries continue the erroneous notion of peakedness. Consulting multiple dictionaries for assistance is reminiscent of the story that says if one has one watch, one knows the time, whereas with two watches one is never sure. Few dictionaries agree on standard terms, so that the hard work of sorting out the essence of a term still requires the construction of concept diagrams and consulting the leading textbooks in statistics as to their use of terms.

GENERAL STATISTICAL TERMS AND TERMS USED IN PROBABILITY

ISO 3534—Part 1 was revised in 2006 to implement the methodology advocated by ISO TC37 in the development of terminology documents. In particular, concept diagrams were constructed to include all terms in the standard. These diagrams as noted in the previous section are designed to delineate the interrelationships among terms. Part 1 consists of two sections: general statistical terms (65 clauses) and terms used in probability (70 clauses). Ten concept diagrams are included in this section. The concept diagrams are an integral part of the standard and in many cases practically dictate the appropriate definition. Each clause corresponds to one concept and consists of a term in bold face followed by the definition and then notes and/or examples. Notes are provided especially for terms with mathematical definitions to enhance the clarity of the definition. The technical definition is required for the discriminating statistical standards writer. On the

TABLE 2 **ISO 3534 Part 1, probability terms**

Sample space	probability mass function	hypergeometric distribution
event	mode of probability mass function	negative binomial distribution
complementary event	probability density function	normal distribution (Gaussian distribution)
independent events	mode of probability density function	standardized normal distribution (standardized Gaussian distribution)
probability of an event A	discrete random variable	lognormal distribution
conditional probability	continuous random variable	t distribution
distribution function of a random variable X	centred random variable	degrees of freedom
family of distributions	standardized probability distribution	F distribution
parameter	standardized random variable	gamma distribution
random variable	moment of order r (rth moment)	chi-squared distribution (c2 distribution)
probability distribution (distribution)	mean	exponential distribution
expectation	variance	beta distribution
p-quantile (p-fractile)	standard deviation	uniform distribution
median	coefficient of variation	Type I extreme value distribution (Gumbel distribution)
quartile	coefficient of skewness	Type II extreme value distribution (Frechet distribution)
univariate probability distribution	coefficient of kurtosis	Type III extreme value distribution (Weibull distribution)
multivariate probability distribution	joint moment of orders r and s	multivariate normal distribution
marginal probability distribution	joint central moment of orders r and s	bivariate normal distribution
conditional probability distribution	covariance	standardized bivariate normal distribution
regression curve	correlation coefficient	sampling distribution
regression surface	multinomial distribution	probability space
discrete probability distribution	binomial distribution	sigma algebra of events
continuous probability distribution	Poisson distribution	probability measure.

other hand, notes provide further clarification on aspects of a term that are helpful in comprehending their use but are not essential as part of the definition. Examples are offered to place the term in context and to highlight its use with other related terms.

The preferred terms among the general statistical terms are shown in Table 1. The preferred terms among the terms used in probability are shown in Table 2.

APPLIED STATISTICS

The terms in ISO 3534—Part 2 support many of the statistical standards that have been published over the years across the subcommittees of TC69. Terms in the areas of statistical process management (including quality control), statistical metrology (covering specifications and test results), inspection, acceptance sampling, and bulk sampling are extensively covered. This standard has been updated multiple times on a roughly 5- to 10-year cycle. There are so many terms included that this standard is approaching the need to be split into more manageable documents. TC 37 recommends a range of 50–200 concepts in a single terminology document. The organization of Part 2 is given in the following outline:

1. Data generation and collection
   - 1.1 Systems of reference values for characteristics
   - 1.2 Sources of data
   - 1.3 Types of sampling
2. Statistical process management
   - 2.1 General process-related concepts
   - 2.2 Variation related concepts
   - 2.3 Control-related charts
   - 2.4 Control chart components
   - 2.5 Fundamental terms related to process performance and process capability
   - 2.6 Process performance ⟨measured data⟩
   - 2.7 Process capability ⟨measured data⟩
3. Specifications, values, and test results
   - 3.1 Specification-related concepts
   - 3.2 Determination of characteristics and quantities
   - 3.3 Properties of test and measurement methods
   - 3.4 Properties of test and measurement results
   - 3.5 Capability of detection
4. Inspection and general acceptance sampling
   - 4.1 Types of inspection
   - 4.2 Types of acceptance sampling inspection
   - 4.3 Acceptance sampling inspection system aspects
   - 4.4 Acceptance criteria
   - 4.5 Types of operating characteristic curves
   - 4.6 Terms relating to operating characteristics
   - 4.7 Outgoing quality concepts and average inspection effort
5. Sampling of bulk materials
   - 5.1 Concepts related to bulk materials
   - 5.2 Bulk sampling aspects
   - 5.3 Bulk sample preparation
   - 5.4 Procedural aspects

We give two definitions from 3534-2 as representative examples from this standard. Note that when an additional is used in the definition or note, that definition references the term being used back to its definition.

- 1.2.1 **population**

  (reference) totality of items (1.2.11) under consideration

  NOTE 1: A population can be real and finite or hypothetical and infinite.

  NOTE 2: Extended sampling (1.3.1) from a finite real population can give rise to the generation of actual relative frequencies or frequency distributions (2.5.1). Alternatively, or arising from this, a theoretical model of the hypothetical population based on probability distributions can be derived. This enables predictions to be made.

  NOTE 3: A population can be the result of an ongoing process that may include future output.

  NOTE 4: A population can consist of distinguishable objects or bulk material.
- 3.3.2 **bias**

difference between the expectation of a test result (3.4.1) or measurement result (3.4.2) and a true value (3.2.5).

NOTE 1: Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the true value is reflected by a larger bias value.

NOTE 2: The bias of a measuring instrument is normally estimated by averaging the error of indication over an appropriate number of repeated measurements. The error of indication is the indication of a measuring instrument minus a true value of the corresponding input quantity.

NOTE 3: In practice, the accepted reference value (3.2.7) is substituted for the true value 3.3.2.

DESIGN OF EXPERIMENTS

The terms in the proposed revision to ISO 3534—Part 3 are organized in three sections consisting of *general terms* (40 clauses), *arrangements of experiments* (31 clauses), and terms associated with *methods of analysis* (12 clauses). There are also nine concept diagrams in Annex A with each term in at least one of these figures. Annex B is a normative annex outlining the methodology used to develop vocabulary (identical to the same annex used with Parts 1 and 2). Annex C contains two checklists identifying key items for use in designing experiments. Annex D provides a description of experimental design from a systems design perspective.

The general terms are shown in Table 3a; terms related to type of experiments in Table 3b and terms related to analysis in Table 3c.

TABLE 3a **ISO 3534—Part 3, design of experiments, general terms**

experiment	curvature
model	contrast
response variable (output variable)	orthogonal contrast
predictor variable	experimental unit
Factor	block
residual error (error term)	blocking
residual	designed experiment
variance component	experimental design
pure random error (pure error)	experimental plan
misspecification error	randomization
design region (design space)	orthogonal array
factor level	degrees of freedom
run (experimental treatment)	one factor experiment
factor effect	two-factor experiment
main effect	k-factor experiment
dispersion effect	replication
interaction	cube point
confounded effect	star point
confounding	centre point
Alias	rotatability

TABLE 3b **ISO 3534—Part 3, design of experiments, terms related to arrangements of experiments**

factorial experiment	Youden square design
full factorial experiment	split-plot design
fractional factorial experiment	two-way split-plot design (split block design)
two-level experiment	response surface design
2k factorial experiment	mixture design
2k-p factorial experiment	balanced nested design (fully nested design)
design resolution	staggered nested design
screening design	optimal design
block design	design matrix
randomized block design	D-optimal design
Latin square design	A-optimal design
Graeco-Latin square design	G-optimal design
incomplete block design	orthogonal design
balanced incomplete block design	saturated design
partially balanced incomplete block design	completely randomized design.

TABLE 3c **ISO 3534—Part 3, design of experiments, terms related to methods of analysis**

graphical method
main effects plot
interaction plot
quantile plot of effects
residual plot
method of least squares
regression analysis
analysis of variance
fixed effects analysis of variance
random effects analysis of variance
mixed model analysis of variance
analysis of covariance

SURVEY SAMPLING AND OTHER ACTIVITIES

The proposed standard ISO 3534-4 on survey sampling is in its formative stages; international experts have worked on this document for just a few years. The official launch of this effort via a new work item (subject to balloting and the cooperation of at least five member countries) was deliberately held up until a complete committee draft was completed. Once launched, the ISO Secretariat following the Directives maintains a strict clock vigil to assure that progress is made in a timely fashion. Part 4 consists of one section containing general terms and a second section on terms related to estimation.

TABLE 4a ISO 3534—Part 4, survey sampling, general terms

population	sampling plan
finite population	sampling frame
subpopulation	dual frame
superpopulation	area frame
sampling unit	sampling design
cluster	systematic sampling
stratum	quasi-random sampling
sample	judgment sampling (purposive sampling)
sample size	stratified sampling
random sample	stratified simple random sampling
random sampling	proportional allocation
simple random sample	optimal allocation
probability sample	Neyman allocation
representative sample	poststratification
selection probability	cluster sampling
sampling	post cluster sampling
sampling with replacement	multi-stage sampling
sampling without replacement	two-stage sampling
census	multi-stage cluster sampling
sample survey	inverse sampling
survey sampling	sampling proportional to size
pilot survey	quota sampling
opinion survey	

The preferred terms in the general term section are shown in Table 4a. Table 4b shows terms related to estimation.

Because the survey sampling standard is in its infancy, future revisions to an eventually adopted international standard will likely expand with respect to the number of terms included. The present proposed version focuses on finite population sampling and future revisions will likely incorporate more terms specific to large scale complex sampling efforts.

TABLE 4b ISO 3534—Part 4, survey sampling, terms related to estimation

population parameter	non-response
Estimate	question bias
estimator	observational error
estimation	regression estimator
standard error	auxiliary data
error of estimation	ratio estimator
Bias	Horvitz-Thompson estimator
unbiased estimator	sampling fraction
sampling error	finite population correction.
non-sampling error	

FUTURE TC69/SC1 ACTIVITIES

Subcommittee one (SC1) on terminology and nomenclature will continue to update, improve, and expand its core documents in the ISO 3534 series. The inclusion of concept diagrams as advocated by TC37 has had major impacts on identifying difficulties with previous definitions and also facilitates the translation of the terminology documents into languages besides the official English and French versions. Part 2 on applied statistics will likely split into several smaller and more manageable standards likely including statistical process management, measurement systems, and acceptance sampling. According to ISO TC37, this natural evolution should take place as documents exceed 200 or so concepts. On an ongoing basis, suggested revisions regarding the existing set of terminology documents are welcome and can be sent to Standards@asq.org.

The former Chair of ISO TC69, Christophe Perruchet (who tragically died in June 2011 in a motorcycle accident), was adamant about keeping the task of terminology and nomenclature under the control of SC1 and avoiding the temptation of drift of this activity to other subcommittees. With this management position in place, SC1 formed a new working group, WG6, at the Berlin July 2011 meeting to work with project leaders from subcommittees 7 and 8 on three new terminology standards on terms related to (1) Lean and Six Sigma methodology, (2) robust parameter design, and (3) quality function deployment.

As the published documents evolve over time and new terminology standards are published, the *desiradata* of these documents will continue to be as follows:

- Provide a coherent and harmonized vocabulary and nomenclature.
- Maintain normative references for use of other SCs within TC69 specifically and throughout ISO generally. Establish consistency of terminology usage.
- Facilitate translation of these definitions into other member bodys via concept diagrams and simplicity and clarity of the English versions.

Contributors to terminology documents or dictionaries are providing a much-needed service to the profession but one with somewhat limited rewards. Relatively few people manage to obtain recognition of authorship of terms; still, these project leaders acknowledge the contributions made to terminology standards from many peers in specialized fields. The compilations of terms most often do not correspond to drawing upon published papers, so that terminology standards are not a compilation of a body of research. Thus, much of the enjoyment in participating in terminology work resides in satisfaction of achieving a high degree of clarity in the vocabulary, in providing invaluable tools to the profession, and in establishing friendships with other contributors. Perhaps, as Winchester (1998) noted in regards to the massive undertaking of the *Oxford Dictionary* of the English language, a certain level of *madness* accompanies the participation in these projects!

ACKNOWLEDGMENT

We are grateful to Dr. Michèle Boulanger for providing insights on the ISO 3534 series regarding contractual implications.

ABOUT THE AUTHORS

Stephen N. Luko is a statistician at Hamilton Sundstrand Corporation, a division of United Technologies. He is a fellow of ASTM International, a member of ASA, a senior member of ASQ and a certified Quality and Reliability Engineer. He is the editor of the Statistical Standards column in *Quality Engineering*.

Dr. Mark E. Johnson is professor, Department of Statistics, University of Central Florida. He is a fellow of the American Statistical Association, a member of ASQ, and a Chartered Statistician with the Royal Statistical Society. He is currently chairman of TC69/Sub-committee 1 on terminology and symbols.

REFERENCES

ASTM E456-08. (2011). *Standard Terminology Relating to Quality and Statistics*. ASTM Committee E11 on Quality and Statistics, ASTM International, West Conshohocken, PA.

Bellos, D. (2011). *Is That a Fish in Your Ear*. New York: Faber and Faber.

Boulanger, M., Johnson, M. E. (2011). Statistical standards and ISO, Part 1. *Quality Engineering*, 23(1).

Dodge, Y., Cox, D. R., Commenges, D., Davison, A. (2006). *The Oxford Dictionary of Statistical Terms.* Oxford, UK: Oxford University Press.

Everitt, B. S. (2006). *The Cambridge Dictionary of Statistics*, 3rd ed. Cambridge, MA: Cambridge University Press.

Freund, J. E., Williams, F. J. (1991). *Outline of Basic Statistics: Dictionary and Formulas.* New York: Dover.

(2004). *Glossary and Tables for Statistical Quality Control*, 4th edition. Milwaukee, WI: ASQ Press.

Marriot, F. H. C. (1990). *A Dictionary of Statistical Terms*, 5th ed. London: The International Statistical Institute.

Tietjen, G. L. (1986). *A Topical Dictionary of Statistics.* New York: Chapman & Hall.

Ullman, N. R., Luko, S. (2010). Statistical standards and ASTM. *Quality Engineering*, 22(4):358–363.

Upton, G., Cook, I. (2008). *A Dictionary of Statistics.* Oxford, UK: Oxford University Press.

Winchester, S. (1998). *The Professor and the Madman (A Tale of Murder, Insanity and the Making of the Oxford English Dictionary)*. New York/Oxford: Harper Collins.