ASME B89.7.4.1-2005
(Technical Report)

Measurement Uncertainty and Conformance Testing: Risk Analysis

AN ASME TECHNICAL REPORT

The American Society of Mechanical Engineers

ASME B89.7.4.1-2005
(Technical Report)

Measurement Uncertainty and Conformance Testing: Risk Analysis

AN ASME TECHNICAL REPORT

The American Society of Mechanical Engineers

Three Park Avenue • New York, NY 10016

Date of Issuance: February 3, 2006

This Technical Report will be revised when the Society approves the issuance of a new edition. There will be no addenda issued to this edition.

The American Society of Mechanical Engineers
Three Park Avenue, New York, NY 10016-5990

CONTENTS

FOREWORD

The ISO Guide to the Expression of Uncertainty in Measurement (GUM) is now the internationally accepted method of expressing measurement uncertainty [1]. The U.S. has adopted the GUM as a national standard [2]. The evaluation of measurement uncertainty has been applied for some time at national measurement institutes; more recently, increasingly stringent laboratory accreditation requirements have increased the use of measurement uncertainty analysis in industrial calibration laboratories. In some cases, measurement uncertainty calculations have even been applied to factory floor measurements.

Given the potential impact to business practices, national and international standards committees are working to publish new standards and technical reports that will facilitate the integration of the GUM approach and the consideration of measurement uncertainty in product conformance decisions. In support of this effort, the ASME B89 Committee for Dimensional Metrology has formed Subcommittee 7 — Measurement Uncertainty.

Measurement uncertainty has important economic consequences for calibration and inspection activities. In calibration reports, the magnitude of the uncertainty is often taken as an indication of the quality of the laboratory, and smaller uncertainty values generally are of higher value and cost. In industrial measurements, uncertainty has an economic impact through the decision rule employed in accepting and rejecting products. ASME B89.7.3.1, Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance to Specifications, addresses the role of measurement uncertainty when accepting or rejecting products based on a measurement result and a product specification.

With significant economic interests at stake, it is advisable that manufacturers guard against accepting bad products and rejecting good ones. Even with a very good measurement system, there will be some risk of decision errors, with cost impacts that vary depending upon the nature of the product and its intended end use. While the evaluation of measurement uncertainty is a technical activity well-described in the GUM, the selection of a decision rule is a business decision that involves cost considerations.

ASME B89.7.3.1 provides uniform, unambiguous terminology for documenting a decision rule. It describes the relationship between the conformance zone (locating conforming characteristics) and the acceptance zone (locating acceptable measurement results). This Technical Report addresses the problem of determining the gauging limits (or test limits) that define the boundaries of the acceptance zone. The limits are chosen to balance the risks of the two types of decision errors, whose relative magnitudes depend upon product-specific economic factors that are outside the scope of this Report.

ASME STANDARDS COMMITTEE B89
Dimensional Metrology

(The following is the roster of the Committee at the time of approval of this Technical Report.)

OFFICERS

B. Parry, *Chair*
D. E. Beutel, *Vice Chair*
M. Lo, *Secretary*

COMMITTEE PERSONNEL

D. E. Beutel, Caterpillar
J. B. Bryan, Bryan Associates
T. E. Carpenter, U.S. Air Force Metrology Labs
T. Charlton, Jr., Charlton Associates
G. A. Hetland, International Institute of Geometric Dimensioning and Tolerancing
R. J. Hocken, University of North Carolina, Charlotte
M. Liebers, Professional Instruments Co.
M. Lo, The American Society of Mechanical Engineers
B. Parry, The Boeing Co.
S. D. Phillips, National Institute of Standards and Technology
J. G. Salsbury, Mitutoyo America
D. A. Swyt, National Institute of Standards and Technology
B. R. Taylor, Renishaw PLC
R. C. Veale, National Institute of Standards and Technology

SUBCOMMITTEE 7 — MEASUREMENT UNCERTAINTY

G. A. Hetland, *Chair,* International Institute of Geometric Dimensioning and Tolerancing
D. A. Swyt, *Vice Chair,* National Institute of Standards and Technology
W. Beckwith, Brown & Sharpe
D. E. Beutel, Caterpillar
B. Borchardt, National Institute of Standards and Technology
J. Buttress, Hutchinson Technology, Inc.
T. E. Carpenter, U.S. Air Force Metrology Labs
T. Charlton, Jr., Charlton Associates
W. T. Estler, National Institute of Standards and Technology
H. Harary, National Institute of Standards and Technology
M. Liebers, Professional Instruments Co.
B. Parry, The Boeing Co.
S. D. Phillips, National Institute of Standards and Technology
J. Raja, University of North Carolina, Charlotte
J. G. Salsbury, Mitutoyo America
C. Shakarji, National Institute of Standards and Technology
P. Stein, P. G. Stein Consultants

PERSONNEL OF WORKING GROUP B89.7.4 — GENERAL PRINCIPLES

W. T. Estler, *Chair,* National Institute of Standards and Technology
W. Beckwith, Brown & Sharpe
J. Buttress, Hutchinson Technology, Inc.
T. Charlton, Jr., Charlton Associates
T. D. Doiron, National Institute of Standards and Technology
H. Harary, National Institute of Standards and Technology
B. Parry, The Boeing Co.
S. D. Phillips, National Institute of Standards and Technology
J. G. Salsbury, Mitutoyo America

MEASUREMENT UNCERTAINTY AND CONFORMANCE TESTING: RISK ANALYSIS

1 SCOPE

This Technical Report provides guidelines for setting gauging (or test) limits in support of accept/reject decisions in workpiece inspections, instrument verifications, and general conformance tests where uncertain numerical test results are compared with specified requirements.

In accepting or rejecting workpieces or instruments based on the results of inspection measurements, the presence of unavoidable measurement uncertainty introduces the risk of making erroneous decisions. By implementing a decision rule that defines a range of acceptable measurement results, one can balance the risks of rejecting conforming workpieces or instruments and accepting nonconforming ones.

2 DEFINITIONS AND TERMINOLOGY

For the purposes of this Technical Report, the following definitions apply [1–4]:

accept–reject measurement: measurement made for the purpose of accepting or rejecting a workpiece, workpiece feature, or measuring instrument [4].

acceptance: decision that the measured value of a characteristic satisfies the acceptance criteria.

acceptance criterion: specification criterion for acceptance of a workpiece, workpiece feature, or measuring instrument based upon the result of a measurement or test.

NOTE: The most common acceptance criterion for accept/reject decisions is acceptance when the measured characteristic lies in the acceptance zone and rejection otherwise.

acceptance zone: set of values of a characteristic, for a specified measurement process and decision rule, that results in product acceptance when a measurement result is within this zone [3].

binary decision rule: decision rule with only two possible outcomes, either acceptance or rejection [3].

characteristic: property that helps to identify or differentiate between items of a given population [5, para. 1.5.1]. In this Report, a characteristic is typically a workpiece feature or the error of a measuring instrument subject to a conformance test.

conformance test: measurement of a characteristic in order to decide conformance or nonconformance with specifications.

conforming: a characteristic is conforming if its true value lies within or on the boundary of the tolerance zone.

NOTE: In ASME B89.7.2-1999, conforming is defined as having a measured value lying within or on the boundary of the allowable tolerance band. This definition would be correct if *measured* were changed to *true.*

consumer's risk: probability of a pass (or Type II) error. (The cost of such an error is generally borne by the consumer.)

decision rule: documented rule that describes how measurement uncertainty will be allocated with regard to accepting or rejecting a product according to its specification and the result of a measurement [3].

fail error: rejection, as a result of measurement error, of a characteristic whose true value is within specified tolerances (also known as a Type I error) [4].

gauging limits: specified limits of a measured value [4].

guard band: magnitude of the offset from a specification limit to an acceptance or rejection zone boundary [3].

inspection: activities such as measuring, examining, testing, and gauging one or more characteristics of a product or service, and comparing with specified requirements to determine conformity [5, para. 1.2.1].

inspection by variables: method that consists in measuring a quantitative characteristic for each item of a population or a sample taken from this population [6, para. 3.1].

NOTE: Inspection by variables may be compared with a related concept, inspection by attributes. In the latter, one simply notes the presence (or absence) of some characteristic of an item, while in the former one measures and records a numerical value of a characteristic, with reference to a continuous scale. In the inspection of a ballpoint pen, for example, an inspection by attributes might consist of noting whether or not the pen will write, while an inspection by variables might require a measurement of the pen's ball diameter and a comparison with a tolerance.

measurand: particular quantity subject to measurement [2, para. 2.6; 7, para. B.2.9].

measured value: value obtained by measurement.

NOTE: The measured value is the result of the measurement [2, para. 3.1] and is the value attributed to the measurand after performing a measurement.

measurement capability index, C_m: in the case of measuring a characteristic for conformance to a two-sided tolerance zone of width T, $C_m = T/4u_m$, where u_m is the standard uncertainty associated with the estimate of the characteristic; for a one-sided tolerance zone of width T, $C_m = T/2u_m$; and in the case of calibration or verification of a measuring instrument with specified maximum permissible error $\pm MPE$, $C_m = MPE/2u_e$, where u_e is the standard uncertainty associated with the estimate of the instrument error.

nonacceptance: decision that the measured value of a characteristic does not satisfy the acceptance criteria.

nonconforming: a characteristic is nonconforming if its true value lies outside the boundary of the tolerance zone.

NOTE: In ASME B89.7.2-1999, nonconforming is defined as having a measured value lying outside the boundary of the allowable tolerance band. This definition would be correct if *measured* were changed to *true.*

pass error: acceptance, as a result of measurement error, of a characteristic whose value is outside specified tolerances (also known as a Type II error) [4].

process distribution: probability distribution characterizing reasonable belief in values of a characteristic resulting from a manufacturing process.

NOTE: The form of this distribution can be inferred from a frequency distribution (usually displayed in a histogram) of measured characteristics from a large sample of items.

producer's risk: probability of a fail (or Type I) error. (The cost of such an error is generally borne by the producer.)

rejection: see *nonacceptance.*

rejection zone: set of values of a characteristic, for a specified measurement process and decision rule, that results in product rejection when a measurement result is within this zone [3].

specification limits: see *tolerance limits.*

test limits: see *gauging limits.*

tolerance: total amount by which a specific characteristic is permitted by specifications to vary.

NOTE: The tolerance is the difference between the upper and lower specification limits [5, para. 1.4.4; 8, para. 1.3.3].

tolerance interval: region between, and including, the tolerance limits [5; para. 1.4.5].

tolerance limits: specified values of the characteristic, giving upper and/or lower bounds of the permissible value [5, para. 1.4.3].

lower tolerance limit (T_L): specification limit that defines the lower conformance boundary for an individual unit of a manufacturing or service operation.

upper tolerance limit (T_U): specification limit that defines the upper conformance boundary for an individual unit of a manufacturing or service operation.

NOTE: For a single-sided conformance test, there is only a single tolerance limit.

tolerance zone: see *tolerance interval.*

3 INSPECTION MEASUREMENTS AND PASS/FAIL DECISIONS

In a typical inspection measurement or conformance test, a characteristic or feature is measured[1] and the result compared with a specified acceptance criterion in order to establish whether there is an acceptable probability that the characteristic conforms to its tolerance requirements. Such a conformance test consists of the following sequence of three operations:

(a) measure a characteristic of interest

(b) compare the result of the measurement with a specified requirement

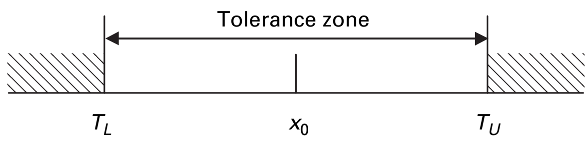
(c) decide on the subsequent action

In practice, once the measurement data are in hand, the comparison/decision operations are typically implemented by way of a decision rule that depends on the measurement result and its associated uncertainty, the specified requirement, and the chances and consequences of making an erroneous decision. The producer is generally responsible for choosing the decision rule to be used when making conformance decisions.

Documentary guidance is available regarding the formulation of a decision rule. ASME B89.7.3.1-2001 [3], for example, provides a unified set of guidelines for documenting a chosen decision rule, including an explicit description of the role of the measurement uncertainty in setting the test limits (or guard bands).

In an industrial and commercial setting, inspection measurement or conformance test procedures are designed to obtain, at reasonable cost, information that will enable rational business decisions to be made. Money spent to reduce uncertainty below the level at which a rational business decision can be made will usually lead to lost revenue. An inspection sequence with its associated decision rule (measure ⇒ compare/decide) is thus necessarily very closely tied to matters such as costs and risks. As such, the design of an effective inspection measurement or conformance test is not a purely technical exercise, but also depends upon economic factors that are specific to the particular enterprise. For this reason, generic or default decision rules (such as those proposed in ISO 14253-1) that are based only on the measurement uncertainty and with no consideration of costs can be inadequate for maximizing return on investment.

[1] This Report considers only scalar characteristics that are measurable on a continuous scale. An inspection measurement of such a characteristic is called *inspection by variables.*



GENERAL NOTE: The tolerance zone [5,8] is equivalent to the specification zone [3].

Fig. 1 Tolerance Zone

4 FREQUENCY DISTRIBUTIONS: VARIABLE PRODUCTION PROCESSES AND NOISY MEASUREMENTS

4.1 Specification and Tolerance

The following simple one-dimensional example will serve to illustrate in detail the development of a pass/fail conformance test procedure for a manufactured workpiece. A manufacturer produces metal spacers of nominal length x_0. The design specification includes a tolerance T and calls for x_0 to lie at the center of a tolerance zone of length T. An acceptable spacer must therefore have a length X in the range $T_L \le X \le T_U$, where the lower tolerance limit $T_L = x_0 - T/2$ and the upper tolerance limit $T_U = x_0 + T/2$. The tolerance is simply related to the tolerance limits by $T = T_U - T_L$, as shown in Fig. 1. A spacer is said to be conforming if its length X lies in the specification zone and nonconforming otherwise.

4.2 Process Variation

By design and adjustment, the manufacturing process can be arranged so that, on average, it produces a spacer whose length equals the nominal value x_0. Due to unpredictable and unavoidable process variations, however, there will be some distribution of actual lengths in any particular batch of parts. The nature of this distribution can be studied by measuring a large sample of spacers and plotting the results in a histogram. In such a study, any nonrepeatability in the measuring system will be superimposed on the variability due to the production process. In studying process variation, the measurement data can be corrected for this effect (see para. 4.5).

Figure 2 shows a histogram for the lengths of a batch of spacers produced by a hypothetical production process.[2] The vertical axis shows the fraction (or relative frequency) of parts whose lengths lie in the various narrow bins distributed along the horizontal (length) axis. The width of the histogram is a measure of the variability of the production process. The data in Fig. 2 show that most of the spacers are conforming, but there are clearly some nonconforming ones in the batch. The goal of a conformance test plan is to detect and remove these bad parts.

[2] The data in Fig. 2 are taken to be the true lengths of the sample.

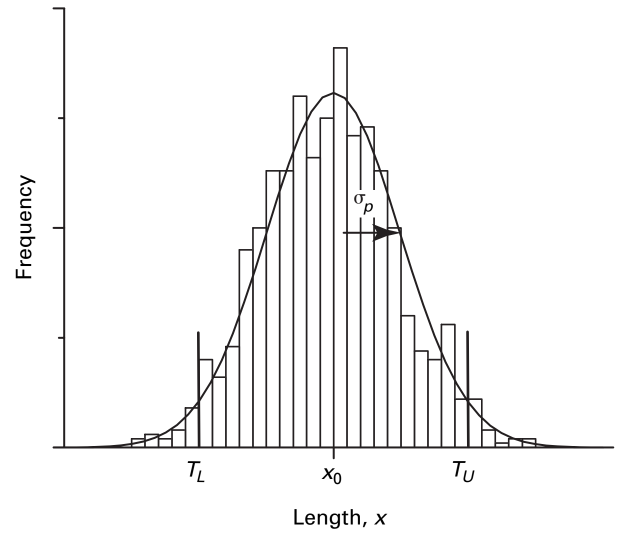


Fig. 2 Frequency Distribution of a Sample of Spacers

Denoting by $x_1, x_2, ..., x_N$ the individual lengths of a sample of N spacers, it is common to summarize the characteristics of the sample by calculating the sample mean, $\overline{x}$, and the sample variance, s^2, given by

$$\overline{x} = \frac{1}{N}\sum_{k=1}^{N} x_k$$

and

$$s^2 = \frac{1}{N-1}\sum_{k=1}^{N}(x_k - \overline{x})^2$$

The square root of the sample variance is called the sample standard deviation

$$s = \sqrt{\sum_{k=1}^{N}\frac{(x_k - \overline{x})^2}{N-1}} \qquad (1)$$

For a stable manufacturing process, the sample parameters $\overline{x}$ and s are, respectively, estimates of the process mean μ_p and process standard deviation σ_p that would characterize the average length and dispersion of a very large ($N \to \infty$) sample of spacers.

In many cases, the observed variability, as displayed in a histogram, can be well-approximated by a Gaussian (or normal) curve. The solid line in Fig. 2 shows such a curve overlaid on the length measurement data.

A Gaussian distribution is uniquely specified by its mean, μ, and standard deviation, σ, and these two numbers provide a convenient way to summarize the production process.

In this Report it is assumed that the frequency distribution of produced spacers is a Gaussian distribution with mean $\mu = x_0$, the design length, and standard deviation $\sigma = \sigma_p$, estimated by Eq. (1). If N workpieces have been measured, with $N > 30$ or so, then the relative uncertainty in the estimate of σ_p will be less than 10% or so.

The histogram in Fig. 2 is a measured frequency distribution of spacer lengths. The Gaussian approximation is given by

$$f(x) = \frac{1}{\sigma_p \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - x_0}{\sigma_p}\right)^2\right] \tag{2}$$

The meaning of $f(x)$ is as follows: Given a large sample of parts, the fraction of them with lengths between x and $x+\Delta x$ is just $f(x)\Delta x$. If the size of the sample is N, then approximately $Nf(x)\Delta x$ of them would have lengths in the interval $[x, x+\Delta x]$.

4.3 Process Capability Index

In statistical quality control, a common measure of the quality of a production process is the inherent process capability index, C_p, defined by

$$C_p \equiv \frac{T_U - T_L}{6\sigma_p} = \frac{T}{6\sigma_p} \tag{3}$$

This particular definition is chosen so that $C_p = 1$ for a process characterized by a value of σ_p equal to one-sixth of the tolerance. The choice of the factor of 6 in Eq. (3), rather than a factor of 3 or 10, is clearly arbitrary, but C_p as defined does give a useful way to compare degrees of variability of various processes.

Given C_p for a centered process [i.e., $x_0 = (T_U + T_L)/2$] with a Gaussian frequency distribution, one can calculate the fraction of spacers that will conform, in the absence of process drift, with specification (see Mandatory Appendix I). This is just the fraction of the area under the process frequency distribution [Eq. (2)] between the tolerance limits of $x_0 \pm T/2 = x_0 \pm 3C_p\sigma_p$.

Figure 3 and Table 1 show how the yield of conforming parts increases with increasing process capability. For $C_p = 1$, the fraction conforming is 0.997 (99.7%), so that 0.003 of them (0.3%) would be expected to be out of tolerance. For a more variable process ($4\sigma_p = T$ or $C_p = 2/3$), the fraction conforming is about 0.96 (96%), so that on average about 0.04 (4%) of manufactured spacers would be out of tolerance.

Numerical Example. Values for the process parameters in the following example are taken from ASME B89.7.2 [4]. The process density for a feature of length is centered at a mean value $x_0 = 1\,500$ mm. The process standard deviation is $\sigma_p = 0.12$ mm. The upper and lower tolerance limits are $T_U = 1\,500.2$ mm and $T_L = 1\,499.8$ mm, so that $C_p = (T_U - T_L)/6\sigma_p \approx 0.55$. For this process, Table 1 shows a fraction 0.902 of conforming parts, so that about 9.8% of production would be out of tolerance. If the manufacturer simply shipped every spacer produced, nearly one in ten would be nonconforming.

By reducing the variability of the process (increasing C_p), the manufacturer could reduce the fraction of spacers that fail to meet specification. Of course, such process improvements cost money and, as shown in Fig. 3, there is a diminishing return on such investment as the process becomes increasingly tightly controlled. At some point it will usually be more economical to invest, not in process improvement, but rather in workpiece inspection. In such a case, it would be cheaper to detect and remove nonconforming parts rather than to try to prevent their production. The exact nature of such tradeoffs between process improvement and workpiece inspection will depend upon the economics of the marketplace.

4.4 Generalizations

In the remainder of this Report, it is assumed that (1) the process is centered, so that the average length of a spacer (the process mean) equals the design length, $\mu_p = x_0 = (T_L + T_U)/2$, and (2) the variability of the process is well-characterized by a Gaussian frequency distribution. In a case where these assumptions are not valid, the risk calculations that are developed in detail in Mandatory Appendix II can be modified to account for the characteristics of the actual production process.

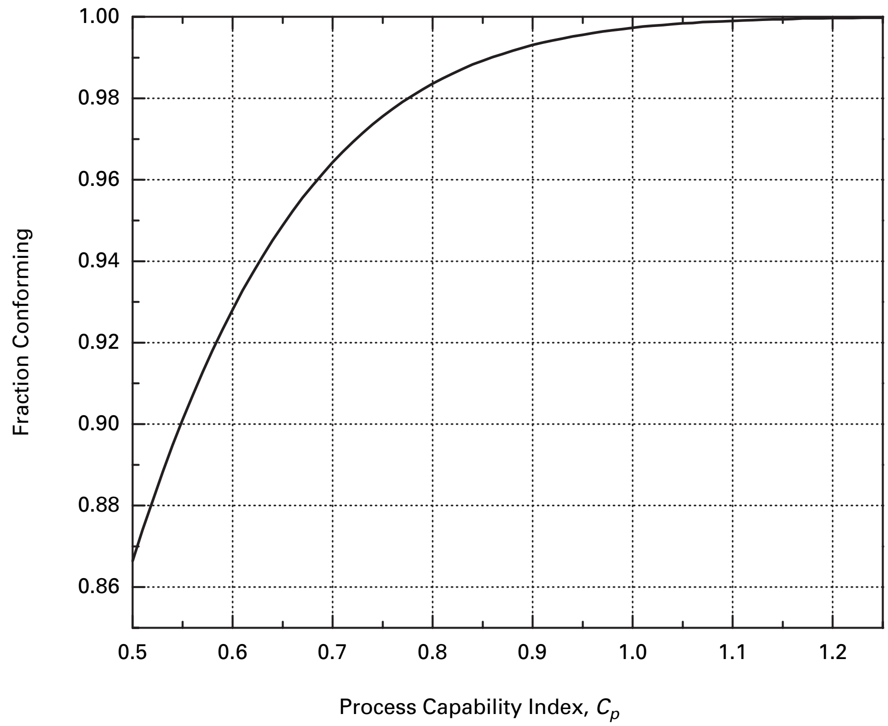
Many process capability indices have been suggested for processes that do not satisfy one or both of the assumptions above. A noncentered Gaussian process, for example, where the average spacer length, μ_p, does not lie at the center of the tolerance zone, can be characterized by a more general process capability index, C_{pk}, defined by

$$C_{pk} = \min\left[\frac{T_U - \mu_p}{3\sigma_p}, \frac{\mu_p - T_L}{3\sigma_p}\right]$$

In general, $C_{pk} \le C_p$, with equality for a centered process, i.e., $\mu_p = x_0$.

It should be recognized that while process capability indices such as C_p and C_{pk} can be useful summary parameters for stable production processes, such parameters add no new information. All such indices are calculated from more basic quantities, such as T_L, T_U, σ_p, and μ_p, that characterize the process and the tolerance requirements. What is needed in order to calculate the risks associated with erroneous accept/reject decisions is a probability density function that characterizes belief in possible values of a workpiece feature (such as a spacer length) before it is measured. Such a probability density is assigned based on knowledge of the process, usually acquired by measurements of a suitable sample of workpieces.

If the probability density is Gaussian, then the use of C_p, C_{pk}, or some other index might be useful in simplifying the notation in calculations and for communicating results. In a case where the probability density is not Gaussian, the risks can still be calculated, given a



GENERAL NOTE: For $C_p = 1$, about 3 parts in 1,000 will be nonconforming. The shape of the curve suggests that for a well-controlled process ($C_p \geq 1$ or so), improving the yield by means of better process control can become increasingly difficult.

Fig. 3 Fraction of Workpieces Conforming Versus Process Capability Index

suitable analytic form for the density. A useful discussion of a variety of capability indices and the effects of non-Gaussian process densities may be found elsewhere [9].

4.5 Nonrepeatable Measurement Results

It is a very common experience in industrial metrology for repeated measurements to yield different results. Among the many sources of measurement variability are small setup variations and instabilities, vibration, electrical noise, dirt, and operator effects. Because of this lack of repeatability, part of the observed variability when measuring a batch of parts will be due to the measurement system.

Measurement repeatability can be studied by repeatedly measuring a stable artifact and examining the frequency distribution of the results. Such repeatability data will typically show a central tendency, with a dispersion characterized by a standard deviation, σ_m. It is important in this kind of study for the measurement system to be calibrated so that results are expressed in units of the measurand.

It should be emphasized that σ_m characterizes measurement variability and is only one component of measurement uncertainty. It is possible for a measurement process to be highly repeatable and yet have a large uncertainty. A perfectly repeatable length measurement, for example, might be performed in an environment where the temperature is stable and uniform, but poorly known. In such a case, the measurement uncertainty could be dominated by this poor knowledge of the workpiece and instrument temperatures and their coefficients of thermal expansion.

Suppose that the sample of spacers in Fig. 2 were measured with a noisy measurement system with a variability characterized by a standard deviation σ_m, with each measurement consisting of a single reading. Then, under very general conditions, the total standard deviation, σ_T, of the frequency distribution of measurement results would be σ_T (single measurement) $= \sqrt{\sigma_p^2 + \sigma_m^2}$.

If the spacers were each measured n times, with the result taken to be the average of the n measurements, then for the frequency distribution of averages, $\sigma_T(n) = \sqrt{\sigma_p^2 + \sigma_m^2/n}$. These results show (1) how the effect of measurement nonrepeatability can be reduced by averaging and (2) how process variation can be distinguished from the variability of the measurement system. The latter point follows from the expressions above. From a histogram of single measurements, one would calculate a total standard deviation, σ_T, from which the

Table 1 Fraction Conforming Versus Process Capability Index

Process Capability Index, C_p	Fraction of Spacers Conforming
0.50	0.866
0.55	0.902
0.60	0.928
0.65	0.949
0.70	0.964
0.75	0.976
0.80	0.984
0.85	0.989
0.90	0.993
0.95	0.996
1.00	0.997
1.05	0.998
1.10	0.9990
1.15	0.9994
1.20	0.9997
1.25	0.9998

process standard deviation follows from $\sigma_p = \sqrt{\sigma_T^2 - \sigma_m^2}$. An analogous result is obtained if each measurement is repeated n times, with σ_m replaced by $\sigma_m/\sqrt{n}$.

5 PROBABILITY DENSITIES: PRIOR INFORMATION AND STANDARD UNCERTAINTY

5.1 Conditional Probabilities

The nature of manufacturing and measurement is such that the value of a quantity of interest, such as the length of a workpiece or the magnitude of a measurement error, cannot be known exactly. In general, there will be an infinite number of possible values that are consistent (in the sense of being plausible) with one's knowledge of the manufacturing and/or measurement processes.

In this common situation, one's confidence in the various possible values of an uncertain quantity is represented by a continuous probability density. It is assumed that the reader is familiar with the concept of probability as a numerical representation of degree of belief, with certainty represented by a probability equal to one and impossibility represented by a probability equal to zero.

All probabilities are conditional on whatever information is available that is relevant to the situation. Consider the following statement: "The length of the spacer is 25.000 ± 0.001 mm, with a 99% level of confidence." This statement might or might not be true, and it could be made in a variety of contexts. For example, it might describe:

(a) a spacer chosen at random from a batch of similar parts produced by a well-characterized manufacturing process

(b) a spacer whose length has just been measured in an inspection operation

(c) a spacer purchased from a vendor based on a published specification

The source and nature of the background information is quite different in these three situations.

In this Report, the symbol I will be used to represent conditioning information and probabilities will be written in a way that explicitly displays their conditional nature. Thus, for some assertion A, the quantity $p(A|I)$ is the probability that A is true, given information I. In such probability expressions, quantities to the right of the vertical bar are assumed to be true. For a quantity y that can assume a continuous range of values, the expression $p(y|I)\Delta y$ will stand for the probability that y lies in the range $[y, y+\Delta y]$, given information I.

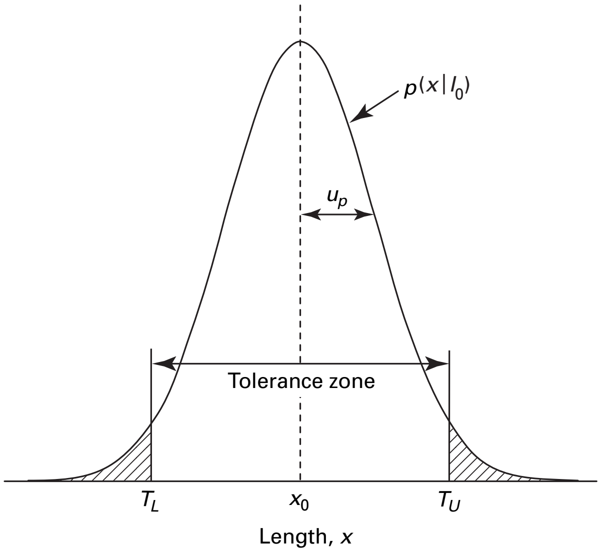
5.2 Probability Density of the Production Process: Prior Information

Consider again the production process described in para. 4.2 and suppose that a spacer is chosen at random during a production run. What can be said about the length of this particular part? Given the information provided by the sample data (Fig. 2), it would seem reasonable to believe that the length of the spacer would be more likely to be near the average length x_0 than to be much larger or smaller than average. It also seems reasonable that the range of plausible lengths could be characterized by the standard deviation, σ_p, of the frequency distribution (see Fig. 2). In the absence of any measurement data, the best that can be done is to estimate or infer the length of the spacer based on information about the process provided by its production history.

Such an inference takes the form of a probability density function (pdf), $p(x|I)$, called the process probability density or, in short, the process density. In the language of probability theory, this density is often called the prior density for the probable lengths of the spacer, since it characterizes a state of knowledge or degree of reasonable belief in the length of the spacer before it is measured.

The prior information, I, that conditions this premeasurement knowledge of the length of a workpiece includes the frequency data of Fig. 2. It might also include other data, such as sample measurements performed as part of a statistical quality control program. Such information is valuable in assuring that the process is free of drift or abnormal variability.

The form of the process density, $p(x|I)$, follows from the information provided by the measurement data of Fig. 2. It seems intuitively reasonable, and in fact can be shown, that the probability $p(x|I)\Delta x$ that the spacer



GENERAL NOTE: This density characterizes what is known about the workpiece length before it is measured. The area under the curve between the tolerance limits is the probability that the workpiece conforms to specification.

Fig. 4 Process Probability Density for the Length of a Randomly Chosen Workpiece

length, X, lies in the interval between x and $x+\Delta x$ is numerically equal to the fraction of the sample of measured spacers in the same interval. This fraction is given by Eq. (2), and so the process density is

$$p(x|I) = \frac{1}{u_p\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - x_0}{u_p}\right)^2\right]$$

with $u_p = \sigma_p$. Figure 4 shows this probability density. The most probable length is just $x_0 = \mu_p$, the process average. The standard deviation, u_p, is a measure of the range of values about x_0 where the probability is appreciable. In the ISO Guide to the Expression of Uncertainty in Measurement (GUM) [7], x_0 is called an estimate of X and u_p is the associated standard uncertainty. The two numbers, x_0 and u_p, together characterize the credible values of the length of an unmeasured workpiece.

5.3 The Difference Between σ_p and u_p

While the quantities σ_p and u_p are both standard deviations and have the same numerical value, they are conceptually different in nature.

The process standard deviation, σ_p, is calculated, to within a relative uncertainty that decreases with increasing sample size, from a sample of measured lengths. This experimentally estimated quantity characterizes the dispersion of the sample of measurements and, as such, is a collective property of the measured sample and the production process.

The standard uncertainty, u_p, by contrast, is an assigned quantity characterizing the dispersion of values that could reasonably be attributed to a particular unmeasured spacer, based on knowledge of the production process acquired via the sample measurements. Thus, u_p characterizes a degree of belief and is not something that could itself be measured. The probability density, $p(x|I)$, and its standard deviation, u_p, are not physical properties of the spacer, but rather they characterize what is reasonable to believe about its length, based on what is known about the production process.

5.4 Conformance of an Unmeasured Workpiece

For the purposes of this Report, it will be convenient to describe the two possible quality states of a workpiece by the symbols C for conformance and $\overline{C}$ for nonconformance. For a spacer of unknown length X, the symbols stand for the following propositions:

(a) C = the spacer conforms to specification, i.e., $T_L \leq X \leq T_U$

(b) $\overline{C}$ = the spacer does not conform to specification, i.e., $X < T_L$ or $X > T_U$

The probability that a spacer conforms (i.e., the probability that C is true) is then written as $p(C|I)$ and the probability that it does not conform is $p(\overline{C}|I)$. Since a spacer either conforms or does not, these probabilities must add up to one: $p(C|I) + p(\overline{C}|I) = 1$.

For a spacer chosen at random but not measured, the conformance probability, $p(C|I)$, is equal to the fraction of the area under the probability density, $p(x|I)$, that lies between the tolerance limits. This fraction, shown as the unshaded portion of the area in Fig. 4, is given by

$$p(C|I) = \int_{T_L}^{T_U} p(x|I)\, dx$$

Numerically, this probability is the same as the fraction of conforming spacers in a large sample as shown in Table 1 and Fig. 3. Thus, unmeasured parts can be accepted for use with acceptable risk so long as the process is controlled and C_p is large enough.

The acceptance of unmeasured workpieces based on knowledge of the process is very common in modern manufacturing. It might seem somewhat unusual to claim a level of confidence in accepting a part that has never been measured — it is a pure inference. But it is conceptually the same as accepting a part based on the result of a measurement. The uncertainty will be smaller in the latter case because of the additional information provided by the measurement, but the true length remains unknown.

6 WORKPIECE INSPECTION: MEASUREMENTS AND MEASUREMENT UNCERTAINTY

6.1 Measurement Probability Density

As part of a quality control system, a spacer is measured in order to decide its conformance to specification. Once corrections have been made for all known significant systematic errors, the result of the measurement is a number, x_m, which is a best estimate of the value of the length, and an associated standard uncertainty, u_m. Even with a high-accuracy measurement, the length cannot be known exactly. Possible values are then represented by a probability density function.

Let I_m stand for the information available after performing the measurement. Symbolically, $I_m = DI$, the prior information, I (what's known before the measurement), updated to include the data, D, acquired in the measurement process. D includes the estimates and associated standard uncertainties of all input quantities that contribute to the evaluation of the estimate, x_m, and its associated measurement uncertainty. The probability density for the length of the spacer following a measurement is called the measurement probability density or, for short, the measurement density.[3]

The modeling of the measurement process, including the assignment of probability densities to the influence quantities and the evaluation of the measurement uncertainty, form the subject of the GUM. In this Report, we assume that the knowledge of the measurand (in this case, the length of the spacer) following a measurement is well represented by the Gaussian probability density

$$p(x|I_m) = \frac{1}{u_m\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - x_m}{u_m}\right)^2\right] \qquad (4)$$

The expectation (or mean) of this density, as well as the most probable value of x, is the estimate x_m. The standard deviation, u_m, characterizes the range of reasonably probable post-measurement lengths and is another way of writing the combined standard uncertainty, $u_c(x)$, associated with the result of the measurement.

The expanded uncertainty, U, is calculated from u_m by multiplying by a coverage factor, k: $U = ku_m$. Unless otherwise stated, in this Report we will use the term measurement uncertainty to mean the expanded uncertainty U with a coverage factor $k = 2$, which is the most common coverage factor used nationally and internationally. For the familiar Gaussian density [Eq. (4)], the expanded uncertainty corresponds to a level of confidence of about 95%. This means that there is a probability of about 95% that the (true) length of a measured spacer lies in the uncertainty interval $[x_m - U, x_m + U]$.

[3] In probability theory, this density is often called the *posterior density* for the probable lengths of the spacer, since it characterizes knowledge of the length of the spacer after it is measured.

6.2 Measurement Capability Index

The definition of the process capability index, C_p, in para. 4.3 illustrates the natural length scale provided by the tolerance. By analogy with C_p, the measurement capability index, C_m, is defined by

$$C_m \equiv \frac{T}{4u_m} = \frac{T}{2U} \qquad (5)$$

which is just the ratio of the tolerance to the width of the uncertainty interval. Just as C_p serves as a useful index of process quality (large $C_p \rightarrow$ low process variability), C_m characterizes the quality of the measurement system (large $C_m \rightarrow$ low measurement uncertainty).

There is a close connection between the measurement capability index, C_m, and various rules and ratios that have been used to characterize measurement quality. Among these are gauging ratio, gauge maker's rule, test accuracy ratio (TAR), test uncertainty ratio (TUR), and others. Sometimes these are stated as numbers, such as a 10-to-1 rule or a TUR of 4:1. One has to be very careful in interpreting these quantities when they are encountered, because they are often ambiguously or incompletely defined.

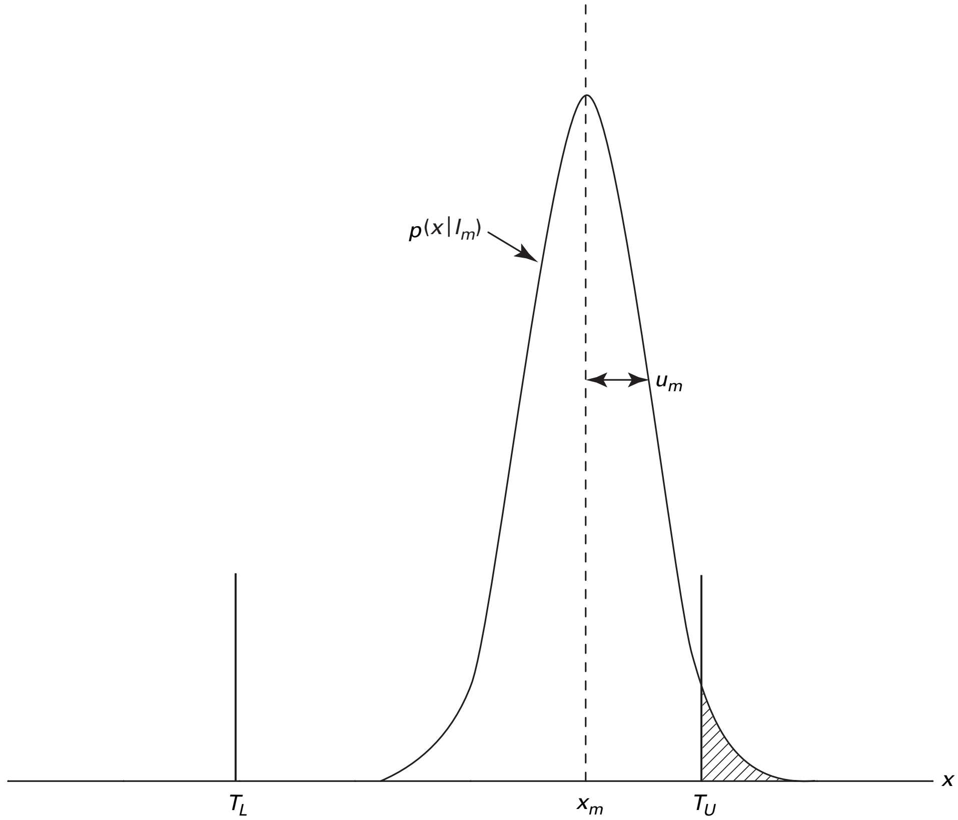
With respect to the TUR, for example, the American Association of Laboratory Accreditation (A2LA) states [10]: "A2LA interprets this ratio to mean that the total uncertainty of the measurement system (as opposed to a simple combination of the uncertainties of the reference standards) does not exceed a given fraction of the specified tolerance." Here the meaning of total uncertainty is ambiguous.

Similarly, the Instrument Society of America (ISA), in a Web-based dictionary [11], defines the test uncertainty ratio (TUR) as "a measure of calibration accuracy — the ratio of observation uncertainty of a unit being calibrated to the output uncertainty of the calibration source." In this case, the terms observation uncertainty and output uncertainty have no clear meanings.

The definition of C_m in Eq. (5) is unambiguous in the case of workpiece inspection with upper and lower tolerance limits. It is consistent with the nomenclature of ASME B89.7.3.1 [3]. In that standard, for example, a 4:1 Decision Rule means that $C_m = 4$.

In the case of a one-sided measurement of a feature such as flatness, there is a lower bound of zero and a single (upper) tolerance limit, T. In this case, the measurement capability index is defined to be $C_m = T/2u_m = T/U$ (one-sided measurement).

In the calibration or verification of measuring instruments, the instrument specification is often in terms of a maximum permissible error (MPE) that should bound the absolute value of instrument errors. In this case, C_m is defined as in Eq. (5) with the replacement of T by $2MPE$, so that $C_m = 2MPE/2U = MPE/U$. The expanded uncertainty, U, of the observed errors will generally have contributions from the imperfect standards used in the



GENERAL NOTE: The best estimate, x_m, lies in the tolerance zone, indicating conformance, but there is a possibility that the spacer is too long. The range of reasonably probable lengths is characterized by the standard uncertainty, u_m. About 95% of the probability lies in the uncertainty interval $[x_m - U, x_m + U]$, where $U = 2u_m$. In this example, the measurement capability index is $C_m = T/2U = 2$.

Fig. 5 Probability Density for the Lengths of a Measured Workpiece

calibration, environmental effects, and from uncertainties in the instrument's readings.

6.3 Conformance of a Measured Workpiece

Figure 5 shows the measurement density, $p(x|I_m)$, of Eq. (4) that characterizes the knowledge of the length of a particular spacer after an inspection measurement. The most probable length is the estimate, x_m. The associated standard uncertainty, u_m, is a measure of the region about x_m where most of the probability is concentrated; 95% of the probable lengths lie in the uncertainty interval $x_m \pm 2u_m$ or $x_m \pm U$. For the example shown in Fig. 5, u_m is one-eighth of the tolerance, so the measurement capability index, C_m, is equal to 2.

Since the measurement result, x_m, lies in the tolerance zone, one might decide to accept the spacer as conforming to specification; this is an example, for $C_m = 2$, of a decision rule called simple 2:1 acceptance (see ASME B89.7.3.1 [3]). Acceptance is not the same as conformance, however; in Fig. 5, for example, there is an obvious fraction (shown hatched) of the probable lengths of the spacer that are outside of the upper tolerance limit, corresponding to a part that is too long.

Given the measurement data, the probability, $p(C|I_m)$, that a measured spacer conforms to its specification equals the fraction of the probable lengths contained between the tolerance limits. Writing $p(C|I_m) = P_C$, the probability of conformance is

$$P_C = \int_{T_L}^{T_U} p(x|I_m)\, dx$$

Inserting the Gaussian measurement density [Eq. (4)] yields explicitly

$$P_C = \frac{1}{u_m\sqrt{2\pi}} \int_{T_L}^{T_U} \exp\left[-\frac{1}{2}\left(\frac{x - x_m}{u_m}\right)^2\right] dx \qquad (6)$$

This integral cannot be evaluated in closed form, but can be expressed (see Mandatory Appendix I) in terms of the well-known standard normal cumulative distribution function (CDF), $\Phi(z)$, defined by

$$\Phi(z) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-t^2/2\right) dt = \int_{-\infty}^{z} f_0(t)\, dt$$

where $f_0(t) = \left(1/\sqrt{2\pi}\right) \exp\left(-t^2/2\right)$ is called the standard normal probability density function.

The cumulative probability, $\Phi(z)$, is a number between 0 and 1, tabulated in most statistics books, and included in commercial spreadsheet and mathematics software. Letting $t = (x - x_m)/u_m$ in Eq. (6) then gives:

$$P_C = \Phi\left(\frac{T_U - x_m}{u_m}\right) - \Phi\left(\frac{T_L - x_m}{u_m}\right)$$

This result expresses the conformance probability, P_C, in terms of the particular product specification limits (T_L, T_U) and the result of a particular measurement (x_m, u_m). Because of the natural length scale provided by the tolerance, T, this result can be rewritten in a form that is suitable for a general inspection problem. Defining a scaled measurement result, $\hat{x}$, by

$$\hat{x} \equiv \frac{x_m - T_L}{T} \tag{7}$$

and using Eq. (5) for the measurement capability index, C_m, the conformance probability, P_C, can be written as follows:

$$\begin{aligned} P_C &= \Phi\left[4C_m\,(1 - \hat{x})\right] - \Phi\,(-4C_m\,\hat{x}) \\ &= P_C\,(\hat{x}, C_m) \end{aligned} \tag{8}$$

The probability, P_C, that a measured spacer conforms to specification thus depends on the two dimensionless numbers $\hat{x}$ and C_m. For a given dimensional measurement plan, the measurement capability index, C_m, is usually a constant. In this case, a question as to whether or not a measured part is in tolerance, given a required level of confidence, may be decided on the basis of the best estimate, x_m, via Eqs. (7) and (8).

Numerical Example. Consider again the example discussed in para. 4.3. Here the tolerance zone for a feature of length is the interval $T_L = 1\,499.8$ mm to $T_U = 1\,500.2$ mm, so that $T = 0.4$ mm. The measurement standard uncertainty is $u_m = 0.04$ mm, so that the measurement capability index is

$$C_m = T/4u_m = 0.4/0.16 = 2.5$$

Suppose that an inspector measures this feature on a particular workpiece, with the resulting estimate $x_m = 1\,500.16$ mm. What is the probability, P_C, that the feature is in tolerance?

From Eq. (7), the scaled measurement result, $\hat{x}$, is

$$\hat{x} = (x_m - T_L)/T = (1\,500.16 - 1\,499.8)/0.4 = 0.9$$

Then, from Eq. (8),

$$\begin{aligned} P_C &= \Phi\left[4 \times 2.5(1 - 0.9)\right] - \Phi\,(-4 \times 2.5 \times 0.9) \\ &= \Phi\,(1) - \Phi\,(-9) \end{aligned}$$

From a table of the normal CDF $\Phi(z)$, we find $\Phi(1) = 0.84$ and $\Phi(-9) \approx 0$. Thus, $P_C = 0.84$ and there is an 84% probability that the feature conforms to specification and a 16% probability that it does not.

For a particular part length measurement, the estimate, x_m, and the associated standard uncertainty, u_m, uniquely determine $\hat{x}$ and C_m, and therefore the conformance probability, P_C, via Eq. (8). There are an infinite number of pairs (x_m, u_m) that yield a given level of confidence, P_C. A useful and informative way of displaying this information is shown in Fig. 6, for a level of confidence $P_C = 95\%$.

In Fig. 6, the vertical axis shows $C_m = T/4u_m$ on a logarithmic scale with values corresponding to various gauging ratios. The horizontal axis shows values of the scaled measurement result, $\hat{x} = (x_m - T_L)/T$, in the range from 0 to 1, corresponding to values of x_m between T_L and T_U, i.e., measurement results within the tolerance zone. The restriction of x_m to this range is a practical one. For a measurement result, x_m, outside of the tolerance limits, the probability of conformance is less than 50% no matter what the measurement uncertainty. It is unlikely that such a workpiece would be found acceptable.[4]

The solid curve in Fig. 6 is a line of constant 95% probability that divides the measurement results into regions of conformance and nonconformance at a 95% level of confidence. A spacer for which the result (x_m, u_m) yields a point in the shaded region below the curve has a conformance probability, P_C, of less than 95%.

7 GAUGING (OR TEST) LIMITS AND GUARD BANDS

7.1 Defining an Acceptance Zone Using Gauging (or Test) Limits

If their lengths could be measured exactly, a batch of spacers could be sorted good from bad without risk of error. Because of the uncertainty of any real measurement process, however, the situation is not so simple. A part whose measured length lay within the tolerance zone might in fact be too long or too short. Similarly, a part measuring too long or too short might well be conforming.

Consider, for example, a spacer whose measured length, x_m, lay right at one of the tolerance limits. Such a spacer would be equally probable of conforming or not conforming to specification. Whether such a part

[4] In the case of inspection during production, such workpieces might be accepted, provided they were sufficiently rare.

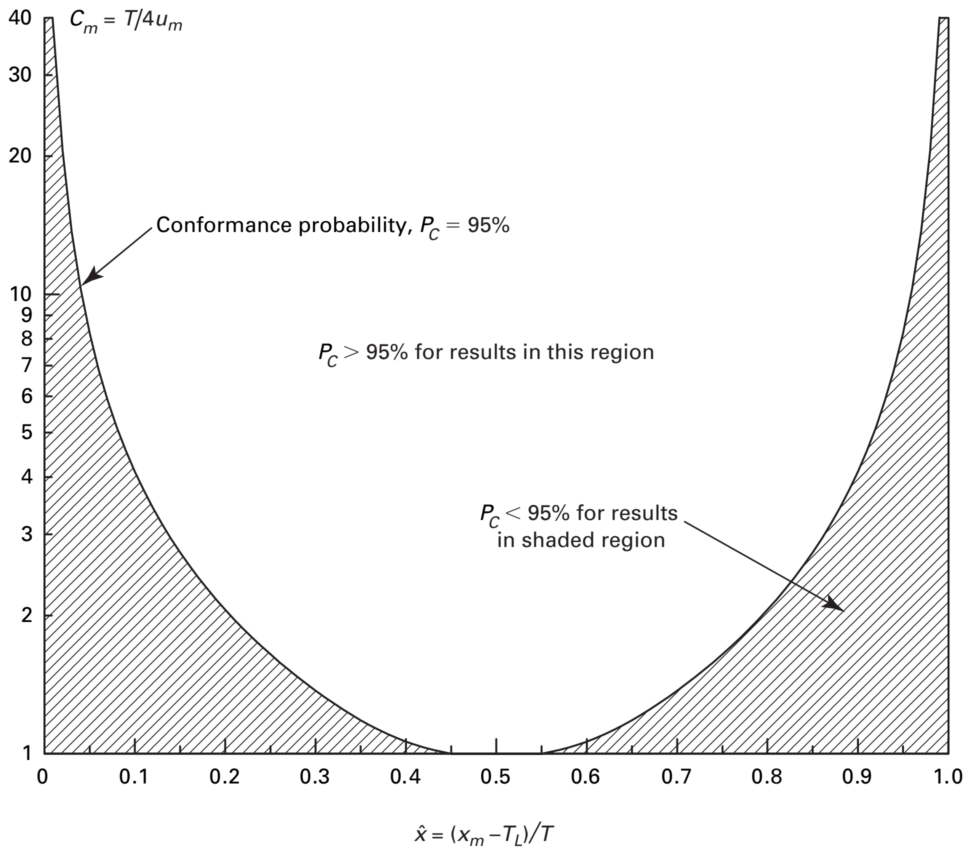


Fig. 6 Measurement Capability Index Versus Scaled Measurement Result

was accepted or rejected, there would be a 50% chance of making a mistake.

The risk of accepting nonconforming workpieces can be reduced by setting a pair (G_L, G_U) of upper and lower gauging limits (also called test limits) inside the tolerance limits. Such gauging limits define a reduced acceptance zone, as shown in Fig. 7. In a typical dimensional measurement plan, a workpiece is accepted (passes inspection) if its measured length lies in the acceptance zone and is rejected otherwise. This is a binary decision rule, where there are only two possible outcomes of a conformance test measurement.[5]

For gauging limits inside the tolerance zone, as in Fig. 7, the resultant acceptance zone is called a stringent acceptance zone. With a binary decision rule, stringent acceptance is accompanied by relaxed rejection, so called because a workpiece can be rejected even though its measured length lies in the tolerance zone (i.e., in one of the regions between the gauging limits and the tolerance limits). In this situation, business economics favor a larger risk of rejecting a good part in order to decrease the probability of accepting a bad one.

If the gauging limits are placed outside of the tolerance zone, the resulting relaxed acceptance zone is accompanied by a stringent rejection zone, as shown in Fig. 8. In this situation, business economics would favor a larger risk of accepting a bad workpiece in order to decrease the probability of rejecting a good one.

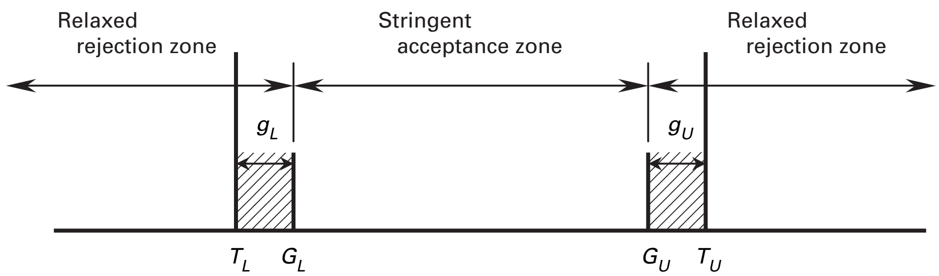
7.2 Guard Bands

The magnitudes of the offsets between the tolerance limits and the gauging limits are called guard bands [12–18]. The function of these offsets, depending on their placement, is to guard against accepting bad workpieces or rejecting good ones.

Figures 7 and 8 show lower (g_L) and upper (g_U) guard bands for the cases of stringent acceptance and stringent rejection, respectively. Depending upon the costs associated with faulty accept/reject decisions, the lower and upper guard bands might have different magnitudes.[6] In the case of a quantity such as roundness error, which

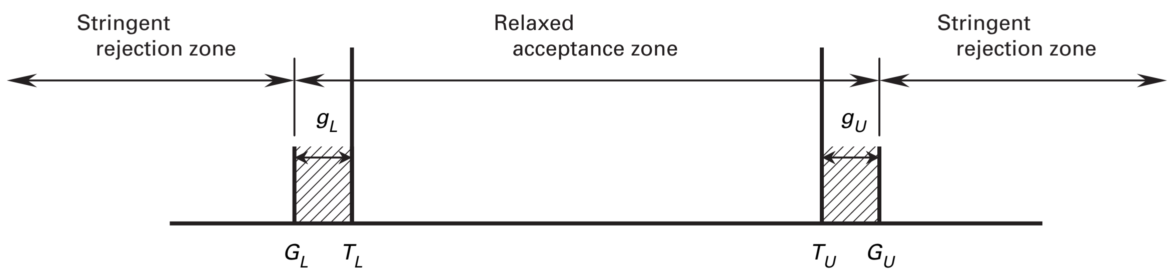
[5] In this and the following paragraphs, the nomenclature follows the terminology of ASME B89.7.3.1, Guidelines for Decision Rules.

[6] In the production of one-dimensional spacers, e.g., workpieces that are too long could be reworked in a downstream operation, while ones that are too short could not be made to function and would have to be scrapped. The decision rule then might favor a higher risk of accepting a nonconforming long spacer and a lower risk of accepting a nonconforming short one.



GENERAL NOTE: The offsets between the tolerance limits and the gauging limits are the guard bands g_L and g_U. A stringent acceptance decision rule reduces the probability of accepting a nonconforming workpiece.

Fig. 7 Stringent Acceptance Zone



GENERAL NOTE: The offsets between the tolerance limits and the gauging limits are the guard bands g_L and g_U. A stringent rejection decision rule reduces the probability of rejecting a conforming workpiece.

Fig. 8 Relaxed Acceptance Zone

is always positive, there would typically be a single tolerance limit and only one guard band.

This Report considers, in detail, symmetric two-sided guard banding where the guard bands are the same size, $g_L = g_U = g$, and are expressed in units of the expanded uncertainty, U

$$g = hU$$

where

$h > 0$ for stringent acceptance
$h < 0$ for relaxed acceptance

The quantity, h, is called a guard band multiplier and its numerical value is used in specifying an unambiguous decision rule. As a particular example, taking $h = +1$ (i.e., $g = +U$) results in a decision rule called stringent acceptance with a 100% guard band, using the nomenclature of ASME B89.7.3.1.

8 CONTROLLING THE QUALITY OF INDIVIDUAL WORKPIECES

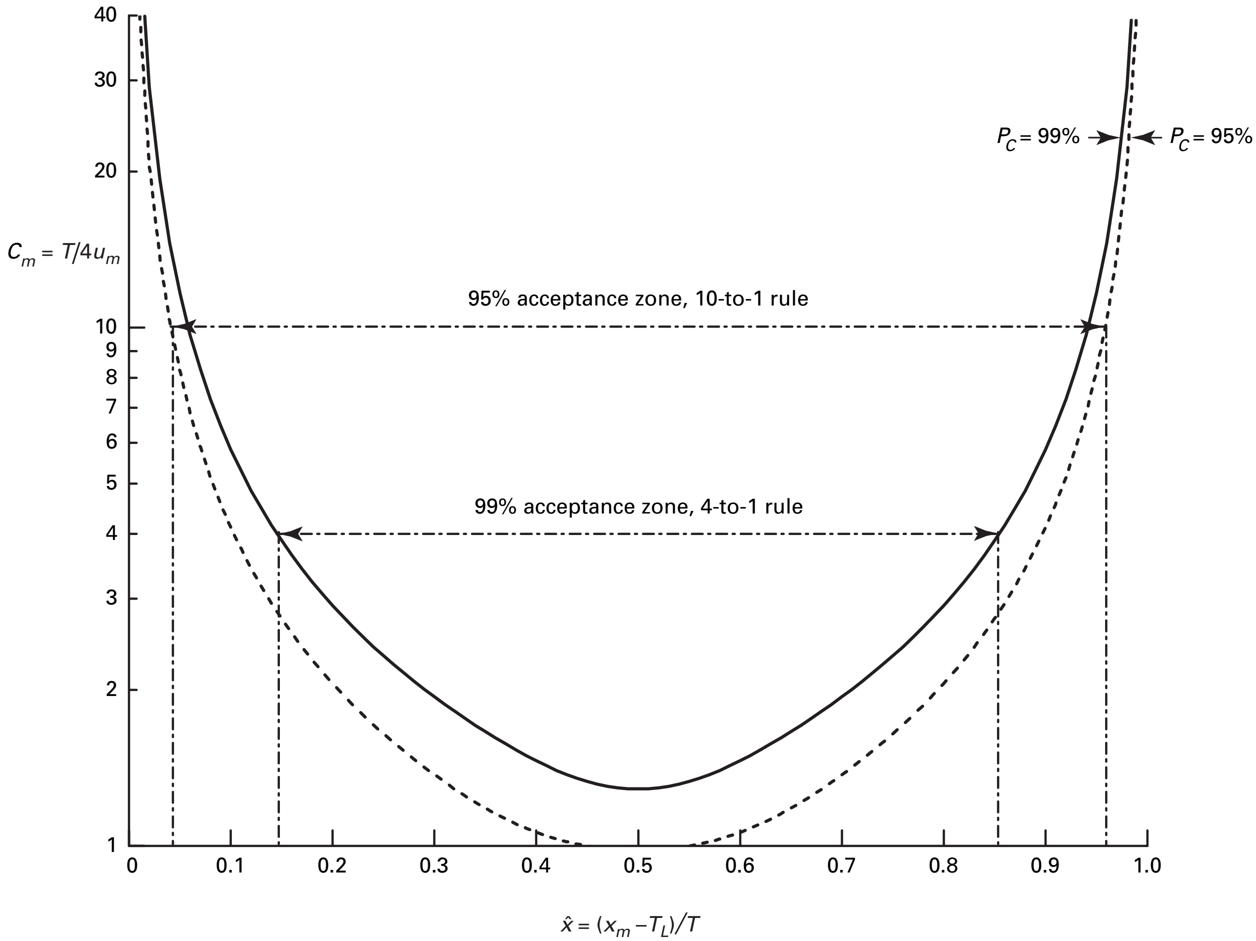
8.1 Acceptance Zones and Levels of Confidence

Consider the measurement of spacers, and suppose that economic considerations require that every spacer measured and accepted for use must have at least a probability P_C of conforming to specification. The size of the appropriate acceptance zone can be understood by reference to Fig. 9, which shows curves of constant 95% and 99% conformance probability.

For a given measurement capability index, $C_m = T/4u_m$, and level of confidence, P_C, the associated acceptance zone, as a fraction of the tolerance, is the width of the curve of constant P_C where it intersects the line of constant C_m. For a given level of confidence, such as $P_C = 99\%$, the acceptance zone shrinks in size with decreasing C_m (increasing measurement uncertainty) and ultimately reduces to zero. In Fig. 9, for example, we see that for C_m less than about 1.4, no measured spacers could be accepted at a 99% level of confidence. At such low measurement capabilities, more than 1% of the probability would lie outside the tolerance zone regardless of the result of the measurement.

8.2 Setting Guard Band Limits for Individual Workpieces

Once the required level of confidence (conformance probability) is chosen, setting the guard band limits is straightforward. Figure 10 shows the measurement probability density for a spacer whose measured length lies exactly at the upper gauging (or test) limit, G_U.



GENERAL NOTE: For a given level of confidence, the width of the acceptance zone increases with better measurement quality (i.e., larger values of measurement capability index $C_m = T/4u_m$).

Fig. 9 Desired Level of Confidence Defines an Acceptance Zone

For stringent acceptance at a level of confidence P_C, the upper gauging limit, G_U, is set inside the upper tolerance limit, creating an upper guard band of magnitude $g = hU$, $h > 0$. In two-sided symmetric guard banding, an equal offset inside the lower tolerance limit fixes the location of the lower gauging limit, G_L. By only accepting spacers whose measured lengths lie in the acceptance zone of width $T - 2g$, those that pass inspection will conform to specification with a probability of at least P_C. Figure 11 shows this stringent acceptance zone.

The measurement probability density shown in Fig. 10, with a measurement result at the upper gauging limit, is given by Eq. (4) with $x_m = G_U$

$$p(x|x_m = G_U, I_m) = \frac{1}{u_m\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - G_u}{u_m}\right)\right]$$

The conformance probability follows from Eq. (8), with $\hat{x} = (G_U - T_L)/T = 1 - g/T$ or $\hat{x} = 1 - h/2C_m$

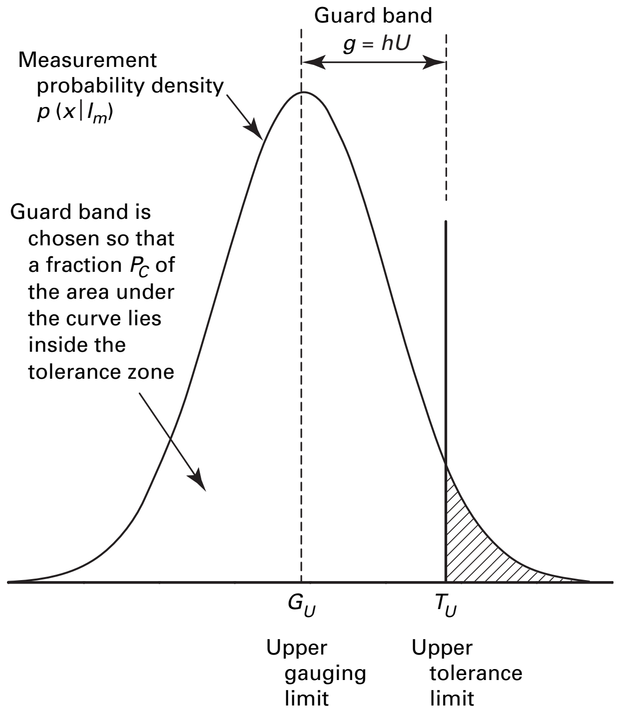
$$\begin{aligned} P_C &= \Phi\left[4C_m\,(1 - \hat{x})\right] - \Phi\,(-4C_m\,\hat{x}) \\ &= \Phi(2h) - \Phi(2h - 4C_m) \end{aligned} \tag{9}$$

The second term in Eq. (9) represents the leakage of a small fraction of the measurement density into the region below the lower tolerance limit. In situations of practical interest, this probability will be very close to zero. Taking $h = 1$, for example, so that $g = U$, and assuming a rather poor measurement capability index, C_m, equal to 2, then $\Phi(2h - 4C_m) = \Phi(-6) \approx 10^{-9}$. Usually C_m is 4 or larger, so that the second term can be safely neglected and we have

$$P_C = \Phi(2h)$$

Then the multiple of U in setting the guard band $g = hU$ is given by

$$h = \frac{1}{2}\,\Phi^{-1}\,(P_C)$$



GENERAL NOTE: The curve shows the measurement density for an estimate x_m at the upper gauging (or test) limit. The guard band magnitude $g = hU$ is chosen so that a fraction P_C of the probability lies inside the tolerance zone.

Fig. 10 Guard Band Chosen to Reduce the Probability of Accepting a Workpiece That Is Too Long

where

Φ^{-1} = inverse of the normal cumulative distribution function

Table 2 gives values for the guard band multiplier, h, for several levels of confidence. To assure a conformance probability of 90%, for example, the gauging (or test) limits should be offset from the tolerance limits by $g = 0.64U$. The corresponding decision rule would be stated as 64% stringent acceptance.

The following examples illustrate choosing guard band limits when conformance probability must be controlled for every measurement.

8.2.1 Example 1. Consider again the process described in para. 4.3. The upper and lower tolerance limits are $T_U = 1\,500.2$ mm and $T_L = 1\,499.8$ mm for a feature of nominal length $x_0 = 1\,500$ mm. The measurement standard uncertainty is $u_m = 0.04$ mm, so that the measurement capability index is $C_m = T/4u_m = 2.5$.

In order for a workpiece to be acceptable, the feature must conform to specification with a level of confidence, P_C, of at least 99%. Where should the guard bands be placed?

Table 2 Conformance Probability Versus Guard Band Multiplier

Conformance Probability, P_C	Guard Band Multiplier, h
0.80	0.42
0.85	0.52
0.90	0.64
0.95	0.82
0.99	1.16
0.999	1.55

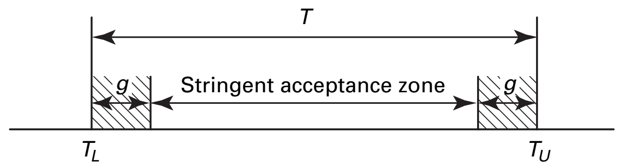


Fig. 11 Stringent Acceptance Zone for Symmetric Two-Sided Guard Banding

Solution. From Table 2, with $P_C = 0.99$, we see that $h = 1.16$. Thus, the guard band limits should be set inside the tolerance limits by 116% of the expanded uncertainty. Then the upper gauging limit is

$$\begin{aligned} G_U &= T_U - hU \\ &= 1\,500.2\text{ mm} - 1.16 \times 0.08\text{ mm} \\ &\approx 1\,500.1\text{ mm} \end{aligned}$$

and the lower gauging limit is

$$\begin{aligned} G_L &= T_L + hU \\ &= 1\,499.8\text{ mm} + 1.16 \times 0.08\text{ mm} \\ &\approx 1\,499.9\text{ mm} \end{aligned}$$

Note that for this relatively poor measurement capability ($C_m = 2.5$) and large required conformance probability ($P_C = 99\%$), the acceptance zone is only one-half the width of the tolerance zone.

8.2.2 Example 2. This example comes from electrical metrology and involves the testing of a measuring instrument for conformance to a maximum permissible error, *MPE,* requirement.

A digital voltmeter is to be tested by applying a 1 V dc input from a precision voltage reference source. For this input, the voltmeter specification states that $MPE = \pm 10.4\ \mu$V. The $k = 2$ expanded uncertainty, U, of the 1 V dc reference input is 4.2 μV. Where should the guard bands be set so that a voltmeter that passes inspection has a probability, P_C, of at least 95% of conforming to specification?

Solution. Here the measurand is the voltmeter error and the tolerance zone is centered at zero with a width equal to 2*MPE.* The measurement capability index is then given by $C_m = 2MPE/2U = MPE/U = 10.4/4.2 = 2.5$. The required level of confidence for an instrument passing inspection is $P_C = 95\%$.

From Table 2, for $P_C = 0.95$, we see that $h = 0.82$. The guard band limits should thus be placed inside the maximum permissible error limits by $0.82U = 3.45$ μV, so that the test (or gauging) limits that define the acceptance zone are set at $\pm(10.4 - 3.45) = \pm6.95$ μV.

9 CONTROLLING THE AVERAGE QUALITY OF WORKPIECES

9.1 Average Versus Individual Level of Confidence

In para. 8, guard banding was used to assure a minimum level of confidence for each *individual* workpiece. In situations where large numbers of parts are produced, it can be economically advantageous to use less restrictive guard banding, with gauging (or test) limits chosen to assure an acceptable *average* level of confidence when workpieces are inspected.

In such a case, it might be acceptable as a business decision for an occasional part that passes inspection to have a higher probability of not conforming to specification than the average accepted part. With this type of guard banding, more parts will pass inspection and fewer will be rejected, so long as the average level of confidence is acceptable.

Unlike the procedure in para. 8, setting the guard band limits in this type of inspection relies on prior knowledge of the process density. Consider a manufacturer who requires a typical workpiece to conform to specification at a 95% level of confidence or greater. The manufacturer can achieve this with a process capability index, $C_P = 0.65$ or greater and no measurement at all, except for an occasional measurement to verify that the process is stable and that $C_P \geq 0.65$.

Now, if the manufacturer decides, for economic reasons, that a typical workpiece must conform at a 99% level of confidence, then a measurement system with appropriate gauging limits can be used to ensure this outcome. Workpiece characteristics with values that are far from the process average (and thus nonconforming) will be more likely to fail inspection than those near the process average. The average conformance probability of accepted workpieces will thus rise, and an acceptance zone can be calculated that will yield an average level of confidence of 99%.

The following paragraphs describe these calculations.

9.2 Consumer's Risk and Producer's Risk

There are four possible outcomes of an inspection measurement with a binary decision rule: a workpiece could be conforming (C) or nonconforming ($\overline{C}$), and it could pass (P) or fail (F) inspection.

The events P and F are introduced by the following definitions:

For a spacer that passes inspection, let

P = the measured length x_m lies in the acceptance zone
$= G_L \leq x_m \leq G_U$

For a spacer that fails inspection, let

F = the measured length x_m does not lie in the acceptance zone
$= x_m < G_L$ or $x_m > G_U$

Combining in pairs each possible quality state (C, $\overline{C}$) with each possible result (P, F) of the length measurement yields the following four possible outcomes of an inspection measurement:

(a) PC (the spacer passes inspection and conforms to specification). This is a desired outcome of an inspection measurement, leading to acceptance of a good part.

(b) $P\overline{C}$ (the spacer passes inspection and does not conform to specification). This is a mistake, variously called a pass error, a Type II error, a false accept, or a false positive. The probability of a pass error, $p(P\overline{C}|I_0) \equiv R_C$, is often called the consumer's risk, since the cost associated with an out-of-tolerance part is usually borne by the customer.

(c) FC (the spacer fails inspection and conforms to specification). This is another mistake, variously called a fail error, a Type I error, a false reject, or a false negative. The probability of a fail error, $p(FC|I_0) \equiv R_P$, is often called the producer's risk, since the cost of rejecting a conforming part is usually borne by the manufacturer.

(d) $F\overline{C}$ (the spacer fails inspection and does not conform to specification). This is a desired outcome leading to rejection of a bad part.

Figure 12 shows a contingency table containing the probabilities of the four possible outcomes of a spacer conformance test. At the bottom are the marginal probabilities of conformance and nonconformance, which depend only on the process distribution. The right-hand column shows the marginal probabilities of passing or failing inspection.

9.3 Consumer's and Producer's Risk Calculations

Evaluation of the consumer's and producer's risks requires numerical integration, an exercise that may be performed manually or with the aid of a computer program. A particular example of the manual approach is given in ASME B89.7.2. This paragraph presents a generalized approach that yields equivalent results. The mathematical details are given in Mandatory Appendix II of this Report.

It should be noted that in the following procedures, the risks are calculated, given a known set of gauging limits. In most real applications, a desired level of risk is chosen and one needs to choose gauging limits that will ensure that the risk target is met. Such a calculation is not straightforward. A practical way to determine

	Spacer conforms, C	Spacer does not conform, $\overline{C}$	
Spacer passes, P	$p(PC\|I_0)$	$p(P\overline{C}\|I_0)$	Probability that spacer passes, $p(P\|I_0)$
Spacer fails, F	$p(FC\|I_0)$	$p(F\overline{C}\|I_0)$	Probability that spacer fails, $p(F\|I_0)$
	Probability that spacer conforms, $p(C\|I_0)$	Probability that spacer does not conform, $p(\overline{C}\|I_0)$	

GENERAL NOTE: The table entries are the probabilities of the various outcomes. The quantity $p(P\overline{C}|I_0)$ is the probability of a pass error, which means accepting a nonconforming spacer. This probability is often called the consumer's risk, written R_C. Similarly, the quantity $p(FC|I_0)$ is the probability of a fail error, which means rejecting a conforming spacer. This probability is often called the producer's risk, written R_P.

Fig. 12 Contingency Table for an Inspection Measurement

gauging limits for a desired level of risk is via graphs such as those in Figs. 16 through 19, as described below.

Calculation of the consumer's risk, R_C, and producer's risk, R_P, requires knowledge of the following quantities:

(a) the process density, assumed to be a Gaussian (or normal) probability density, characterized by the estimate (expectation) x_0 and associated standard uncertainty $u_p = \sigma_p$, where σ_p is an estimated standard deviation that characterizes the process variability. The process is centered, meaning $x_0 = (T_L + T_U)/2$.

(b) the measurement density, also assumed to be a Gaussian or normal probability density, with estimate x_m and associated standard uncertainty u_m.

(c) the upper and lower tolerance limits, T_U and T_L.

(d) the upper and lower gauging (or test) limits, G_U and G_L.

9.3.1 Procedure. Once the above quantities are known, the procedure is as follows:

(a) Compute the tolerance, $T = T_U - T_L$.

(b) Compute the guard band multiplier, $h = (T_U - G_U)/2u_m$.

(c) Compute the inherent process capability index, $C_p = T/6u_p$.

(d) Compute the measurement capability index, $C_m = T/4u_m$.

(e) Compute $r = u_p/u_m = \frac{2}{3}(C_m/C_p)$.

(f) Compute $\gamma = 2(C_m - h)$.

NOTE: γ is the width of the acceptance zone in units of expanded uncertainty U.

(g) Form the function $F(z) = \Phi(\gamma - rz) - \Phi(-\gamma - rz)$. Here, Φ is the standard normal cumulative distribution function (see Mandatory Appendix I).

(h) Compute the probability of a fail error that is the producer's risk, R_P

$$R_P = p(FC|I_0) = \int_{-3C_p}^{3C_p} \left[1 - F(z)\right] f_0(z)\, dz \tag{10}$$

where

$$f_0(z) = \left(1/\sqrt{2\pi}\right)\exp\left(-z^2/2\right)$$

is the standard normal probability density function (see Mandatory Appendix I).

(i) Compute the probability of a pass error that is the consumer's risk, R_C

$$R_C = p(P\overline{C}|I_0) = \int_{-\infty}^{-3C_p} F(z)\, f_0(z)\, dz + \int_{3C_p}^{\infty} F(z)\, f_0(z)\, dz$$

9.3.2 Numerical Example. Consider again the example from para. 4.3. Assume that the data refer to a steel spacer of nominal length $x_0 = 1\,500$ mm and that such parts are measured in an environment where the mean temperature is 25°C. The coefficient of thermal expansion of the workpiece material is 12×10^{-6}/°C. The measurement plan specifies that spacer length measurements are to be corrected for nominal thermal expansion (a systematic error), which in this case amounts to $\Delta L = \alpha L \Delta T = (12\times10^{-6}/°\text{C})(1\,500\text{ mm})(5°\text{C}) = 0.09$ mm. Evaluation of the combined standard uncertainty of the measurement process according to the GUM should include a term that accounts for the uncertainty of this correction.

(a) The following data apply to the production and inspection processes:

(1) The process density is a Gaussian with mean value $x_0 = 1\,500$ mm and standard uncertainty $u_p = \sigma_p = 0.12$ mm.

(2) The measurement density is a Gaussian with standard uncertainty $u_m = 0.04$ mm.

(3) The upper and lower tolerance limits are $T_U = 1\,500.2$ mm, $T_L = 1\,499.8$ mm.

(4) The upper and lower gauging limits are $G_U = 1\,500.18$ mm, $G_L = 1\,499.82$ mm.

(b) With this information, the steps leading to the associated risks are as follows:

(1) Tolerance, $T = (1\,500.2 - 1\,499.8)$ mm $= 0.4$ mm

(2) $h = (1\,500.2 - 1\,500.18)/(2 \times 0.04) = 0.25$

(3) $C_p = T/6u_p = 0.4/(6 \times 0.12) = 0.55$

(4) $C_m = T/4u_m = 0.4/(4 \times 0.04) = 2.5$

(5) $r = 2C_m/3C_p = 3.025$

(6) $\gamma = 2(2.5 - 0.25) = 4.5$

(7) $F(z) = \Phi(4.5 - 3.025z) - \Phi(-4.5 - 3.025z)$

(8) $$p(FC|I_0) = \int_{-1.653}^{1.653} \left[1 - F(z)\right] f_0(z)\, dz$$

Carrying out the numerical integration yields the producer's risk, R_P

$$R_P = p(FC|I_0)\; 0.0694 = 6.94\%$$

(9) $$p(P\overline{C}|I_0) = \int_{-\infty}^{-1.653} F(z) f_0(z)\, dz + \int_{1.653}^{\infty} F(z) f_0(z)\, dz$$

Carrying out the numerical integration yields the consumer's risk, R_C

$$R_C = p(P\overline{C}|I_0) = 0.0101 \approx 1\%$$

In para. 4.3, it was shown that simply accepting all parts produced by this process, with no inspection, would result in a 9.8% defect rate or consumer's risk, meaning nearly one out of every ten spacers produced would be out of tolerance. This example shows how the risk is reduced by the conformance test procedure and associated decision rule,[7] with a post-measurement consumer's risk of about 1%, at the cost of rejecting about 7% of conforming spacers. Whether or not this is an acceptable situation is a business decision that depends on the costs associated with accept/reject errors.

Another way to reduce the risks would be to improve the process by reducing its variability. If the process standard deviation were reduced from 0.12 mm to 0.08 mm (a reduction of about 35%), then all spacers could be shipped with a fraction nonconforming of about 1% and no retention of costly scrap. The manufacturer would have to compare the cost of this process improvement with the costs of inspection and the subsequent generation of scrap spacers.

Once the consumer's and producer's risks have been calculated, the other probabilities in the contingency table can be easily found. Since $p(PC|I_0) + p(FC|I_0) = p(C|I_0)$, and since $p(C|I_0) = 0.902$ or 90.2% (see para. 4.4), the probability that a spacer conforms and passes inspection is just

$$p\,(PC|I_0) = p\,(C|I_0) - p\,(FC|I_0) = 90.2\% - 6.9\% = 83.3\%$$

Similarly, the probability that a spacer does not conform and fails inspection is

$$p\,(F\overline{C}|I_0) = p(\overline{C}|I_0) - p\,(P\overline{C}|I_0) = 9.8\% - 1.0\% = 8.8\%$$

Figure 13 shows the completed contingency table for this example.

[7] In this example, with $h = 0.25$, the decision rule according to ASME B89.7.3.1 would be called 25% stringent acceptance.

	Spacer conforms, C	Spacer does not conform, $\overline{C}$	
Spacer passes, P	83.3%	Consumer's risk 1.0%	Probability that spacer passes, 84.3%
Spacer fails, F	Producer's risk 6.9%	8.8%	Probability that spacer fails, 15.7%
	Probability that spacer conforms, 90.2%	Probability that spacer does not conform, 9.8%	

GENERAL NOTE: The probabilities for the four possible outcomes sum to 100%, as do the marginal probabilities for pass/fail and conform/nonconform.

Fig. 13 Contingency Table for the Worked Example

(c) The following features of this example conformance test procedure can be noted:

(1) The manufacturing process continues to produce 90.2% conforming and 9.8% nonconforming spacers.

(2) The inspection measurements serve to detect and remove 8.8% out of the 9.8% bad parts, the remaining 1% being falsely accepted as conforming.

(3) 84.3% of the manufactured spacers pass inspection; of these, 83.3/84.3 ≈ 99% conform to specification, while about 1% are out of tolerance.

(4) Of the 15.7% of spacers that fail inspection, 6.9/15.7 ≈ 44% are in tolerance. This is one of the prices to be paid for passing only 1% bad product.

Figure 14 graphically displays the producer's risk and consumer's risk versus the measurement capability index, $C_m = T/4u_m$, for an inherent process capability index $C_p = 0.55$, the value used in the worked numerical example. The various curves correspond to different choices of guard band $g = T_U - G_U$; the heavy solid curve corresponds to the value used in the worked example: $g = +0.02$ mm $= +0.25U$. Positive values of g indicate guard bands located inside the tolerance limits (i.e., stringent acceptance).

A study of Fig. 14 shows that acting to reduce the acceptance of nonconforming spacers by increasing the guard band (reducing the consumer's risk) always results in an increased number of conforming spacers that are falsely rejected (increased producer's risk). This inverse relationship between the producer's and consumer's risks is well-illustrated in Fig. 15, which shows R_P versus R_C for this example.

Figures 16–19 show graphs of R_P versus R_C for C_p = 1.5, 1, ⅔, and ⅓, respectively, for values of measurement capability index in a range from $C_m = 2$ to $C_m = 10$ and guard bands in a range from $g = -U$ (100% relaxed acceptance) to $g = +U$ (100% stringent acceptance). These graphs can be useful in choosing an economically acceptable decision rule [19].

9.4 Guide to Use of the Graphs

The basic quantities needed to use Figs. 16–19, in addition to the specified tolerance, T, are the process capability index, C_p, and the measurement capability index, C_m.

The process capability index is evaluated by studying the distribution of characteristics (such as lengths) produced by the process and estimating the process standard deviation, σ_p. Then $C_p = T/6\sigma_p$ and the appropriate figure can be chosen. It is unlikely that C_p will be exactly equal to ⅓ or any of the other three values shown in the four graphs; one can interpolate between the graphs in order to choose appropriate guard bands.

The measurement capability index is evaluated by performing an uncertainty analysis of the measurement process and calculating the standard uncertainty, u_m, associated with the measured values of characteristics. Then $C_m = T/4u_m = T/2U$, which fixes the particular curve in the figure corresponding to the value of C_p. One can interpolate between these curves for values of C_m different from those shown.

Example. Suppose a process is characterized by C_p = ⅓ and the measurement capability is such that $C_m = 4$. In order to maximize return on investment, the consumer's risk, R_C, must be held to 2% or less. Where should the guard bands be located in order to satisfy the risk requirement?

Figure 19 shows the risks, R_C and R_P, for C_p = ⅓. A vertical line upwards from $R_C = 2\%$ intersects the curve for $C_m = 4$ near the point $g = 0$. In this case, the gauging limits coincide with the tolerance limits, so there are no guard bands and the acceptance zone coincides with the tolerance zone. The decision rule is then 4:1 simple acceptance. This operating point has a producer's risk of $R_P \approx 3\%$, so that about 3% of measured workpieces would fail inspection and yet conform with specification.

Now suppose the same process capability (C_p = ⅓) and a less-accurate measurement process with measurement capability index $C_m = 2$. The $R_C = 2\%$ vertical line intersects the curve for $C_m = 2$ near the point $g = +0.25$. With the less-capable measurement system, a consumer's risk of 2% can still be achieved by using a 25% stringent acceptance decision rule. The price to be paid for the less-accurate measurements is an increase in producer's risk from about 3% to more than 10%, resulting in more costly rejection of conforming workpieces.

This example shows that a desired quality level for accepted workpieces can be achieved using a variety of measurement systems. In general, the lower the measurement capability, the lower the cost of measurement. But less-accurate measurements will require a reduced acceptance zone and more rejection of conforming workpieces. The higher cost associated with the rejection of these conforming workpieces must be balanced against the lower cost of the measurements. The optimum choice of an accept/reject decision rule is thus a matter of business economics.

REFERENCES

[1] ISO Standards Handbook — Statistical methods for quality control, Volume 1: Terminology and symbols, Geneva, 1995.

[2] International vocabulary of basic and general terms in metrology, ISO, Geneva, 1993.

[3] ASME B89.7.3.1-2001, Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance to Specifications.

[4] ASME B89.7.2-1999, Dimensional Measurement Planning.

[5] ISO 3534-2:1993, Statistics — Vocabulary and symbols — Part 2: Statistical quality control.

[6] ISO 3951:1989, Sampling procedures and charts for inspection by variables for percent nonconforming.

[7] Guide to the expression of uncertainty in measurement, ISO, Geneva, 1995. This document is also available as ANSI/NCSL Z540.2-1997, U.S. Guide to the Expression of Uncertainty in Measurement.

[8] ASME Y14.5M-1994, Dimensioning and Tolerancing.

[9] Kotz, K. and Johnson, L. J., Process Capability Indices, Chapman & Hall, 1993.

[10] A2LA News: The Newsletter of the American Association for Laboratory Accreditation, July 1998.

[11] www.isa.org/Content/NavigationMenu/Members_and_Leaders/Dictionary/Definitions/DefineT.htm, ISA — The Instrumentation, Systems, and Automation Society.

[12] Eagle, A. R., "A Method for Handling Errors in Testing and Measuring," Ind. Qual. Control 10:3, pp. 10–15, 1954.

[13] Grubbs, F. A. and Coon, H. J., "On Setting Test Limits Relative to Specification Limits," Ind. Qual. Control 10:3, pp. 15–20, 1954.

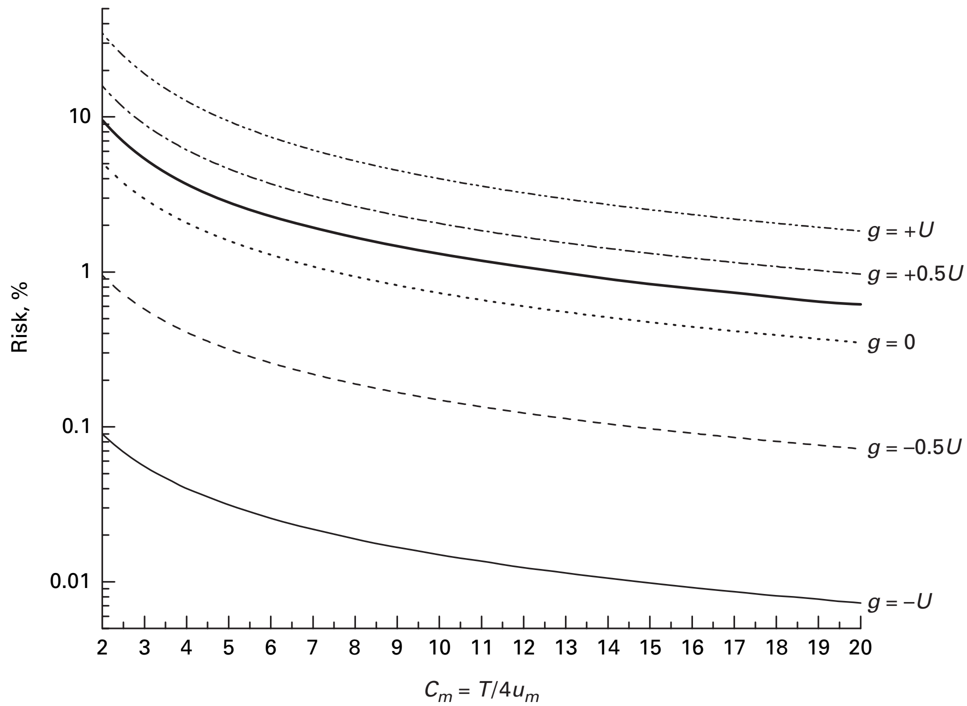
[14] Deaver, D., "How to Maintain Your Confidence (in a World of Declining Test Uncertainty Ratios)," 1993 NCSL Workshop and Symposium, pp. 133–153.

[15] Deaver, D., "Guardbanding with Confidence," 1994 NCSL Workshop and Symposium, pp. 383–394.

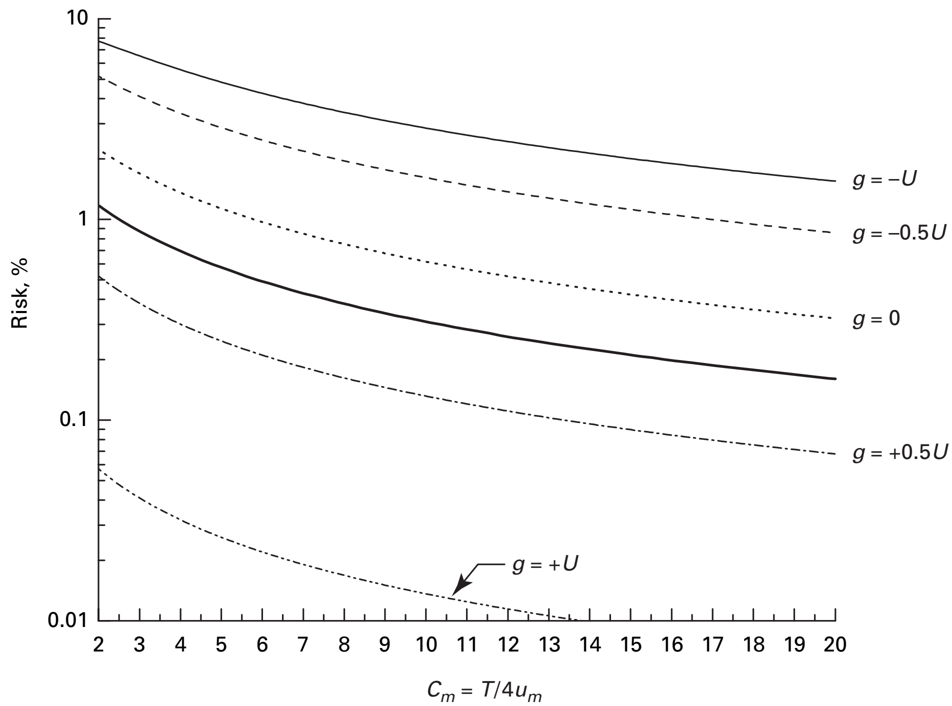
[16] Deaver, D., "Managing Calibration Confidence in the Real World," 1995 NCSL Workshop and Symposium, pp. 1–17.

[17] Deaver, D., "Guardbanding and the World of ISO Guide 25: Is There Only One Way?," 1998 NCSL Workshop and Symposium, pp. 319–332.

[18] Williams, R. and Hawkins, C., "The Economics of Guardband Placement," Proceedings of the 24th IEEE International Test Conference, Baltimore, 1993.



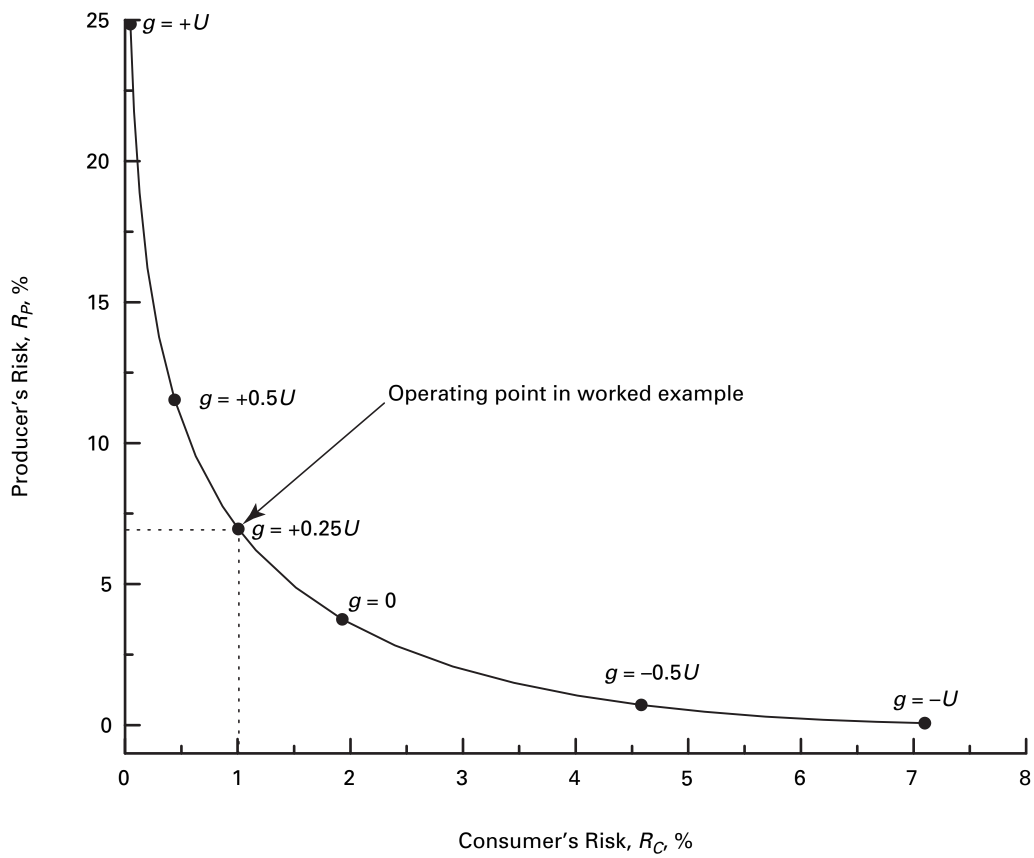
(a) Producer's Risk for C_p = 0.55



(b) Consumer's Risk for C_p = 0.55

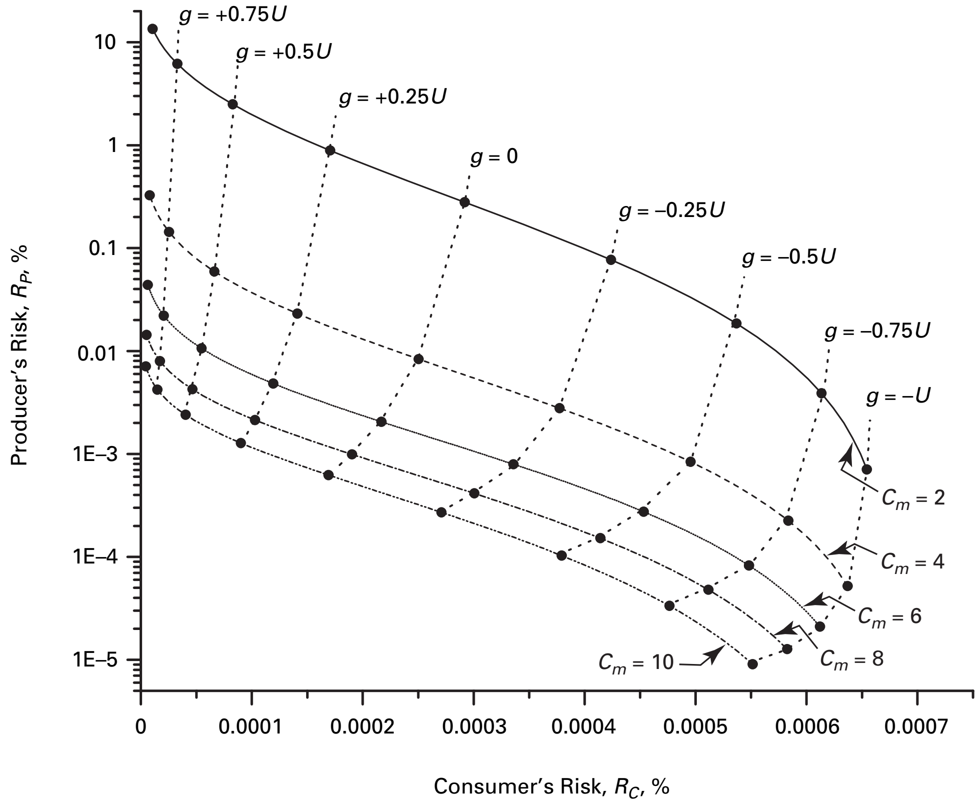
GENERAL NOTE: The risks are plotted versus the measurement capability index, $C_m = T/4u_m$, for various values of guard band magnitude $g = T_U - G_U$. The thick solid curve corresponds to $g = 0.25U$, the value used in the worked example. Positive values of g correspond to guard bands inside the tolerance limits, implying a stringent acceptance decision rule.

Fig. 14 Producer's and Consumer's Risks for the Worked Example



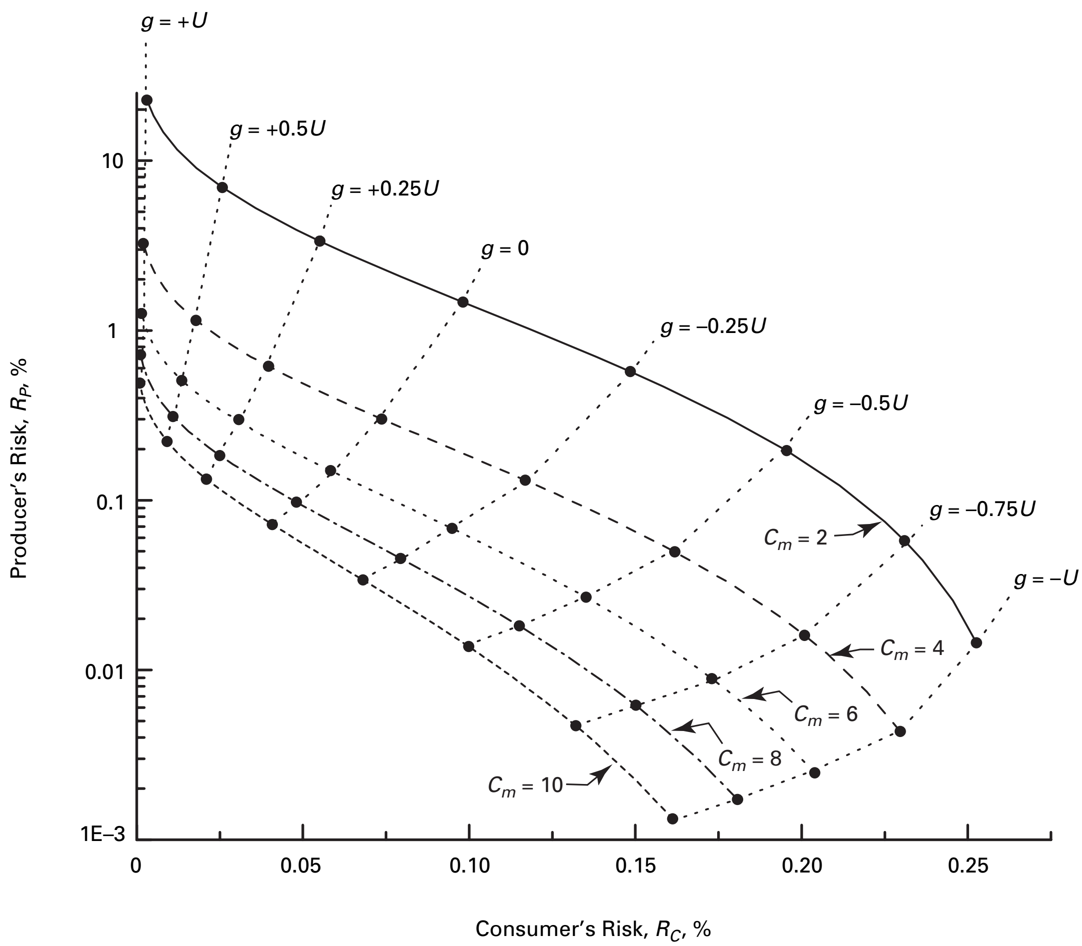
GENERAL NOTE: Several values of guard band magnitude g are shown. The point $g = 0$ corresponds to a simple 2.5:1 acceptance decision rule (no guard bands), while values of $g > 0$ correspond to stringent acceptance. Moving the location of the guard bands invokes a tradeoff between the two kinds of risks. The choice of a particular value of g depends upon the costs associated with accepting bad spacers or rejecting good ones. Analysis of these costs is a matter of business economics, with guard bands chosen to maximize profit [19]. In this example, the producer is willing to scrap about 7% of conforming spacers in order to reduce the fraction of falsely accepted nonconforming spacers to 1%. The operating point that achieves this objective, shown above, is $g = +0.25U$, stringent acceptance with a 25% guard band.

Fig. 15 Producer's Risk Versus Consumer's Risk for the Worked Example With $C_p = 0.55$ and $C_m = 2.5$



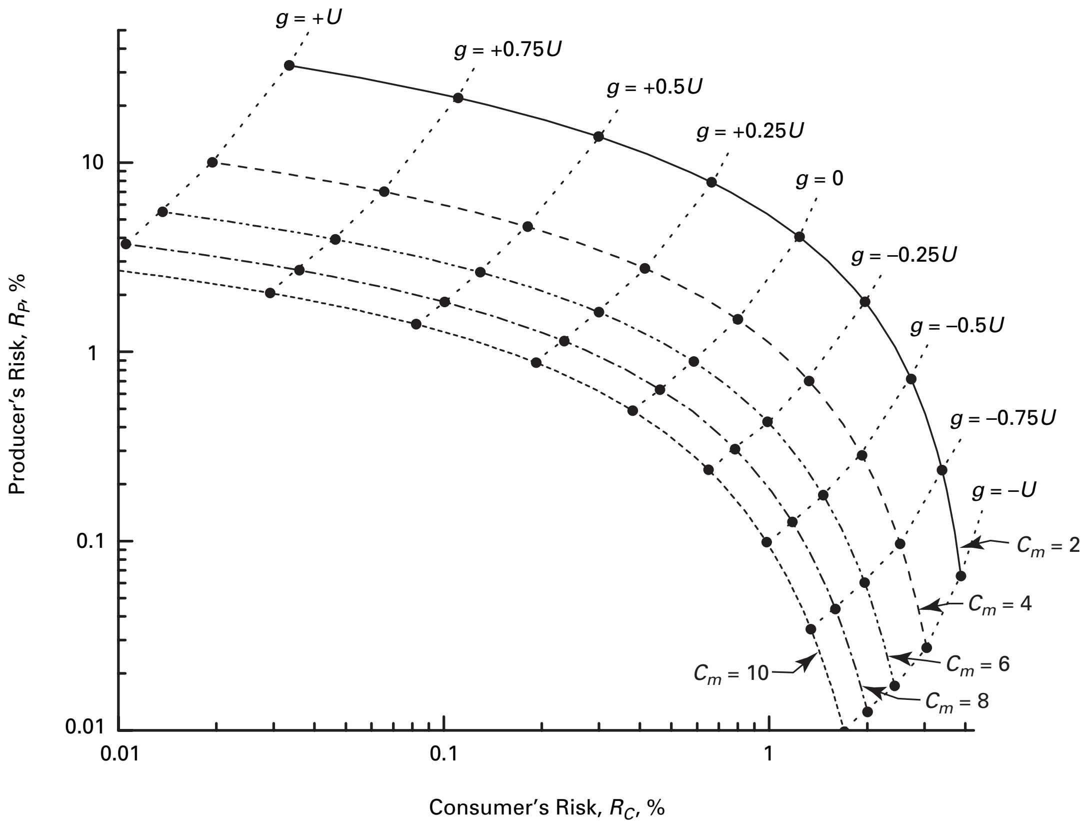
GENERAL NOTE: The five curves correspond to values of measurement capability index C_m in a range from 2 to 10. The solid points locate guard bands ranging from $g = -U$ (100% relaxed acceptance) to $g = +U$ (100% stringent acceptance). The curves can be useful in choosing a decision rule after an economic analysis has provided an acceptable balance of risks. For example, if $C_m = 8$, then choosing a relaxed acceptance rule with a 25% guard band ($g = -0.25U$) would result in a consumer's risk of about 0.0003% and a producer's risk of about 0.0004%. Note that the R_P scale is logarithmic.

Fig. 16 Producer's Risk Versus Consumer's Risk for $C_p = 1.5$



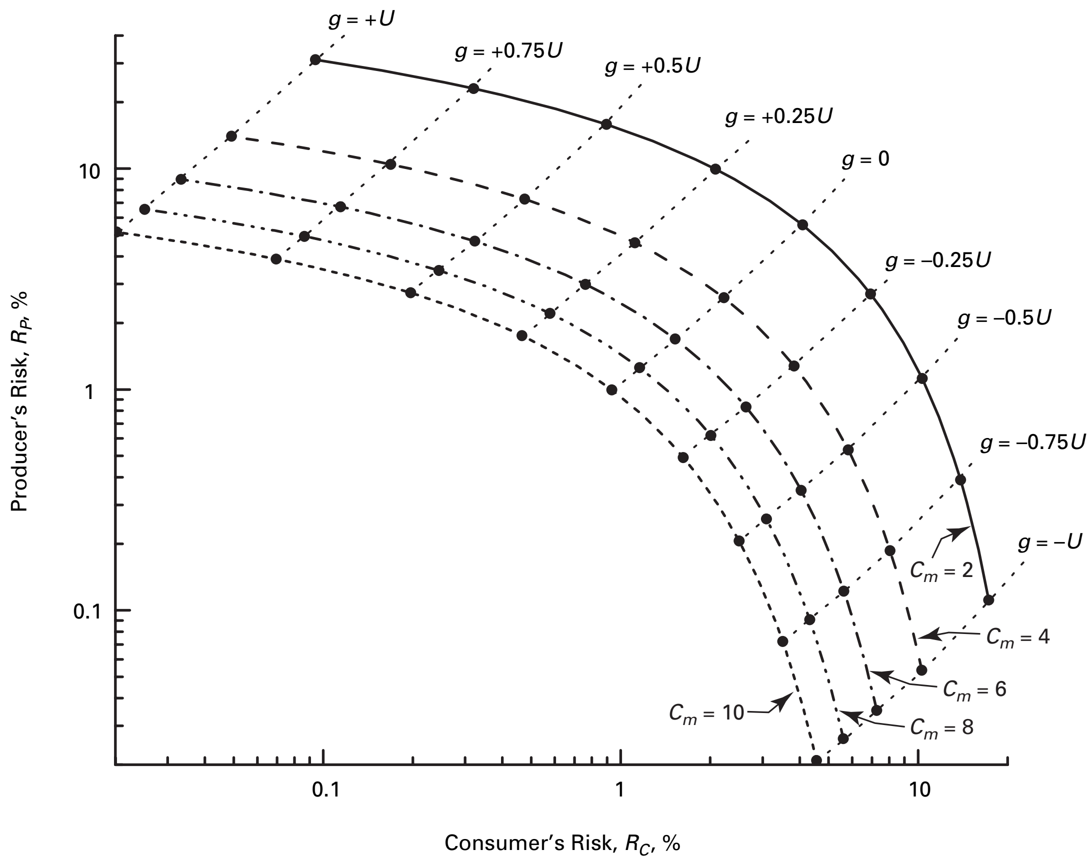
GENERAL NOTE: The five curves correspond to values of measurement capability index C_m in a range from 2 to 10. The solid points locate guard bands ranging from $g = -U$ (100% relaxed acceptance) to $g = +U$ (100% stringent acceptance). Note that the R_P scale is logarithmic.

Fig. 17 Producer's Risk Versus Consumer's Risk for $C_p = 1$



GENERAL NOTE: The five curves correspond to values of measurement capability index C_m in a range from 2 to 10. The solid points locate guard bands ranging from $g = -U$ (100% relaxed acceptance) to $g = +U$ (100% stringent acceptance). Both scales are logarithmic.

Fig. 18 Producer's Risk Versus Consumer's Risk for $C_p = 2/3$



GENERAL NOTE: The five curves correspond to values of measurement capability index C_m in a range from 2 to 10. The solid points locate guard bands ranging from $g = -U$ (100% relaxed acceptance) to $g = +U$ (100% stringent acceptance). Both scales are logarithmic.

Fig. 19 Producer's Risk Versus Consumer's Risk for $C_p = {}^1/_3$

MANDATORY APPENDIX I
PROPERTIES OF GAUSSIAN PROBABILITY DENSITIES

I-1 GAUSSIAN PROBABILITY DENSITY

Assume that knowledge of the length, X, of a workpiece, after performing a measurement, is well-characterized by a Gaussian (normal) probability density

$$p(x|I_m) = \frac{1}{u_m\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-x_m}{u_m}\right)^2\right] \qquad (1)$$

where the result, x_m, is the best estimate (expectation) of X and u_m is the standard deviation of the density function.[1] Information, I_m, includes the measurement data as well as prior knowledge of the characteristics of the production process. The density [Eq. (1)] expresses the fact that, since X cannot be known exactly, there are an infinite number of possible lengths consistent with what is known, summed up in I_m. The density means that $p(x|I_m)\Delta x$ is the probability that X lies in the interval $(x, x+\Delta x)$. Because the length is certain to have some value, the density is normalized, which means that

$$\int_{-\infty}^{\infty} p(x|I_m)\,dx = 1$$

For a coverage factor, k, the expanded uncertainty is defined to be $U = ku_m$. The probability that the length of the measured workpiece lies in an expanded uncertainty interval $[x_m - U, x_m + U]$ about the measurement result is just the fraction of the area under the density [Eq. (1)] between these limits, given by

$$p(|x - x_m| \le U|I_m) = \int_{x_m - ku_m}^{x_m + ku_m} p(x|I_m)\,dx \qquad (2)$$

The probability [Eq. (2)] is called a containment probability, coverage probability, or (in the GUM) a level of confidence.

I-2 GAUSSIAN INTEGRALS

In computing probabilities and the risks of quantities such as pass and fail errors, one needs to evaluate integrals of Gaussian functions between finite limits. Such integrals cannot be evaluated in a simple closed form, and are therefore evaluated numerically and tabulated. In order to simplify the notation, it is convenient to introduce a standard normal probability density function, $f_0(z)$, defined by

$$f_0(z) \equiv \frac{1}{\sqrt{2\pi}}\exp(-z^2/2) \qquad (3)$$

There are two common ways that one finds Gaussian integrals evaluated, either in tabular form or computed numerically in computer software. These are

(a) the standard normal cumulative distribution function, $\Phi(k)$, defined by

$$\Phi(y) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\exp(-z^2/2)\,dz$$
$$= \int_{-\infty}^{y} f_0(z)\,dz \qquad (4)$$

(b) the error function, erf(y), defined by

$$\mathrm{erf}(y) \equiv \frac{2}{\sqrt{\pi}}\int_{-\infty}^{y}\exp(-z^2)\,dz$$

These functions are simply related. From their definitions it can be seen that

$$\Phi(y) = \frac{1}{2}\left[1 + \mathrm{erf}(y/\sqrt{2})\right]$$

Given these definitions, consider the probability that the value of X lies in the interval $a \le X \le b$. This is

$$p(a \le X \le b|I_m) = \int_a^b p(x|I_m)\,dx$$

Given the Gaussian density [Eq. (1)], this is

$$p(a \le X \le b \mid I_m) = \frac{1}{u_m\sqrt{2\pi}}\int_a^b \exp\left[-\frac{1}{2}\left(\frac{x-x_m}{u_m}\right)^2\right]dx$$

Now, making the substitutions $z = (x - x_m)/u_m$, $dz = dx/u_m$, this equation becomes

[1] In the nomenclature of the GUM, the quantity u_m is called the combined standard uncertainty, denoted $u_c(x)$. The simpler notation u_m is used in this Report.

$$p(a \le X \le b | I_m) = \int_{\frac{a-x_m}{u_m}}^{\frac{b-x_m}{u_m}} f_0(z)\, dz$$

$$= \Phi\left(\frac{b - x_m}{u_m}\right) - \Phi\left(\frac{a - x_m}{u_m}\right) \quad (5)$$

using Eqs. (4) and (3) for $f_0(z)$.

I-3 LEVELS OF CONFIDENCE FOR GAUSSIAN DENSITIES

In the special case where a and b define an expanded uncertainty interval about the measurement result x_m, which means $a = x_m - ku_m$ and $b = x_m + ku_m$, Eq. (5) reduces to

$$\begin{aligned} p(|X - x_m| \le ku_m | I_m) &= \int_{-k}^{k} f_0(z)\, dz \\ &= \Phi(k) - \Phi(-k) \\ &= \mathrm{erf}\left(k/\sqrt{2}\right) \\ &\equiv P_0(k) \end{aligned}$$

Any good text on statistics, computational software package, or commercial spreadsheet software will show the familiar results for these symmetric Gaussian containment probabilities or levels of confidence

$$\begin{aligned} P_0(1) &= \Phi(1) - \Phi(-1) \\ &= \mathrm{erf}\left(1/\sqrt{2}\right) \\ &\equiv 0.683 \end{aligned}$$

$$\begin{aligned} P_0(2) &= \Phi(2) - \Phi(-2) \\ &= \mathrm{erf}\left(2/\sqrt{2}\right) \\ &\equiv 0.955 \end{aligned}$$

$$\begin{aligned} P_0(3) &= \Phi(3) - \Phi(-3) \\ &= \mathrm{erf}\left(3/\sqrt{2}\right) \\ &\equiv 0.997 \end{aligned}$$

These containment probabilities are often called 1-sigma, 2-sigma, and 3-sigma levels of confidence.

I-4 FRACTION OF WORKPIECES CONFORMING FOR A GAUSSIAN FREQUENCY DISTRIBUTION

A production process produces workpieces whose length frequency distribution is well-characterized by the Gaussian function

$$f(x) = \frac{1}{\sigma_p \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - x_0}{\sigma_p}\right)^2\right] \quad (6)$$

where x_0 is the average length of a workpiece and σ_p is the measured standard deviation of the process, calculated from the measured lengths of a large sample of workpieces. Given the distribution [Eq. (6)], after a very long production run, the fraction of workpieces with lengths in a small range $[x, x+\Delta x]$ would be approximately $f(x)\Delta x$. The fraction of workpieces with lengths in any desired range from a minimum value $x_{\min}$ to a maximum value $x_{\max}$ can then be calculated by integrating the distribution $f(x)$ over this interval

$$\text{fraction of lengths between } x_{\min} \text{ and } x_{\max} = \int_{x_{\min}}^{x_{\max}} f(x)\, dx \quad (7)$$

If the process has been adjusted so that the average length, x_0, lies at the center of a specified tolerance zone of width T, the fraction, f_C, of workpieces that conform to specification is given by Eq. (7) with $x_{\min} = T - x_0/2$ and $x_{\max} = T + x_0/2$

$$f_C = \frac{1}{\sigma_p \sqrt{2\pi}} \int_{x_0 - T/2}^{x_0 + T/2} \exp\left[-\frac{1}{2}\left(\frac{x - x_0}{\sigma_p}\right)^2\right] dx$$

Now, letting $z = (x - x_0)/\sigma_p$ and $dz = dx/\sigma_p$, and defining the inherent process capability index by $C_p = T/6\sigma_p$, the fraction conforming, f_C, becomes

$$\begin{aligned} f_C &= \int_{-3C_p}^{3C_p} f_0(z)\, dz \\ &= \Phi(3C_p) - \Phi(-3C_p) \\ &= \mathrm{erf}\left(\frac{3C_p}{\sqrt{2}}\right) \quad (8) \end{aligned}$$

Consider the numerical example in para. 4.3 of this Report. For this process, the tolerance is $T = 0.4$ mm and the process standard deviation is $\sigma_p = 0.12$ mm, so that $C_p = 0.551$. From Eq. (8), it follows that the desired probability is $\Phi(1.653) - \Phi(-1.653) = 0.902$. Thus, 90.2% of workpieces produced by this process would have lengths in conformance to the tolerance requirement and 100% − 90.2% = 9.8% would be nonconforming.

MANDATORY APPENDIX II
RISK CALCULATIONS

II-1 CONSUMER'S RISK

A procedure for calculating the consumer's risk, R_C, was given in para. 9. The details of these calculations are presented here.

The consumer's risk is the probability of a pass error or false accept, meaning that a nonconforming characteristic passes a measurement inspection. Let $P\overline{C}$ denote the joint proposition that the measured characteristic passes inspection and does not conform to specification.

Conditioned on the available information, I_0, that characterizes knowledge of the production and measurement processes, R_C is just equal to the probability that $P\overline{C}$ is true

$$R_C = p(P\overline{C}|I_0)$$

Denoting by x the possible values of the characteristic X, the risk above can be written as a marginal probability

$$\begin{aligned} R_C &= \int_{x \in R} p(P\overline{C}x|I_0)\, dx \\ &= \int_{x \in R} p(P\overline{C}|xI_0) \cdot p(x|I_0)\, dx \qquad (1) \end{aligned}$$

where the range of integration, R, includes all values of X that are outside of the conformance zone defined by the tolerance limits (T_L, T_U): $R = [X < T_L \text{ and } X > T_U]$. The first integral in Eq. (1) has been rewritten in the second line by using the product rule of probability theory.

The quantity $p(x|I_0)$ in Eq. (1) is the prior density for the values of characteristic X. It is assumed that this prior knowledge is well-characterized by a Gaussian distribution

$$\begin{aligned} p(x|I_0) &= \frac{1}{u_p\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - x_0}{u_p}\right)^2\right] \\ &= N(x; x_0, u_p^2) \qquad (2) \end{aligned}$$

where

x_0 = $(T_L + T_U)/2$
= the nominal value (assumed to lie at the center of the tolerance zone)

u_p = associated standard uncertainty that characterizes the range of reasonably probable values of X prior to performing a measurement

Together, the tolerance, $T = T_U - T_L$, and standard uncertainty, u_p (taken equal to the process standard deviation, σ_p), define the inherent process capability index, $C_p = T/6u_p$.

The quantity $p(P\overline{C}|xI_0)$ in Eq. (1) is the probability that a characteristic known to be nonconforming nevertheless yields a measurement result, x_m, within the acceptance zone, defined by the gauging (or test) limits $G_L \le x_m \le G_U$. This situation is illustrated in Fig. II-1.

For a given assumed known value, x, and given measurement process, there will be a range of reasonably probable measurement results, x_m, that are consistent with the available information, I_0. For a measurement process corrected for all known significant systematic errors, one's degree of belief in this range of probable results will be characterized by a probability density function, $p(x_m|xI_0)$, taken to be a Gaussian density whose standard deviation is equal to the measurement combined standard uncertainty, u_m

$$\begin{aligned} p(x_m|xI_0) &= \frac{1}{u_m\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_m - x}{u_m}\right)^2\right] \\ &= N(x_m; x, u_m^2) \qquad (3) \end{aligned}$$

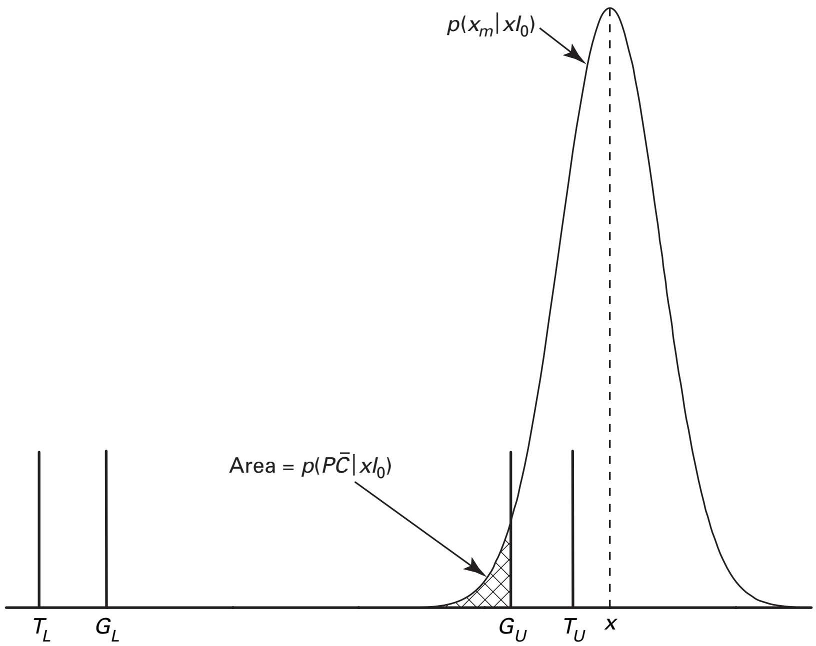
Together, the tolerance, T, and standard uncertainty, u_m, define the measurement capability index, $C_m = T/4u_m$.

The probability density [Eq. (3)] is shown in Fig. II-1. The conditional probability, $p(P\overline{C}|xI_0)$, of a pass error for this particular value of x is equal to the fraction of the area under $p(x_m|xI_0)$ contained between the gauging limits, shown cross-hatched in the figure. This probability is

$$\begin{aligned} p(P\overline{C}|xI_0) &= \int_{G_L}^{G_U} p(x_m|xI_0)\, dx \\ &= \frac{1}{u_m\sqrt{2\pi}} \int_{G_L}^{G_U} \exp\left[-\frac{1}{2}\left(\frac{x_m - x}{u_m}\right)^2\right] \end{aligned}$$

Substituting $w \equiv (x_m - x)/u_m$ and using $f_0(z) = (1/\sqrt{2\pi}) \exp(-z^2/2)$ in this expression gives

$$\begin{aligned} p(P\overline{C}|xI_0) &= \int_{W_L}^{W_U} f_0(z)\, dw \\ &= \Phi(w_U) - \Phi(w_L) \qquad (4) \end{aligned}$$



GENERAL NOTE: For this particular item, x is too large, lying beyond the upper tolerance limit, T_U. The curve shows the distribution $p(x_m|xI_0)$ of probable values of x_m that might reasonably result when measuring a characteristic X with true value x. The probability that the characteristic passes inspection and is accepted is equal to the fraction of the area under the curve $p(x_m|xI_0)$, shown cross-hatched, within the acceptance zone defined by the gauging limits (G_L, G_U).

Fig. II-1 Probability of Accepting a Nonconforming Workpiece

where

$w_U = (G_U - x)/u_m$

$w_L = (G_L - x)/u_m$

$\Phi(w)$ = standard normal cumulative distribution function

Substituting the results from Eqs. (2) and (4) in Eq. (1) gives:

$$R_C = \frac{1}{u_p\sqrt{2\pi}}\int_{-\infty}^{T_L}\left[\Phi(w_U) - \Phi(w_L)\right]\exp\left[-\frac{1}{2}\left(\frac{x - x_0}{u_p}\right)^2\right]dx$$

$$+\frac{1}{u_p\sqrt{2\pi}}\int_{T_U}^{\infty}\left[\Phi(w_U) - \Phi(w_L)\right]\exp\left[-\frac{1}{2}\left(\frac{x - x_0}{u_p}\right)^2\right]dx$$

Letting $z = (x - x_0)/u_p$, this becomes

$$R_C = \int_{-\infty}^{z_L}\left[\Phi(w_U) - \Phi(w_L)\right]f_0(z)\,dz$$

$$+\int_{z_U}^{\infty}\left[\Phi(w_U) - \Phi(w_L)\right]f_0(z)\,dz \qquad (5)$$

where

$$Z_L = \frac{T_L - x_0}{u_p} = -3C_p$$

$$Z_U = \frac{T_U - x_0}{u_p} = +3C_p$$

Now define the guard band multiplier $h \equiv (T_U - G_U)/U = (T_U - G_U)/2u_m$ and let $r = u_p/u_m$. Then, since $x = x_0 + zu_p$, the quantities w_U and w_L can be written as functions of z as follows:

$$\begin{aligned} w_U &= \frac{G_U - x}{u_m} \\ &= 2(C_m - h) - rz \\ &= \gamma - rz \end{aligned}$$

$$\begin{aligned} w_L &= \frac{G_L - x}{u_m} \\ &= -2(C_m - h) - rz \\ &= -\gamma - rz \end{aligned}$$

where the constant $\gamma = 2(C_m - h)$. Then the quantity in brackets in the integrals in Eq. (5) can be replaced by the function, $F(z)$, defined by

$$F(z) \equiv \Phi(\gamma - rz) - \Phi(-\gamma - rz) \qquad (6)$$

Finally, the consumer's risk [Eq. (5)] becomes

$$R_C = \int_{-\infty}^{-3C_p} F(z)\, f_0(z)\, dz + \int_{3C_p}^{\infty} F(z)\, f_0(z)\, dz \tag{7}$$

Equation (7) here is the same as Eq. (11) in the main text.

II-2 PRODUCER'S RISK

The producer's risk, R_P, is the probability of a fail error or false reject, meaning that a conforming characteristic fails a measurement inspection. Let FC denote the joint proposition that the measured characteristic fails inspection and conforms to specification.

Conditioned on the available information, I_0, that characterizes knowledge of the production and measurement processes, R_P is just equal to the probability that FC is true

$$R_P = p(FC|I_0)$$

In analogy with Eq. (1) for the consumer's risk, the producer's risk can be written as a marginal probability

$$\begin{aligned} R_P &= \int_{T_L}^{T_U} p(FCx|I_0)\, dx \\ &= \int_{T_L}^{T_U} p(FC|xI_0) \cdot p(x|I_0)\, dx \end{aligned} \tag{8}$$

where the limits of integration cover the range of conforming values of X, i.e., the tolerance zone, $T_L \le x \le T_U$.

The quantity $p(FC|xI_0)$ in Eq. (8) is the probability that a characteristic known to be conforming nevertheless yields a measured value, x_m, outside of the acceptance zone defined by $G_L \le x_m \le G_U$. This situation is illustrated in Fig. II-2.

For the particular value of x shown in Fig. II-2, the conditional probability, $p(FC|xI_0)$, of a pass error is equal to the fraction of the area under the curve $p(x_m|xI_0)$ that lies outside of the acceptance zone defined by the gauging limits (G_L, G_U). For the Gaussian density [Eq. (3)], this probability is

$$\begin{aligned} p(FC|xI_0) &= \frac{1}{u_m\sqrt{2\pi}} \int_{-\infty}^{G_L} \exp\left[-\frac{1}{2}\left(\frac{x_m - x}{u_m}\right)^2\right] dx_m \\ &\quad + \frac{1}{u_m\sqrt{2\pi}} \int_{G_U}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{x_m - x}{u_m}\right)^2\right] dx_m \end{aligned} \tag{9}$$

Letting $w = (x_m - x)/u_m$, $f_0(z) = (1/\sqrt{2\pi}) \exp(-z^2/2)$, and substituting in Eq. (9) yields

$$\begin{aligned} p(FC|xI_0) &= \int_{-\infty}^{w_L} f_0(z)\, dw + \int_{w_U}^{\infty} f_0(z)\, dw \\ &= \Phi(w_L) + 1 - \Phi(w_U) \end{aligned} \tag{10}$$

where

$$w_L = (G_L - x)/u_m$$
$$w_U = (G_U - x)/u_m$$

Using the result [Eq. (10)] together with the prior probability density [Eq. (2)], the producer's risk [Eq. (8)] becomes

$$R_P = \frac{1}{u_p\sqrt{2\pi}} \int_{T_L}^{T_U} [1 - \Phi(w_U) + \Phi(w_L)] \exp\left[-\frac{1}{2}\left(\frac{x - x_0}{u_p}\right)^2\right] dx$$

and letting $z = (x - x_0)/u_p$, this becomes

$$R_P = \int_{-3C_p}^{3C_p} [1 - \Phi(w_U) + \Phi(w_L)]\, f_0(z)\, dz \tag{11}$$

From the steps leading to the definition of $F(z)$ in Eq. (6), it can be seen that

$$1 - \Phi(w_U) + \Phi(w_L) = 1 - F(z)$$

so that the producer's risk [Eq. (11)] is

$$R_P = \int_{-3C_p}^{3C_p} [1 - F(z)]\, f_0(z)\, dz \tag{12}$$

Equation (12) here is the same as Eq. (10) in the main text.

II-3 ONE-SIDED MEASUREMENTS

Some conformance tests involve characteristics with a single specification (or tolerance) limit. Examples include

(a) the roundness error of a cylindrical shaft, specified to be no greater than 0.1 μm

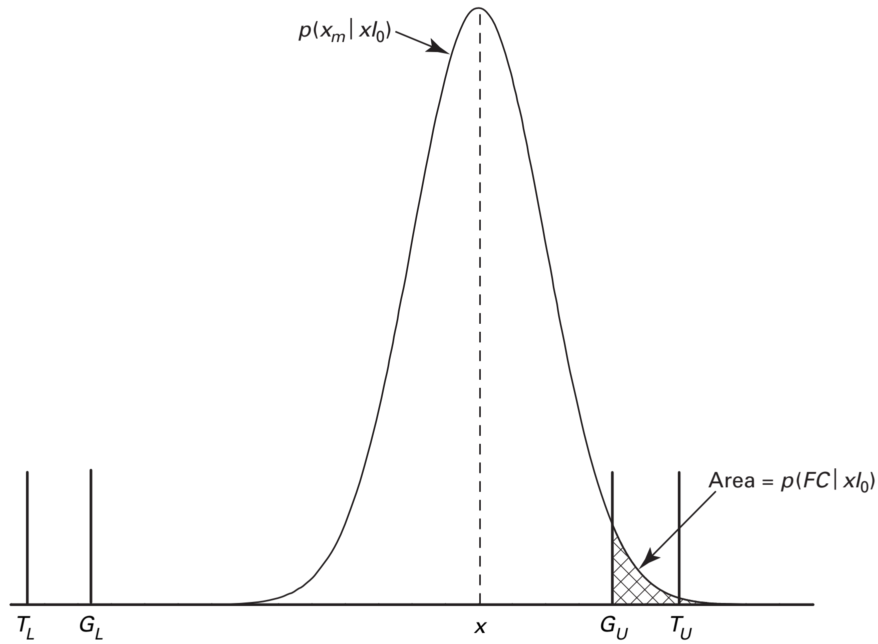
(b) the concentration of mercury in a sample of industrial wastewater, required to be less than 10 ng/L

(c) a particulate air filter, specified to remove no less than 99.97% of particles 0.3 μm in diameter

(d) the insertion loss of a fiber optic connector, specified to be less than 0.2 dB

A typical example of a single-sided specification zone and associated guard band is shown in Fig. II-3.

The probabilities, R_C (consumers' risk) and R_P (producer's risk), of pass errors and fail errors in such a case can



GENERAL NOTE: For this particular item, x lies within the tolerance zone. The curve shows the distribution, $p(x_m|xI_0)$, of probable values of x_m that might reasonably result when measuring a characteristic X with true value x. The probability that the characteristic fails inspection and is rejected is equal to the fraction of the area under the curve $p(x_m|xI_0)$, shown cross-hatched, that lies outside of the acceptance zone defined by the gauging limits (G_L, G_U). For this particular item, there is a negligible probability that x_m would be less than the lower gauging limit, G_L.

Fig. II-2 Probability of Rejecting a Conforming Workpiece

be calculated in a manner analogous to the procedures derived above for two-sided measurements. For the consumer's risk,

$$R_C = \int_T^{\infty} p(P\overline{C}|xI_0) \cdot p(x|I_0)\, dx \qquad (13)$$

where the range of nonconforming values of X is $T \le x < \infty$.

The conditional consumer's risk, $p(P\overline{C}|xI_0)$, following the development leading to Eq. (4), is

$$p(P\overline{C}|xI_0) = \frac{1}{u_m\sqrt{2\pi}} \int_0^G \exp\left[-\frac{1}{2}\left(\frac{x_m - x}{u_m}\right)^2\right] dx$$

$$= \int_{w_1}^{w_2} F_0(z)\, dw$$

$$= \Phi(w_2) - \Phi(w_1)$$

where

$w_1 = -x/u_m$

$w_2 = (G - x)/u_m$

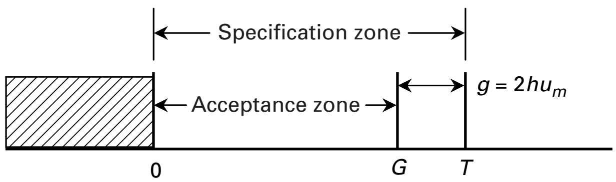
Then Eq. (13) becomes

$$R_C = \int_T^{\infty} \left[\Phi(w_2) - \Phi(w_1)\right] \cdot p(x|I_0)\, dx \qquad (14)$$

A similar analysis, following the development leading to Eq. (10), yields the producer's risk

$$R_P = \int_0^T \left[1 - \Phi(w_2) + \Phi(w_1)\right] \cdot p(x|I_0)\, dx \qquad (15)$$

The function $p(x|I_0)$ in Eqs. (13) and (15) is the prior probability density and characterizes knowledge of X before performing a measurement. For a quantity restricted to the range $x \ge 0$, such as the concentration of mercury in a sample of water, one could represent prior knowledge by a Gaussian function. Such a function, however, would have to be truncated and set equal to zero for impossible values of x, namely $x \le 0$, requiring



GENERAL NOTE: Here a characteristic of interest, X (such as the concentration of a water contaminant), is greater than or equal to zero and specified to have a value less than an upper limit, T. A gauging (or test) limit, G, is set inside the specification limit, T, creating a stringent acceptance zone. The guard band has width $g = 2hu_m$, where u_m is the standard uncertainty associated with the result x_m of the test measurement and h is the guard band multiplier chosen in the course of formulating a decision rule.

Fig. II-3 One-Sided Specification Zone

the calculation of a new normalization constant so that the truncated Gaussian density integrates to one.[1]

Just as in the two-sided case, assigning a prior probability density in one-sided decision problems is commonly based upon a measured frequency distribution (histogram) of characteristics (flatness errors, contaminant concentrations, etc.) acquired from a representative sample. The prior probability density, $p(x|I_0)$, will then follow the measured frequency distribution, $f(x)$.

In a case where values of the characteristic near zero are rarely observed, such a measured frequency distribution can often lead to the assignment of a gamma probability density, defined by

$$p(x|I_0) = g(x; a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} \text{ for } x > 0 \qquad (16)$$

Here a and b are two positive parameters, and $\Gamma(a)$ is the gamma function

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x}\, dx \text{ for } a > 0$$

The expectation, $E(X)$, and variance, $\mathrm{Var}(X)$, of the gamma density [Eq. (16)] are simply related to the parameters a and b

$$E(X) = \int_0^\infty x g(x; a, b)\, dx = \frac{a}{b} \qquad (17)$$

$$\mathrm{Var}(X) = \int_0^\infty [x - E(x)]^2 g(x; a, b)\, dx = \frac{a}{b^2} \qquad (18)$$

Given a particular state of prior information, appropriate values for a and b can be easily calculated using these results.

In the most common case, prior information about possible values of X is obtained by measuring a large sample of characteristics, and calculating a sample mean and standard deviation (or variance). Assuming that the process is stable, the best estimate and associated standard uncertainty of a future, unmeasured characteristic are assigned to be equal to the measured sample statistics as described above. Denoting the estimate (or expectation) of X by $\bar{x}$ and the associated variance (whose square root is the standard uncertainty) by $u^2(x)$, then Eqs. (17) and (18) can be solved for the appropriate values of a and b

$$a = \frac{\bar{x}^2}{u^2(x)} \text{ and } b = \frac{\bar{x}}{u^2(x)} \qquad (19)$$

II-4 EXAMPLE: RISK CALCULATIONS FOR BALL-BEARING PRODUCTION

A manufacturer produces large numbers of precision ball bearings. The performance specification for these bearings requires that the radial error motion[2] be less than 2 μm. In order to characterize the production process, the radial error motions of a large sample of bearings are measured, using a high-accuracy test apparatus with negligible measurement uncertainty. For this sample, the average observed radial error motion is $\bar{x} = 1$ μm, with an associated sample standard deviation $s = 0.5$ μm.

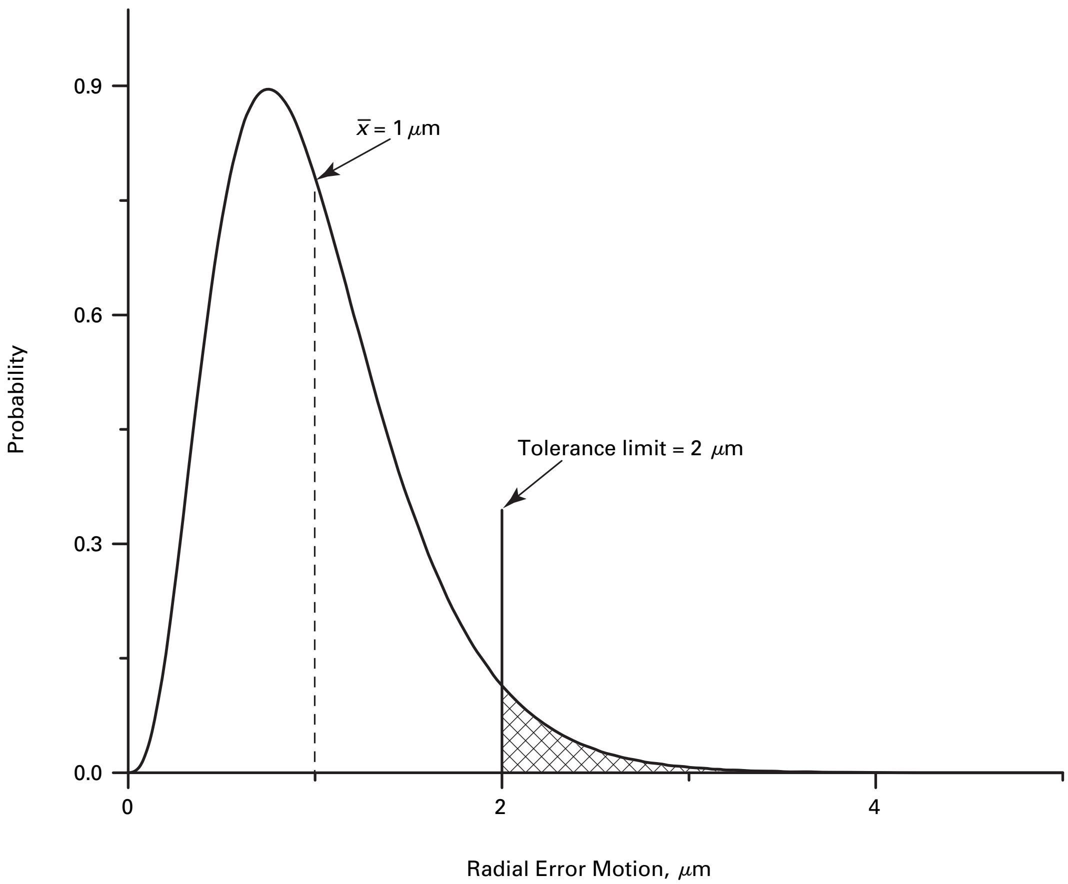
Prior to shipment, bearings are tested for conformance to specification. In these tests, the radial error motion is measured using a calibrated test apparatus. The standard uncertainty of the test measurements is $u_m = 0.25$ μm. For economic reasons, the fraction of nonconforming bearings sold to customers as conforming must be held to 0.1% or less. How can a gauging limit, G, be chosen to achieve this level of consumer's risk?

Solution. Since the radial error motion is always positive, the prior density for values of radial error motion will be modeled using a gamma probability density. Based on the sample measurements, the expectation and standard uncertainty of the prior density are assigned: $\bar{x} = 1$ μm, $u(x) = s = 0.5$ μm. Then, using Eq. (19), the parameters a and b are calculated

$$a = \frac{1^2}{(0.5)^2} = 4 \text{ and } b = \frac{1}{(0.5)^2} = 4$$

[1] A Gaussian pdf for any inherently positive quantity (such as the length of a spacer) will assign a positive belief to impossible values. For real spacers, the probability of negative lengths is infinitesimal; for quantities such as flatness errors or contaminant concentrations, a sizable fraction of the total probability might be distributed over impossible negative values. Thus, a Gaussian pdf might reasonably model belief in the length of a spacer but be an unreasonable model of belief in a quantity whose value is very close to zero.

[2] Radial error motion of a bearing is undesired motion perpendicular to the axis of rotation. For a perfect bearing, the radial error motion would be zero; any real bearing will have a positive radial error motion.



GENERAL NOTE: The specification zone is the region $0 \le x \le T$, where the tolerance is $T = 2$ μm. The mean of the distribution is the estimate $\bar{x} = 1$ μm and the associated standard uncertainty is $u(\bar{x}) = 0.5$ μm. The most probable value of X is the mode of the distribution, which in this case is equal to 0.75 μm. Because the distribution is not symmetric, the mean and mode do not coincide. For this state of prior knowledge, there is a probability of about 4.2% that a roller bearing chosen at random would display a radial error motion outside of the 2 μm tolerance, a region shown cross-hatched in the figure. If all bearings produced were shipped without being measured, about 4.2% of them would be nonconforming. The post-process inspection system is designed to reduce the probability of shipping nonconforming bearings. A gauging limit is desired that will reduce this risk (the consumer's risk, R_C) to 0.1% or better.

Fig. II-4 Prior Probability Density for Radial Error Motion of Ball Bearing

The prior density, $p(x|I_0)$, for values of radial error motion is then

$$p(x|I_0) = g(x;\,4,\,4) = \frac{128}{3}\,x^3 e^{-4x} \tag{20}$$

This probability density function is shown in Fig. II-4.

Given the prior probability density [Eq. (20)], the risks can be calculated using Eqs. (14) and (15). The quantities w_1 and w_2 in these expressions are given by

$$w_1 = -x/u_m = -4x$$

$$w_2 = \frac{G - x}{u_m} = \frac{T - 2hu_m - x}{u_m} = 8 - 2h - 4x$$

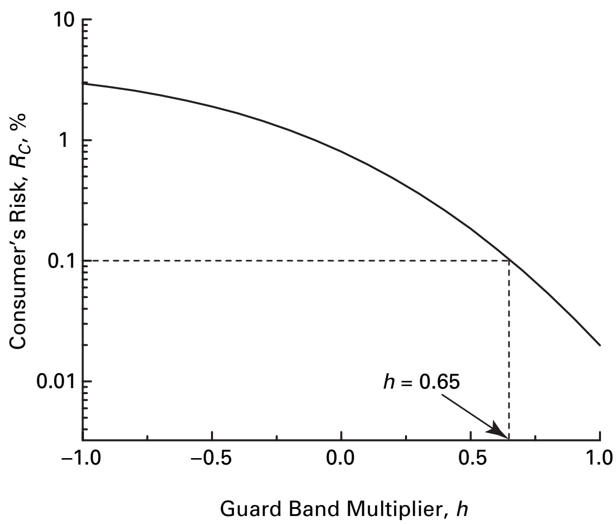
These quantities have been written explicitly in terms of h, the guard band multiplier (see Fig. II-3). The consumer's and producer's risks, as functions of the location of the guard band, are thus given by

$$R_C(h) = \frac{128}{3}\int_2^{\infty} [\Phi(8 - 2h - 4x) - \Phi(-4x)]x^3 e^{-4x}\,dx$$

$$R_P(h) = \frac{128}{3}\int_0^{2} [1 - \Phi(8 - 2h - 4x) + \Phi(-4x)]x^3 e^{-4x}\,dx$$

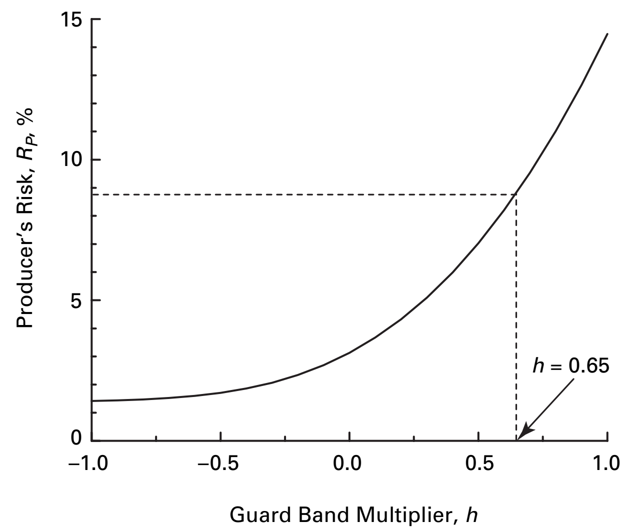
These integrals cannot be evaluated in closed form, but they can be calculated numerically for any chosen values of h.

The risks $R_C(h)$ and $R_P(h)$ are shown in Figs. II-5 and II-6, for $-1 \le h \le 1$. Positive h means $G < T$ (stringent



GENERAL NOTE: For $h = 0.65$ ($G = T - 2hu_m = 1.7$ μm), the desired risk, $R_C = 0.1\%$, is achieved.

Fig. II-5 Consumer's Risk Versus Guard Band Multiplier for Ball-Bearing Example



GENERAL NOTE: The choice $h = 0.65$ that limits consumer's risk to 0.1% results in a producer's risk of nearly 9%.

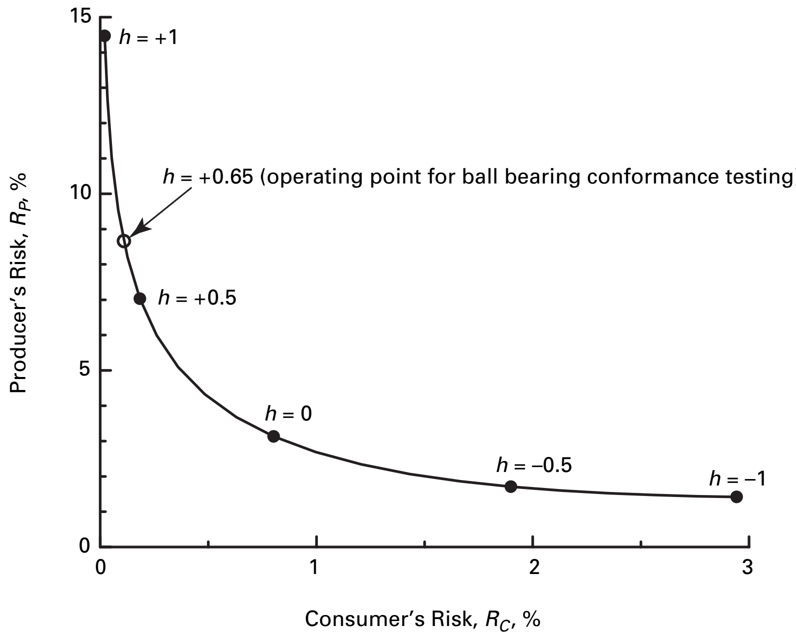
Fig. II-6 Producer's Risk Versus Guard Band Multiplier for Ball-Bearing Example

acceptance) and negative h means $G > T$ (relaxed acceptance). For $h = 0$, there is no guard band ($G = T$), leading to the decision rule called simple acceptance.

Figure II-5 shows that the desired level of consumer's risk, $R_C = 0.1\%$, can be achieved by setting the guard band multiplier $h = +0.65$. This results in a stringent acceptance zone with a gauging (or test) limit at $G = T - 2hu_m = 1.7$ μm. This solves the decision problem.

Figure II-6 shows the producer's risk, R_P, that results from the chosen decision rule. For $h = +0.65$, the producer's risk is nearly 9%. This means that 9 out of every 100 bearings that fail inspection are actually conforming, resulting in the loss of revenue that would accrue if these good bearings were sold. The generation of an increasing amount of conforming scrap is a cost of stringent acceptance rules, which guard against accepting nonconforming products. This general rule is well-illustrated by Fig. II-7, which shows R_P versus R_C for the ball-bearing example.

As seen in Fig. II-7, acting to reduce the consumer's risk, R_C, by reducing the size of the acceptance zone (increasing h) will always result in an increase in the producer's risk, R_P. There are costs associated with accepting a nonconforming bearing (probability R_C) and with rejecting a conforming bearing (probability R_P). In general, the producer must choose an operating point along a curve such as that shown in Fig. II-7 that will balance these risks and yield a maximum profit. The choice of such an operating point is a business decision that requires an economic analysis of the decision problem [18].



GENERAL NOTE: Any point on the curve corresponds to a particular value of h, the guard band multiplier, with several particular values identified. Acting to reduce the consumer's risk by moving the gauging limit farther inside the tolerance zone (increasing h) always increases the risk of falsely rejecting conforming bearings. An economic analysis is required to choose an optimal decision rule that maximizes profit. The operating point in the example is identified by the open circle.

Fig. II-7 Producer's Risk Versus Consumer's Risk for Ball-Bearing Example

ASME B89.7.4.1-2005

ISBN 0-7918-2943-X

9 780791 829431

L08105