How to Maintain Your Confidence
(in a World of Declining Test Uncertainty Ratios)

David K. Deaver
John Fluke Mfg. Co., Inc.
Everett, Washington

ABSTRACT:

Setting test limits different than specification limits influences the risk of accepting defective units (consumer risk) and rejecting conforming units (producer risk). Much has been written about setting limits to accomplish various strategies such as maintaining a minimum consumer risk, equalizing consumer risk with producer risk, minimizing total risk, or equalizing the cost of faulty test decisions between the producer and consumer.

This paper reviews the statistical foundation for making decisions as to where to place test limits and includes a multitude of charts to simplify what used to be tedious calculations of the test limit, consumer risk, and producer risk. The implications of various test strategies can be seen very quickly using the charts.

The MathCAD ® [1] formulas used to generate the charts are included so MathCAD users can duplicate or customize the charts. Representative formulas are shown in Appendix C.

INTRODUCTION:

Despite the efforts of Design to eliminate it, Production to minimize it, and Sales to deny it, variability exists in all manufacturing processes. The uncertainty of a product is dependent on the variability of the individual units produced, the variability of the process that manufactures them, and the systematic errors that can shift the mean of the resulting distribution such as the systematic component of calibration standard uncertainties, interpolation errors, non-linearity, etc. The bulk of the literature, including this paper, deals with the variability, or random errors, as the major contributor to the product uncertainty.

Specifications must be established strategically, positioning them to balance the need to make the product easy to produce with having specifications competitive with products from other manufacturers. While they are an indication of the variability of a product and their associated manufacturing process, they do not describe the variability explicitly because different manufacturers make different assessments as to the best place to assign specifications with respect to the variability of their product. In addition, similar products from different manufacturers may still have enough differences in operating characteristics to make comparisons of the specifications difficult. References [2-4] deal with the setting of specifications and their relationship to the product uncertainty.

PRODUCT UNCERTAINTY:

The output of many manufacturing processes can be described with a normal probability distribution. The probability distribution about the mean is shown below in Fig. 1.

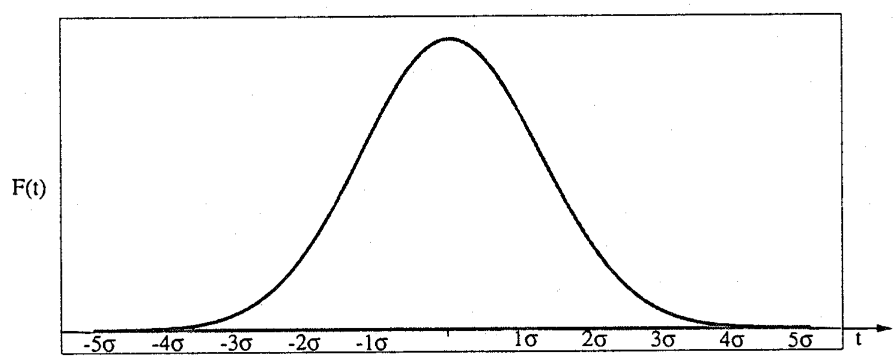


Fig. 1 Normal Probability Distribution

The probability that the performance of the unit under test (UUT) is within its specifications is the area under the curve between the specification limits (SL), assumed to be centered about the mean. The risk of a unit being outside of its specifications (OOT) is the area under the curve outside the specification limits and depends on how conservatively the product is specified with respect to its variability. The probability of an in-tolerance condition is the integral of the probability distribution from the lower specification limit to the upper specification limit. Eq. 1 shows the integral and Fig. 2, a tabulation of the probabilities for a number of specification limits with respect to the product's standard deviation, σ.

Eq. 1

$$P(\text{conforms}) := \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_{-SL}^{SL} \exp\left(-\frac{t^2}{2}\right) dt$$

Specification Limits	Probability Unit Conforms (%)	Probability Unit Doesn't Conform (%)
±1.0σ	68.3	31.7
±1.5σ	86.6	13.4
±2.0σ	95.4	4.6
±2.5σ	98.8	1.2
±3.0σ	99.7	0.3

Fig. 2 In and Out of Tolerance Probabilities for Specification Limits from 1σ to 3σ

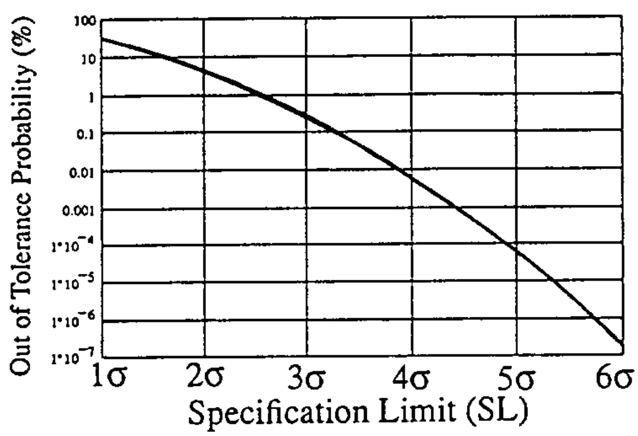


Fig. 3 Out-of-Tolerance Probabilities for Specification Limits from 1σ to 6σ

Fig. 3 is a plot of the probability that a unit is out of tolerance for symmetric specification limits (SL) set from 1σ to 6σ. Some manufacturers are setting goals of 6σ control of their processes with respect to their product specifications. As can be seen from the plot, defect rates of 6σ processes centered on the mean are at parts per billion levels. If mean shifts of 1.5σ are allowed, the defect rate is 3.4 ppm [5].

TEST DECISIONS:

When a product is tested for conformance to its specifications, a test standard (STD) is used to determine if it is in or out of tolerance. The test standard has its own probability distribution, however, producing uncertainty in the determination of an out-of-tolerance condition. The probability of accepting a defective unit is the joint probability of a unit being defective, combined with the probability that the test standard reports such a unit as being in tolerance. This condition is shown graphically in Fig. 4.

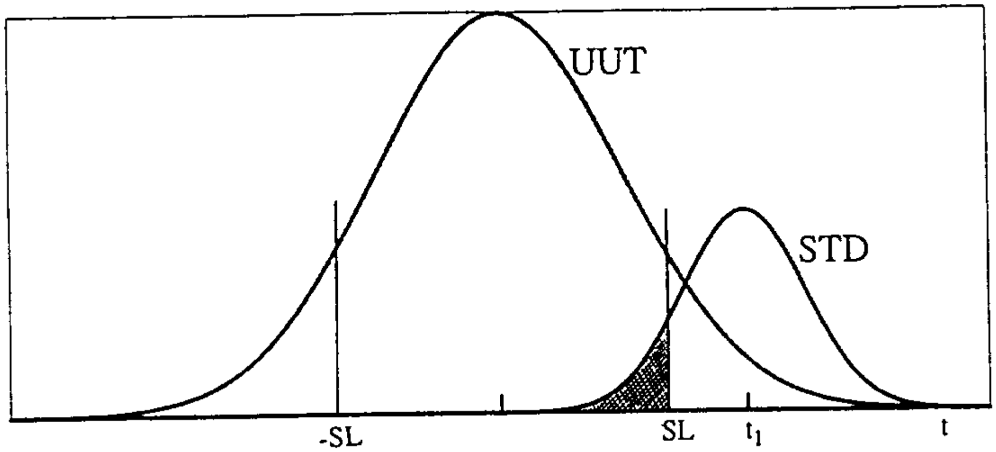


Fig. 4 Out-of-Tolerance Unit Reported as Conforming

The curve labeled UUT is the probability distribution of the unit under test. STD is the distribution of readings reported by the test standard for a UUT at value t_1. The shaded area is the probability that a unit with value t_1 will be reported as being in tolerance; that is, between the lower specification limit and the upper specification limit here assumed to be symmetric about the mean at -SL and +SL respectively. The probability function represented by the shaded area is described by Eq. 2.

Eq. 2

$$F(CR, t_1) := \frac{1}{\sqrt{2\cdot\pi}}\cdot\exp\left[-\frac{(t_1)^2}{2}\right]\cdot\int_{-R\cdot(t_1+SL)}^{-R\cdot(t_1-SL)}\frac{1}{\sqrt{2\cdot\pi}}\cdot\exp\left(-\frac{s^2}{2}\right)ds$$

where R is the test uncertainty ratio (TUR) defined as the uncertainty of the UUT divided by the uncertainty of the STD. It is important to note that the TUR is the ratio of UUT's specifications to the STD's specifications only if both devices were specified with the same confidence. The portion of Eq. 2 to the left of the integral represents the probability that the UUT has value t_1 and the portion within the integral, the probability that t_1 is reported inside the specification limits.

If the "shaded area" is calculated for all values of t outside the specification limits, a probability distribution for the consumer risk is obtained. Fig. 5 shows the distribution for symmetric test limits and test uncertainty ratios from 1 to 4.

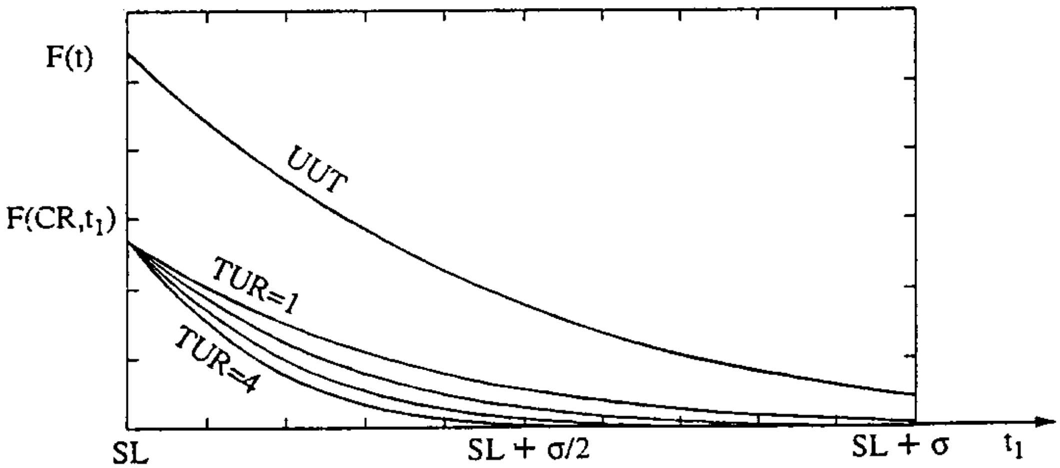


Fig. 5 Probability Distribution of Consumer Risk

As one would expect, at the specification limit, half of the units are reported as conforming and half defective regardless of the TUR. As the units under test exceed the specification limit by greater amounts, test standards with lower uncertainty (higher TUR) report fewer of the UUTs as being in tolerance.

Similarly, the uncertainty of the test standard can result in conforming units being rejected (producer risk). This is illustrated in Fig. 6 where a conforming unit at t_2 has a distribution of values which are measured by the STD.

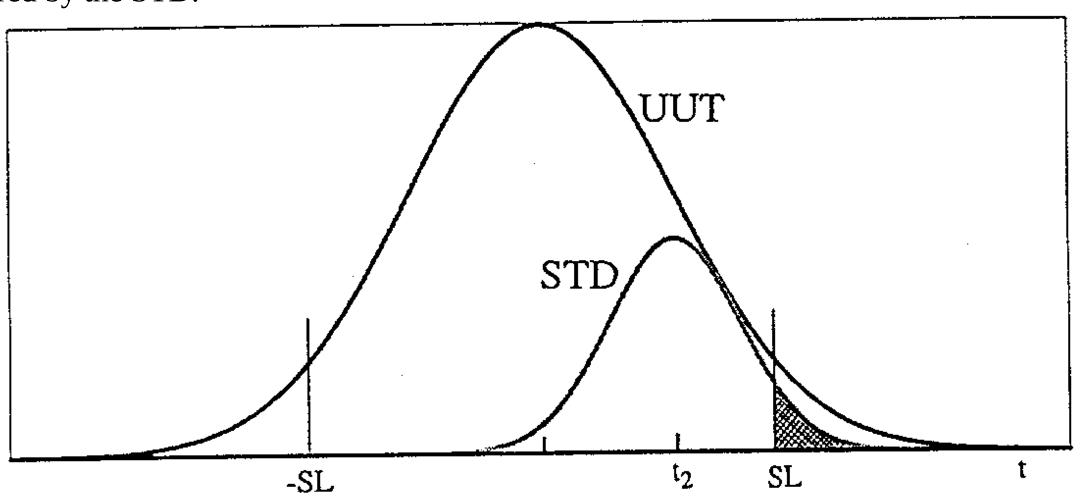


Fig. 6 Conforming Unit Reported Out-of-Tolerance (Producer Risk)

The Producer Risk at t_2 is represented by the shaded area and can be calculated for each value to t located between the specification limits as shown in Eq. 3.

Eq. 3

$$P(PR, t_2) := \frac{1}{2\cdot\pi}\cdot\exp\left[-\frac{(t_2)^2}{2}\right]\cdot\int_{R\cdot(SL - t_2)}^{\infty}\frac{1}{\sqrt{2\cdot\pi}}\cdot\exp\left(-\frac{s^2}{2}\right) ds$$

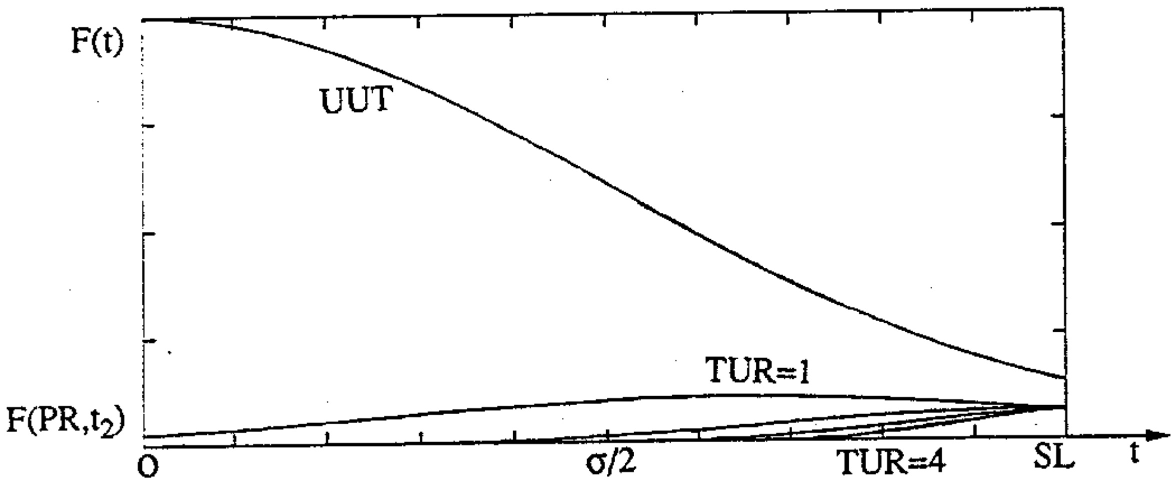


Fig. 7 Probability Distribution of Producer Risk

Fig. 7, the evaluation of Eq. 3 for all values of t within the specification limits, shows the probability distribution of the producer risk. Again, it can be seen, at the specification limit, half of the units under test will be reported out of tolerance and half in, with the reporting errors less for higher TURs.

Integrating the consumer distribution for values of t outside the specification limits and the producer distribution for values of t inside the specification limits yields the Consumer Risk (CR), Eq. 4 and Producer Risk (PR), Eq. 5.

Eq. 4

$$CR := \frac{1}{\pi}\cdot\int_{SL}^{\infty}\int_{-R\cdot(t+SL)}^{-R\cdot(t-SL)}\exp\left[-\frac{(s^2+t^2)}{2}\right] ds\, dt$$

Eq. 5

$$PR := \frac{1}{\pi}\cdot\int_{-SL}^{SL}\int_{R\cdot(SL-t)}^{\infty}\exp\left[-\frac{(s^2+t^2)}{2}\right] ds\, dt$$

where, R is the TUR and s is the local variable for the STD and it can be assumed the specification limits of the UUT and the STD are centered on the means of their respective distributions and can be represented as -SL and +SL respectively.

Using MathCAD, the double integrals of Eq. 4 and Eq. 5 were calculated numerically and plotted in Fig. 8. Note that decreasing the TUR increases the risk of faulty test decisions for both the consumer and producer. Also significant is the sensitivity of the risk to the setting of the specifications of the UUT; how conservatively it is specified with respect to it's variability. With a 4:1 TUR, and specification limits set at 2σ, the consumer risk is 0.8%. However, with a more conservatively specified unit at SL=2.5σ, the chance of accepting defective units would be 0.25%. Even if the TUR was reduced to 1:1, the consumer risk would be only 0.5%. No testing at all would only result in the acceptance of 1.2% defective units, about the same as the less conservatively specified unit (SL=2σ) tested with a TUR of 2:1.

The uncertainty associated with the testing of a unit's conformance to its specifications is dependent both on the product variability and the precision of the conformance test. In the past, more attention has been given to the standardization of conformance testing than to the standardization of the setting of specifications as a function of the product's variability; presumably, because product variability information is more accessible to the manufacturer than the end user. It is clear that a consumer must have confidence in the manufacturer of its equipment as well as in its own incoming inspection process. Some companies have been able to improve product quality and significantly reduce inspection costs by partnering with their suppliers through vendor quality programs such as Motorola's SSPC, one of the components of their 6 sigma quality effort, and Fluke's Aim for Excellence program.

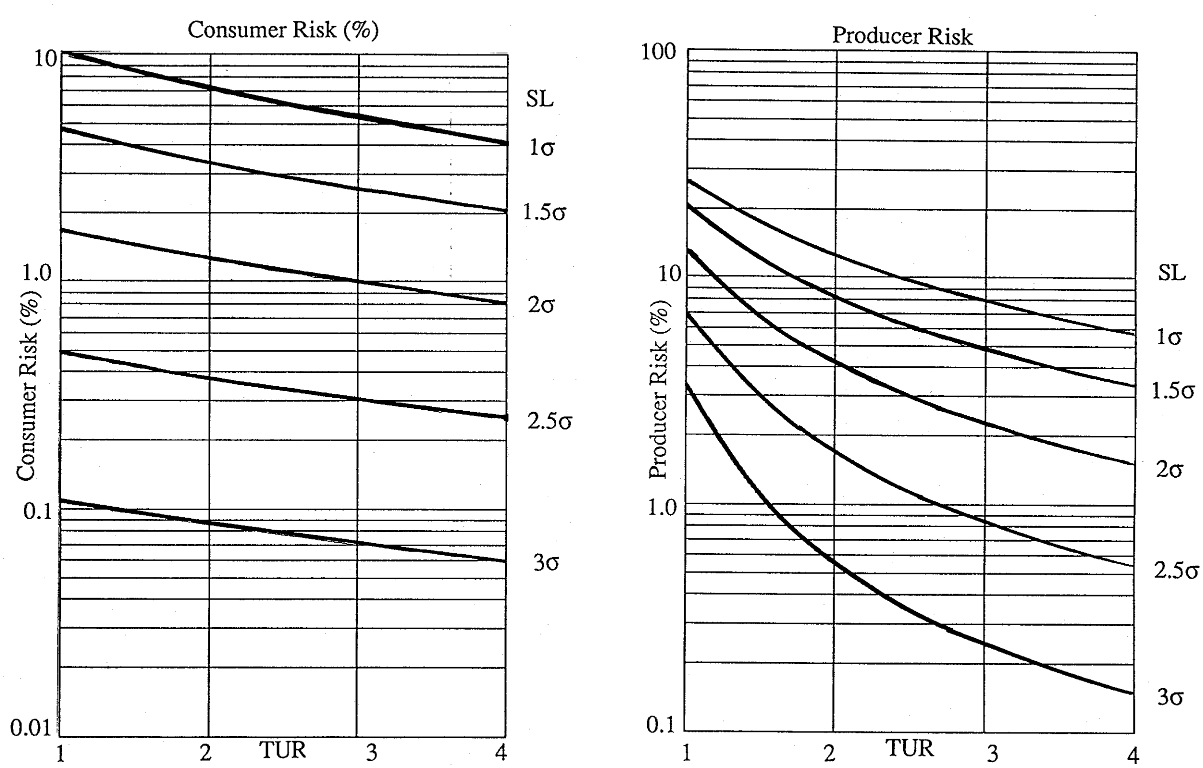


Fig. 8 Consumer Risk and Producer Risk as a Function of TUR

GUARDBANDING

Most calibration labs face the difficulty of having calibration standards which will not meet the desired or required TUR for some of the workload. The metrologist must choose to lower the level of confidence in the measurement, invest in more precise standards, or undergo an analysis of the uncertainties and document the deviations from the required TUR.

Guardbanding, the technique of setting test limits different from specification limits, offers an additional alternative. Though the probability of making faulty test decisions increases with decreasing TURs, the test limits can be placed to set the desired level of consumer risk or producer risk. For example, it is possible, with a 2:1 TUR, to keep the same risk of accepting defective units as a 4:1 TUR by setting the test limits (TL) inside the specification limits. The price to be paid for controlling the consumer risk is that the producer risk can be much higher than for a 4:1 TUR.

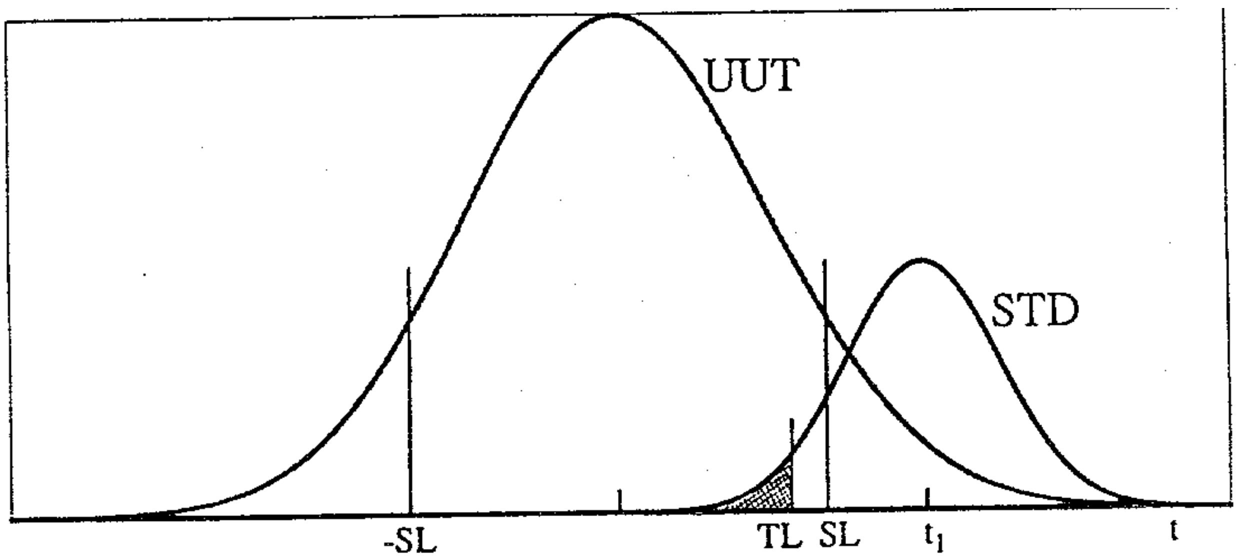


Fig. 9 Out-of-Tolerance Unit Reported as Conforming Despite Guardband

Fig. 9 shows the effects of having a TL inside the SL for symmetrical limits. The shaded area to the left of t_1 illustrates the probability that a unit outside the SL will be accepted. Compared with Fig. 4, the smaller shaded area shows the reduced probability of false accepts since units measuring inside the SL but greater than the TL will be rejected.

Rejecting these additional units increases the chances of rejecting conforming units, however. The shaded area in Fig. 10 associated with t_2 is increased over that of Fig. 6 by including the units falling between the TL and the SL.

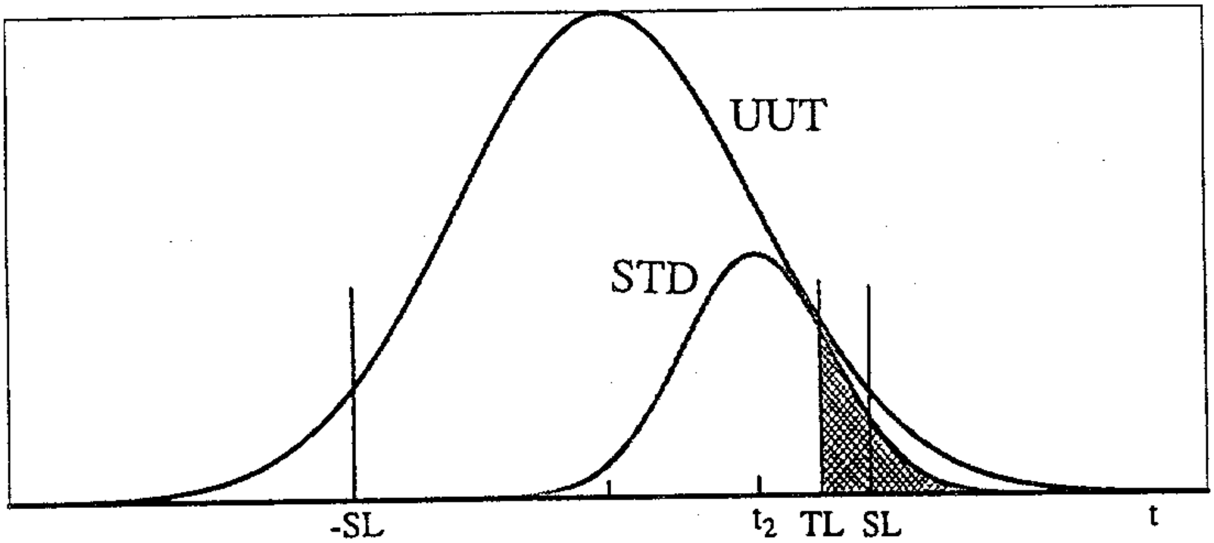


Fig. 10 Conforming Unit with Guardband Reported Non-Conforming

The shaded area in Fig. 10 (consumer risk) can be calculated by evaluating the double integral of Eq. 6. This is obtained by integrating the consumer risk probability function for all values of t lying outside the SL. K is the factor by which the specification limit is reduced to obtain the test limit. (TL = K * SL)

Eq. 6

$$CR := \frac{1}{\pi} \cdot \int_{SL}^{\infty} \int_{-R\cdot(t+K\cdot SL)}^{-R\cdot(t-K\cdot SL)} \exp\left[-\frac{\left(s^2+t^2\right)}{2}\right] ds\, dt$$

Similarly, the Producer Risk with guardband is shown in Eq. 7. It is obtained by integrating the shaded area of Fig. 10 for all values of t between the specification limits.

Eq. 7

$$PR := \frac{1}{\pi} \cdot \int_{-SL}^{SL} \int_{R\cdot(K\cdot SL - t)}^{\infty} \exp\left[-\frac{\left(s^2+t^2\right)}{2}\right] ds\, dt$$

Eagle presented the effects of setting a TL different than the SL in a classic paper in 1954 [6]. The contribution of the present paper is to present the guardband as a multiplier of the SL rather than a multiplier of the uncertainty of the test standard, and to provide sets of curves for UUT uncertainties of 1σ to 3σ rather than just the 2σ curves presented by Eagle. Eq. 6 is essentially the same consumer risk as Eagle's consumer risk equation with the nomenclature change. The producer risk, Eq. 7, is presented in a different form than Eagle's to make it a little more intuitive (especially to the author).

The risks with guardband, calculated from Eq. 6 and Eq. 7, are shown in Appendix A. Note that the curves, in each figure, for K=1.0 are the same as those shown in Fig. 8. The additional curves on each figure in the appendix show the consumer risk and producer risk for test limits set 5% to 30% inside the specification limits.

Hutchinson [7] pointed out that there is an implied consumer risk associated in standards such as MIL-STD-45662A. If it assumed that the specification limits are set at 2σ, the consumer risk is 0.8% for a 4:1 TUR, as shown in Fig. 8. Hutchinson calculated the guardband as a SL multiplier to keep the consumer risk constant at 0.8% independent of the TUR. Constant consumer risk is represented by the horizontal dashed lines on the consumer risk charts in Appendix A. The producer risk can be determined by noting the multiplier K and the TUR on the consumer risk chart, and finding the risk associated with the same K and TUR on the producer risk curve. From the 2σ curve, to maintain the consumer risk at 0.8%, the TL is set to 91% of the SL if one has only a 2:1 TUR. The resulting producer risk for K=0.91 and a 2:1 TUR is 6.8% as compared to 1.5% for K=1. As an aid to finding the producer risk for constant consumer risk, dashed lines are shown on the producer risk curves for consumer risk held constant at the 4:1, 3:1, and 2:1 TUR levels.

Fig. 11 shows, on a single chart, the guardband factors (K) for consumer risk held constant at the TUR=3 level and the TUR=4 level for specification limits from 1σ to 3σ. If it is desired to hold the consumer risk constant with declining TURs, Fig. 11 provides more resolution than the curves in Appendix A. However, Fig. 11 does not show the associated risks that are in the appendix.

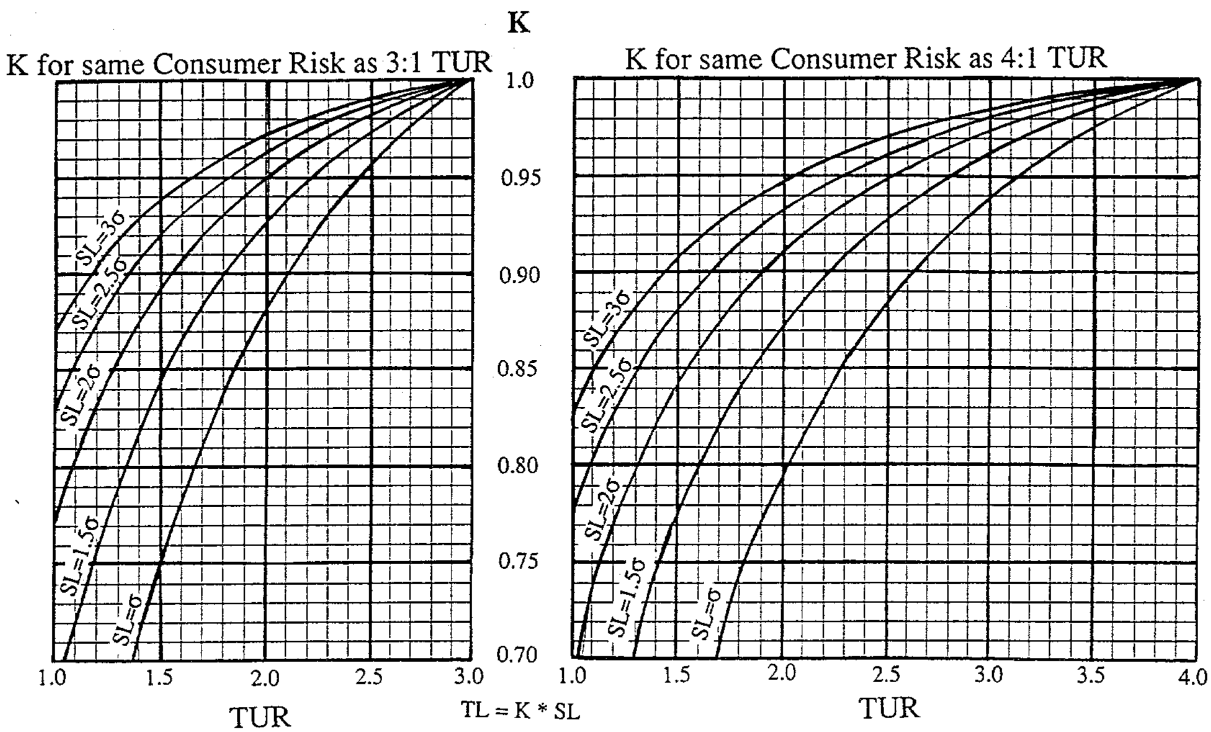


Fig. 11 Guardband Factors (K) to Maintain Consumer Risk at the Same Levels as with 3:1 and 4:1 TUR

Grubbs and Coon [8], as well as Weber and Hillstrom [9], discuss in more depth, some of the economic strategies of setting guardbands. Capricious setting of test limits to reduce consumer risk at the expense of producer risk may dramatically increase the cost of ownership of test equipment as manufacturer's costs are passed on to the consumer and as the consumer must bear the higher cost of maintaining calibration due to false rejects. However, judicious setting of guardband limits, while keeping product variability under control, can be a means to significantly reduce calibration costs without substantially increasing the costs due to false rejects.

CONCLUSIONS

It takes more than maintaining high TUR to maintain measurement quality at high levels. Ensuring that equipment is within specification and stays within specification requires control of variability as well as TUR. The charts in this paper provide the means to assess the risks associated with a wide range of product variability, TUR, and guardband factors. A means of justifying and documenting lower than 4:1 TURs for standards such as MIL-STD-45662A has been supported statistically for products whose variability is well controlled. Additionally, the implications of proposed test, calibration, purchasing, or incoming inspection strategies may be analyzed quickly using the charts.

ACKNOWLEDGMENTS

The author gratefully acknowledges the time taken from exceedingly busy schedules by a number of co-laborers to edit, critique, tutor, and encourage; especially David Agy, Norm Heyerdahl, and Les Huntley.

REFERENCES

[1] MathCAD is a registered trademark of MathSoft, Inc. The author has no interest nor affiliation with MathSoft except that of a satisfied user of the MathCAD software program.

[2] Agy, David, "Interpreting Specifications of Calibration Instruments", NCSL Workshop & Symposium, 1987

[3] Capell, Frank, "How Good is Your TUR?", Evaluation Engineering, January, 1991, pp. 80-84

[4] Read, Sherry L. and Timothy R. C., "Statistical Issues in Setting Product Specifications", Hewlett Packard Journal, June, 1988, pp. 6-11

[5] Harry, Mikel J., "The Nature of Six Sigma Quality", Motorola Inc., Government Electronics Group

[6] Eagle, Alan R. "A Method for Handling Errors in Testing and Measuring", Industrial Quality Control, March, 1954, pp. 10-15

[7] Hutchinson, Bill, "Setting Guardband Test Limits to Satisfy MIL-STD-45662A Requirements", NCSL Workshop & Symposium, pp. 305-309

[8] Grubbs, Frank E. and Coons, Helen J., "On Setting Test Limits Relative to Specification Limits", Industrial Quality Control, March, 1954, pp. 15-20

[9] Weber, Stephen F. and Hillstrom, Anne P., "Economic Model of Calibration Improvements for Automatic Test Equipment", NBS Special Publication 673, 1984

APPENDIX A

CONSUMER RISK
and
PRODUCER RISK
with GUARDBANDS

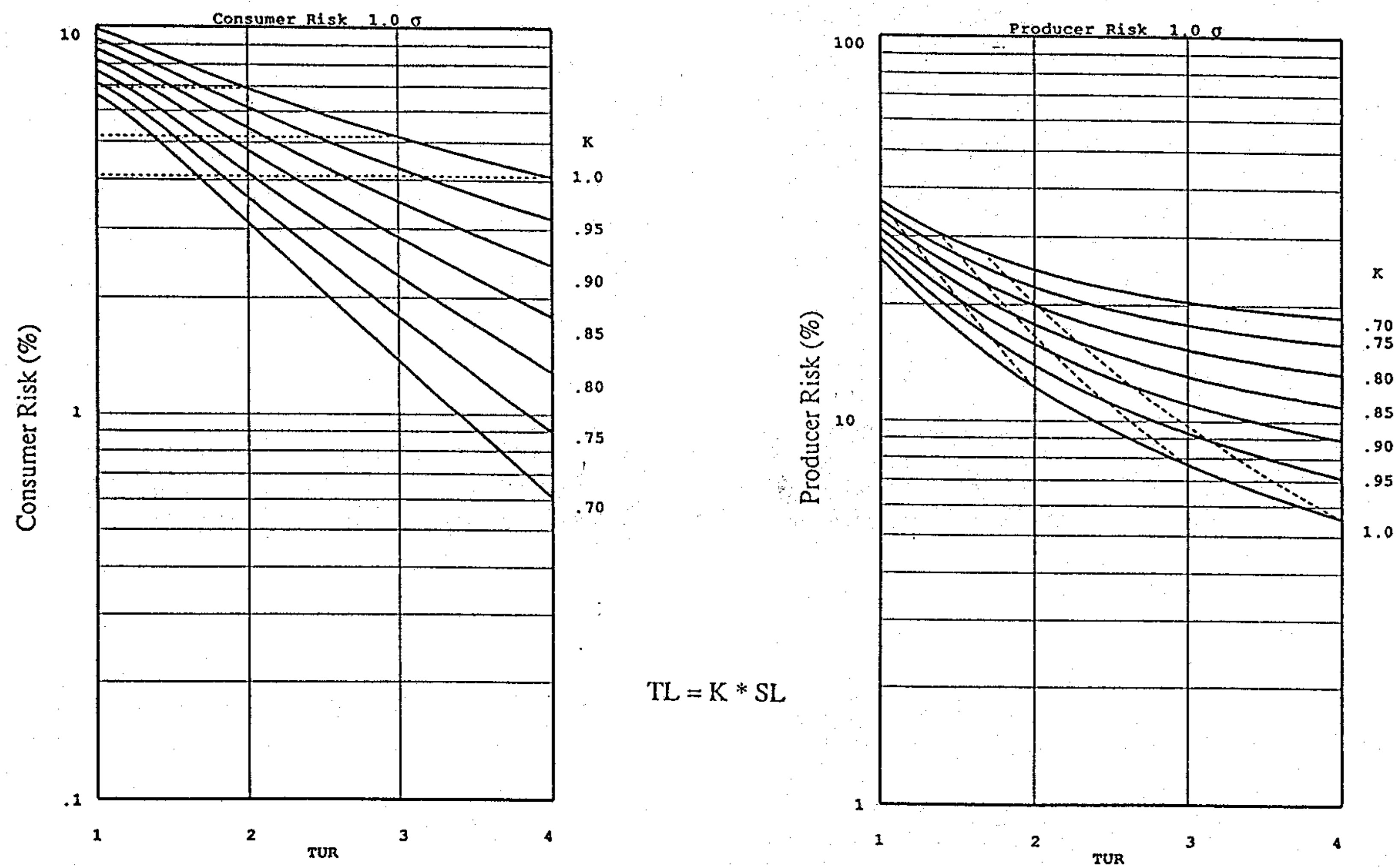
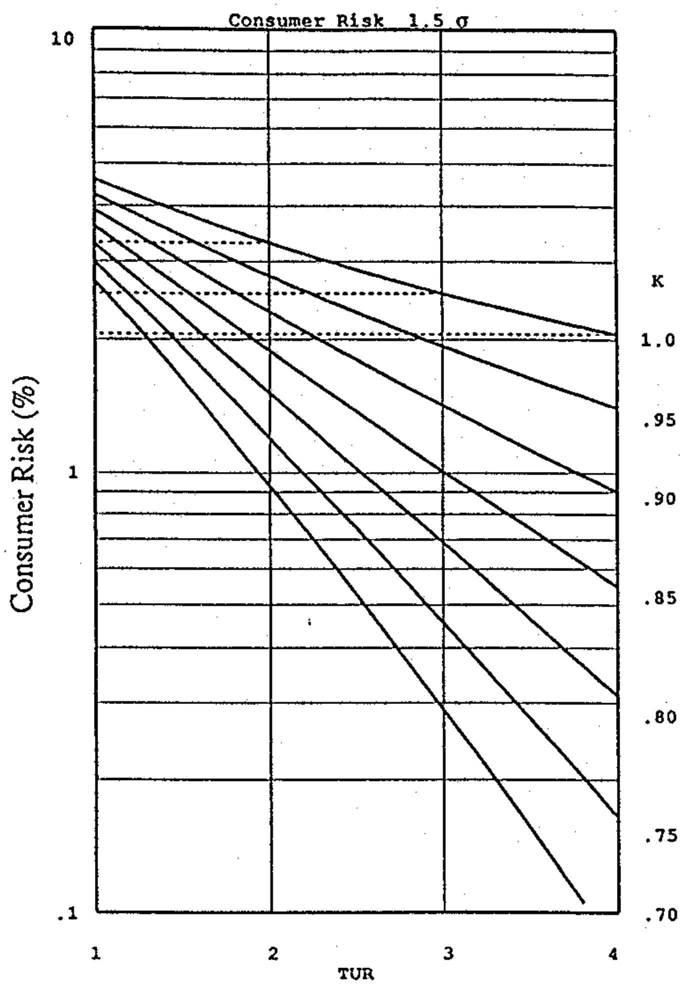
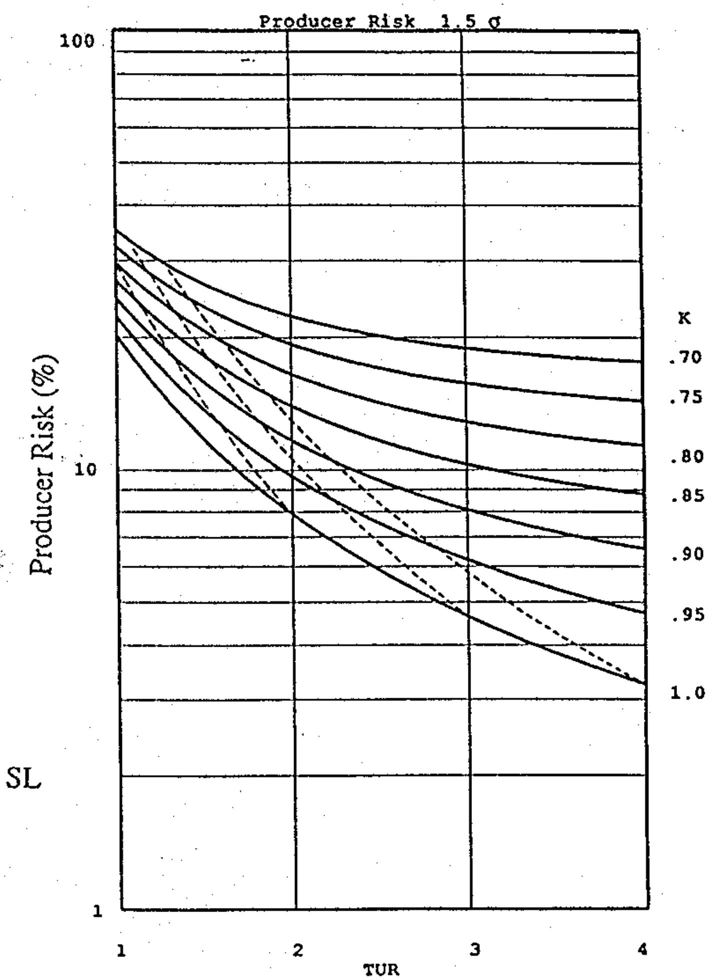


Fig. 12 Consumer Risk and Producer Risk for SL=σ





TL = K * SL

Fig. 13 Consumer Risk and Producer Risk for SL=1.5σ

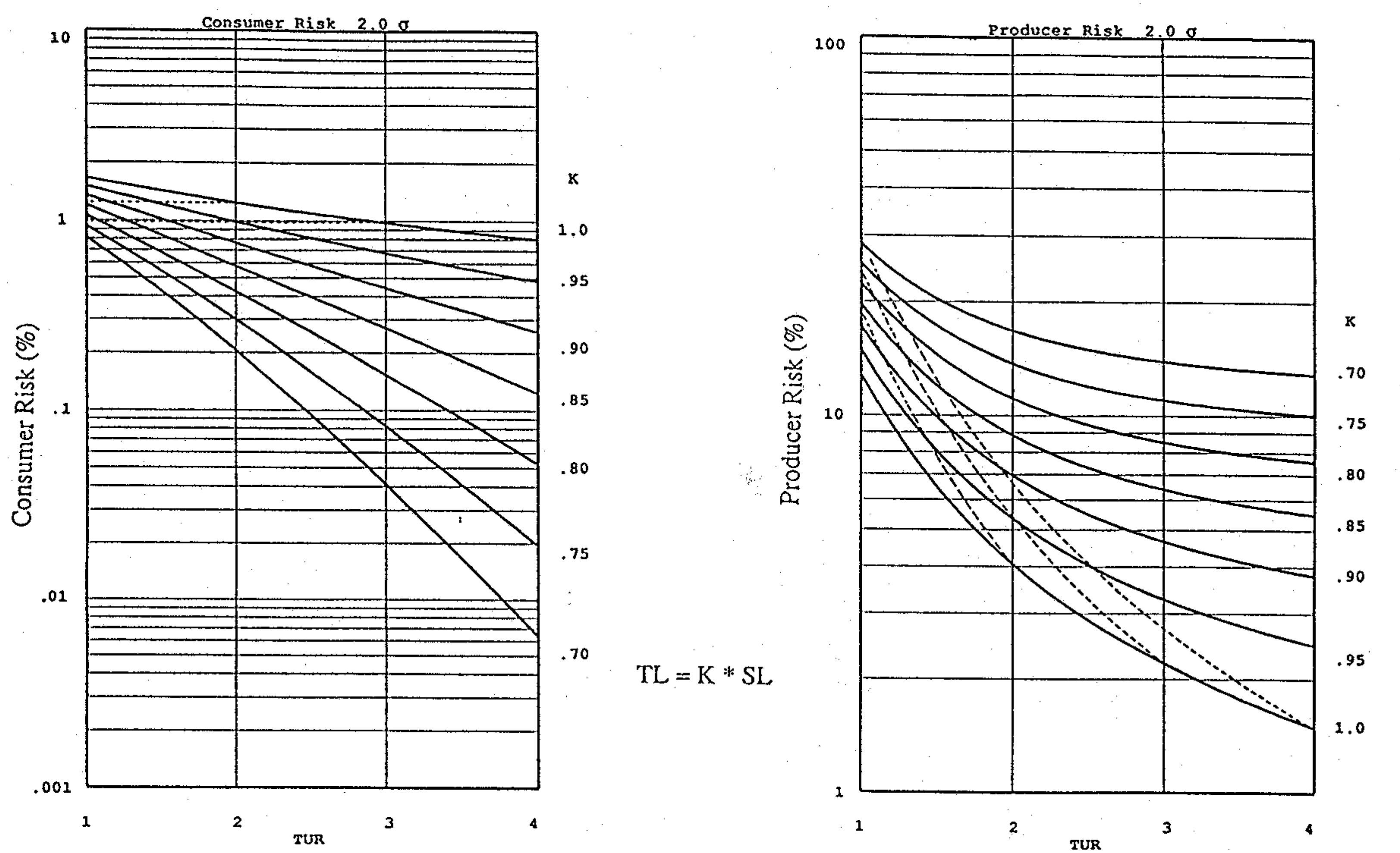


Fig. 14 Consumer Risk and Producer Risk for SL=2σ

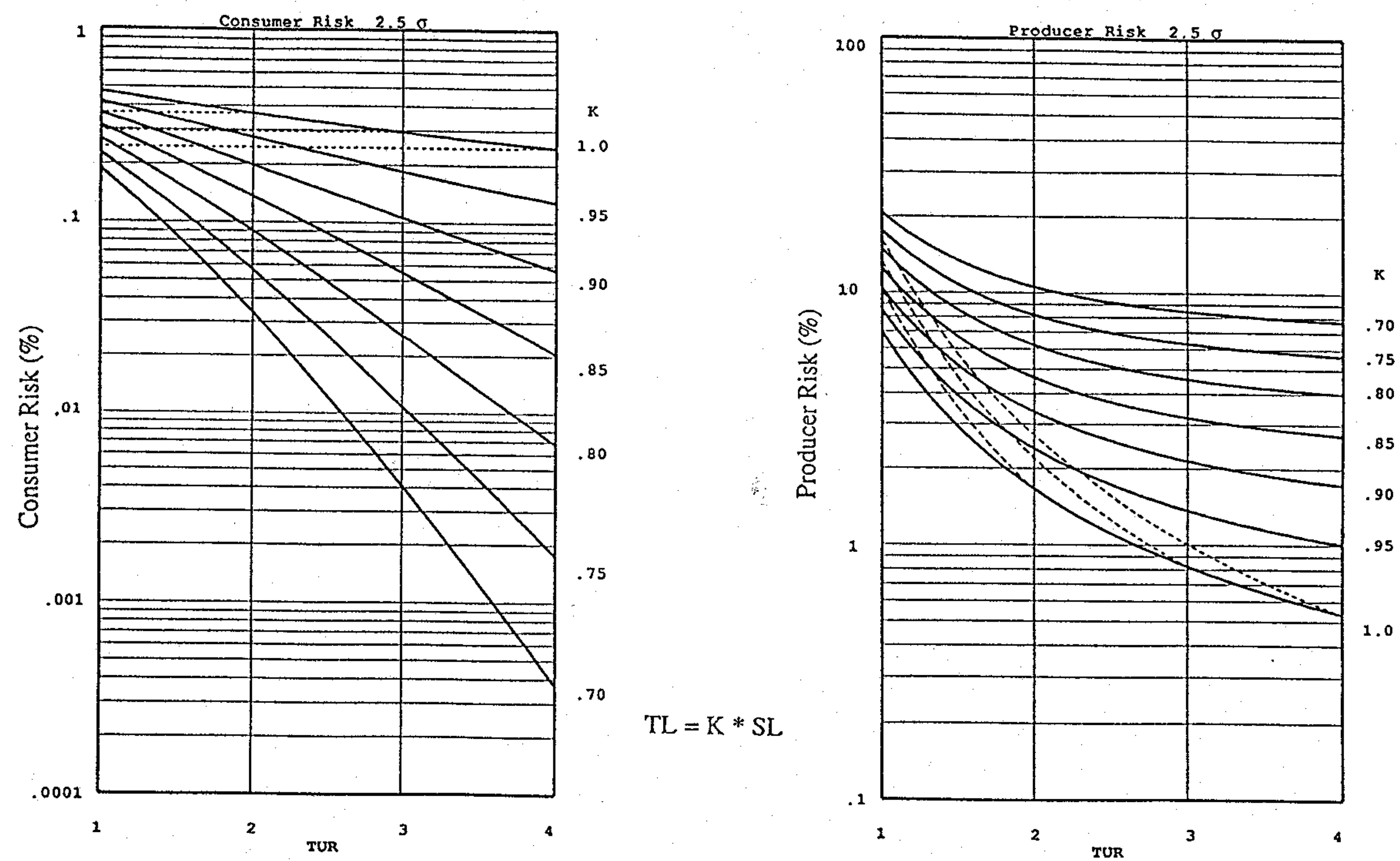


Fig. 15 Consumer Risk and Producer Risk for SL=2.5σ

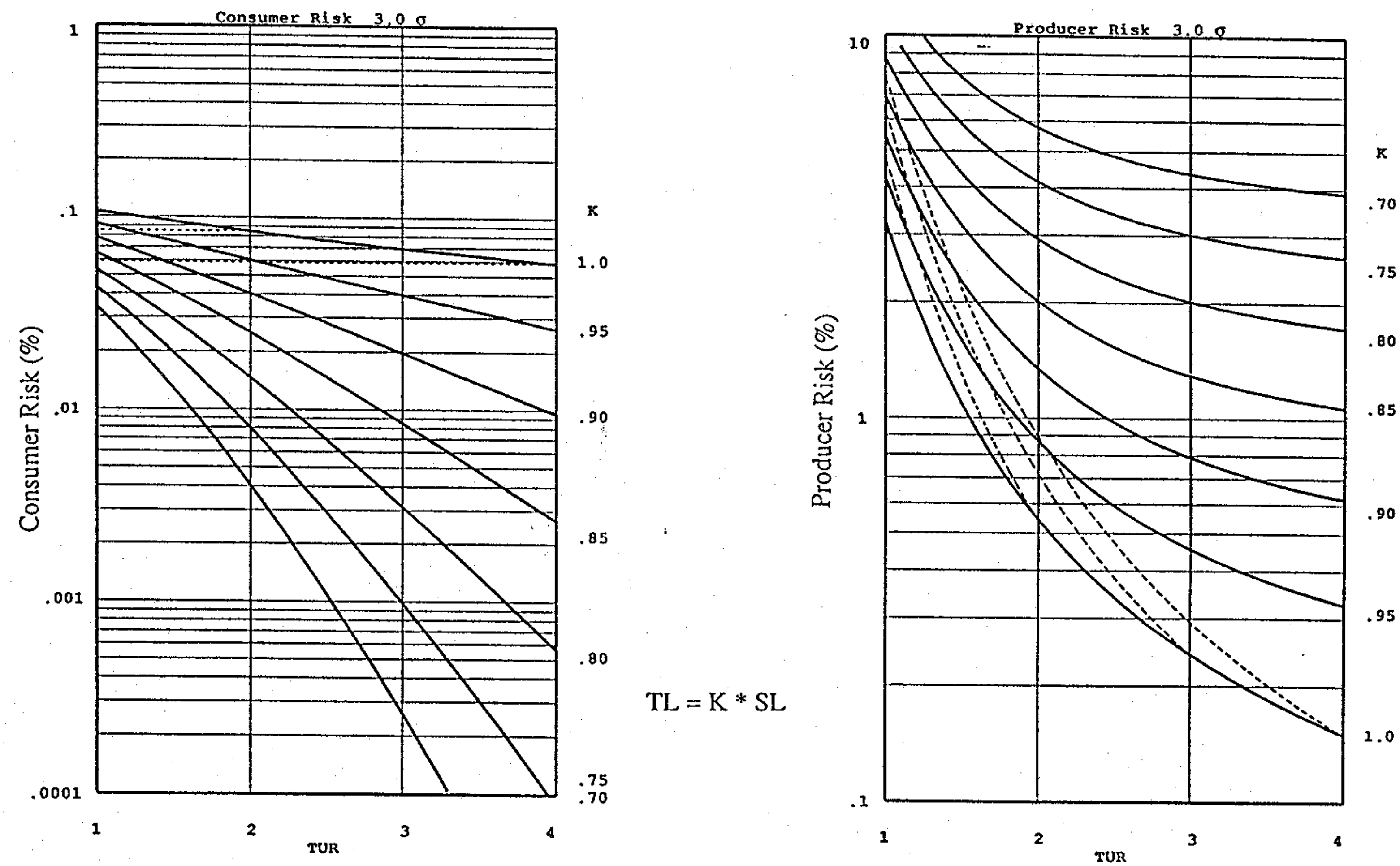


Fig. 16 Consumer Risk and Producer Risk for SL=3σ

APPENDIX B

EXAMPLES

EXAMPLE 1:

A Cal Lab needs to verify that a piece of test equipment is within its specification of 100 ppm. It would like to use a standard with a TUR of 4:1. However, the most accurate standard it has is specified at 50 ppm. If it is assumed both pieces of equipment are specified at an uncertainty of 2σ, this results in a TUR of only 2:1. Referring to Fig. 14 in Appendix A, it can be seen that the consumer risk for a 4:1 TUR is 0.8%. Moving left horizontally keeps the consumer risk constant at 0.8%. At a TUR of 2:1 we can interpolate to obtain a K of 0.91. Thus, setting a TL of 91 ppm will ensure that no more defective units will be accepted than with a 4:1 TUR. This result could be obtained from Fig. 11 as well.

EXAMPLE 2:

A Cal Lab purchases a 2nd standard of the same make and model as one already in service. The manufacturer claims that the SL=2σ which implies a 95.4% confidence that the instrument is within specifications. When the new instrument arrives, it is compared with the first instrument and the two instruments agree within the published specifications. Can it be claimed with a higher degree of confidence than 95.4% that the instrument being received is within its specifications?

Yes, If it assumed that the uncertainties are largely random (the systematic errors are small), referring to Fig. 8, the consumer risk for a TUR of 1:1 and SL=2σ is 1.7%, resulting in a confidence of 98.3% that the new instrument is within its specifications.

EXAMPLE 3:

A Cal Lab maintains a minimum TUR of 3:1. However a few points can only be checked with a 2:1 TUR. What TL should be used to guarantee the same Consumer Risk as for a 3:1 TUR?

Assume the specifications of the UUT and STD are to a 2σ confidence level. From Fig. 11, it can be seen that a TL=0.95*SL should be specified. From Fig. 14, Appendix A, the confidence level being maintained can be seen to be about 99% (100% - 1% consumer risk).

EXAMPLE 4

The goal of a manufacturer is to control its internal process to a 3σ level with respect to its published specification. At final audit, products are tested to 80% of specifications. 5% of the products are rejected at this point and have to be reworked. The test equipment is specified at SL=3σ and a TUR of 4:1 is maintained with respect to the presumed variation of the process. Is the manufacturer meeting its goal of a 3σ process?

No. If the process was, indeed, meeting the 3σ goal, we would expect 0.3% of the units to be defective (Fig. 2) and an additional 0.4% of the conforming units to be rejected (producer risk from Fig. 16, Appendix B, K=0.8, TUR=4), resulting in a reject rate of 0.7%. It would appear that the process is running much closer to a 2.5σ process which would result in 1.2% defective units plus the rejection of some conforming units. If the process were 2.5σ, the TUR would be 4.8:1 (since the standards were selected to provide a TUR of 4:1 for a 3σ process). Fig. 15, Appendix B shows the Producer Risk only to a TUR of 4:1 but visually extrapolating the curve yields an estimate of about 3.5%. This would indicate a total reject rate of around 4.7% which is near what the manufacturer is experiencing.

APPENDIX C

REPRESENTATIVE MathCAD EXAMPLES

MathCAD calculations for Fig. 3

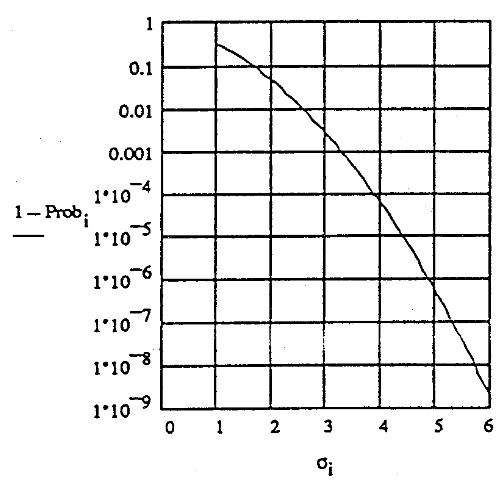
$N := 50$ Number of points in the plot

$i := 0..N$

$\sigma_i := 1 + 5 \cdot \frac{i}{N}$ Scales for sigma from 1 to 6

$TOL := .000001$ Reduces the calculation tolerance (necessary for a smooth plot at the low end)

$$Prob_i := \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_{-\sigma_i}^{\sigma_i} \exp\left[\frac{-(s^2)}{2}\right] ds$$



MathCAD calculations for Fig. 5

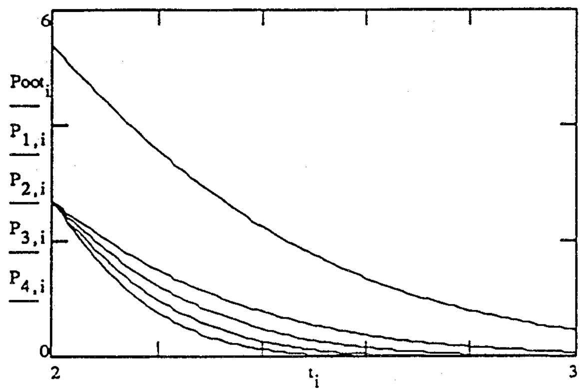
$R := 1..4$ Test Uncertainty Ratio

$SL := 2$ Specification limit in test standard deviations

$i := 0..100$

$t_i := SL + .01 \cdot i$ Scales t from the SL to the SL + 1 sigma

$$Poot_i := \frac{100}{\sqrt{2 \cdot \pi}} \cdot \exp\left[\frac{-(t_i)^2}{2}\right] \qquad P_{R,i} := \frac{100}{2 \cdot \pi} \cdot \exp\left[\frac{-(t_i)^2}{2}\right] \cdot \int_{-R \cdot (t_i + SL)}^{-R \cdot (t_i - SL)} \exp\left(\frac{-s^2}{2}\right) ds$$



MathCAD calculations for Consumer Risk, Fig. 14

$r := 0..60$ $R_r := 1 + .05 \cdot r$ R is scaled for TUR from 1 to 4

$i := 0..6$ $K_i := .70 + .05 \cdot i$ $SL := 2.0$ K is scaled from 70% to 100% of SL

$$P_{r,i} := \frac{1}{\pi} \cdot \int_{SL}^{10} \int_{-R_r \cdot (t + K_i \cdot SL)}^{-R_r \cdot (t - K_i \cdot SL)} \exp\left[\frac{-(s^2 + t^2)}{2}\right] ds\, dt$$

$y := 100 \cdot P$