How to Establish Manufacturing Specifications

Donald J. Wheeler

1. Introduction

The idea behind manufacturing specifications is to define those measurement values which correspond to conforming items. That is, when the measurement falls within the manufacturing specifications we want to be able to say that the product is within the customer specifications. In the absence of measurement error we could achieve this objective by simply using the customer specifications as the manufacturing specifications. But when we have to make allowance for measurement error there will need to be some gap between the manufacturing specifications and the customer specifications. How to determine the size of the gap, D, and the kind of statements that we can make about the product is the topic of this paper.

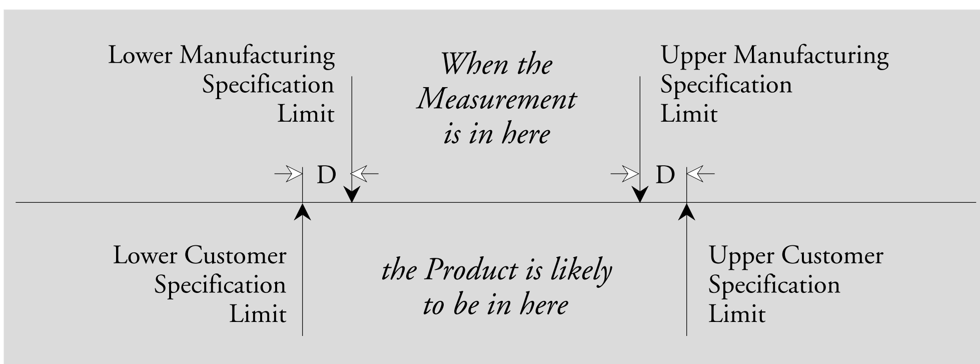

Figure 1: The Idea Behind Manufacturing Specifications

If we choose a large value for D we can be very confident that the product is conforming, but we could end up with very tight manufacturing specifications. When this happens there will be an increased risk of rejecting conforming product. By choosing a small value for D we will not have as much confidence that the product is conforming, but we will have looser manufacturing specifications with a correspondingly smaller risk of rejecting conforming product.

Thus the main problem of obtaining manufacturing specifications is the problem of striking a balance between the size of the adjustment, D, and the likelihood of having a conforming item. While the exact values will also depend on other aspects of the situation, it is these two properties that dominate the problem. In this paper I give a set of rules for constructing manufacturing specifications that have different likelihoods of conforming product, and then I show how these rules were developed.

2. Watershed Specifications

The concept of Watershed Specifications is fundamental to obtaining appropriate Manufacturing Specifications. No matter how we collect our data, at some level it will always become granular and there will be some increment in our recorded values. Traditionally specifications express this granularity by stating the minimum and the maximum acceptable values. While this is sufficient for routine usage, it can create a problem when calculations are performed using specification values. Since we are going to be making adjustments to the specifications, it is important to start off with the correct baseline values. To this end we shall have to make a correction for discrepancy between the granularity of our data and the continuity of the underlying scale.

While the minimum acceptable value is considered to be in-spec, a value that is one measurement increment below the minimum acceptable value will be out-of-spec. Thus the actual watershed point between an acceptable value and an unacceptable value is halfway in between these two values, and the Lower Watershed Specification Limit is:

$LWSL =$ minimum acceptable value – one-half of a measurement increment

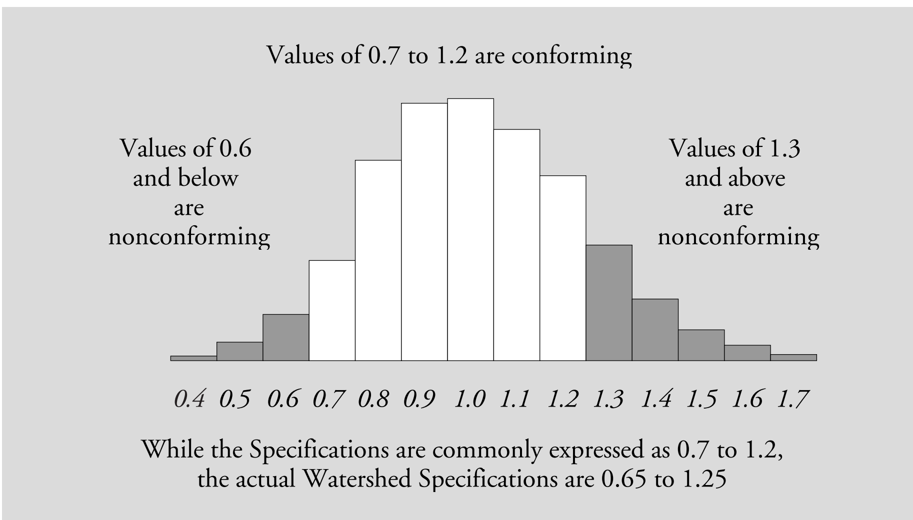

Figure 2: Watershed Specifications

Likewise, a value that is one measurement increment above the maximum acceptable value will also be out of spec. and the upper watershed point between an acceptable value and an unacceptable value is halfway in between these two values, making the Upper Watershed Specification Limit:

$UWSL =$ maximum acceptable value + one-half of a measurement increment

While we commonly express specifications by stating the minimum and maximum acceptable values, the actual specifications are the Watershed Specifications defined here. Their use is assumed throughout the computations and characterizations that follow.

3. Manufacturing Specifications

85% Manufacturing Specifications

If you want to have at least a 85 percent chance that the measured item is in spec when the measurement falls within the Manufacturing Specifications, then you will need to tighten up the Watershed Specification Limits by One Probable Error on each end, and use these tightened specs as your 85% Manufacturing Specifications. If we use the usual notation and let σ_e denote the standard deviation of repeated measurements of the same thing, then One Probable Error will be $0.675\,\sigma_e$. For two-sided specifications this adjustment for measurement error consumes $1.35\,\sigma_e$ of the difference between the Watershed Specification Limits, which shall be referred to in the following as the Watershed Tolerance.

96% Manufacturing Specifications

If you want to have at least a 96 percent chance that the measured item is in spec when the measurement falls within the Manufacturing Specifications, then you will need to tighten up the Watershed Specification Limits by Two Probable Errors, or $1.35\,\sigma_e$ on each end, and use these tightened specs as your 96% Manufacturing Specifications. For two-sided specifications this adjustment for measurement error consumes $2.70\,\sigma_e$ of the Watershed Tolerance.

99% Manufacturing Specifications

If you want at least a 99 percent chance that you are shipping good stuff, your 99% Manufacturing Specifications will be the Watershed Specification Limits tightened by Three Probable Errors, or $2.02\,\sigma_e$ on each end. For two-sided specifications this adjustment for measurement error consumes $4.05\,\sigma_e$ of the Watershed Tolerance. (Not the $5.15\,\sigma_e$ of popular fiction.)

99.9% Manufacturing Specifications

If you want at least a 99.9 percent chance that you are shipping good stuff, your 99.9% Manufacturing Specifications will be the Watershed Specification Limits tightened by Four Probable Errors, or $2.70\,\sigma_e$ on each end. For two-sided specifications this adjustment for measurement error consumes $5.40\,\sigma_e$ of the Watershed Tolerance. (Rather than the $6.0\,\sigma_e$ promoted by some.)

As we will see in the following sections these adjustments to specifications are not some arbitrary guidelines given without rationale or justification. They are the proper way to adjust the specifications to allow for the effects of measurement error.

In those cases where the customer specifications are so close together that your preferred manufacturing specifications are not at least one measurement increment apart, you will have to be content with manufacturing specifications with a lower likelihood. When the specifications are so tight that 85% Manufacturing Specifications cannot be used, you will be left with a substantial possibility that you are shipping nonconforming product. When the Watershed Specifications are used as the Manufacturing Specifications you will have a likelihood of conforming product that varies from approximately 64% at a capability of 0.10 to 83% at a capability of 2.0.

Another strategy for dealing with excessively tight manufacturing specifications is to use the average of multiple determinations. Such averages of n measurements will have a Probable Error that is smaller by the square root of n. This will result in slightly wider manufacturing specifications. While this may be expensive, it can be cheaper than a fine from a regulatory body, or an unhappy customer.

There are many different concepts tied up in these rules for obtaining manufacturing specifications. The following sections will explain each of these concepts in turn, and eventually will they will all be combined to develop a probability model that will justify the rules given here.

4. Predictable Measurement Systems

Until a measurement process is operated in a predictable manner, it cannot be said, in any logical sense, to measure anything at all. For this reason, the first thing to do when studying, or operating, any measurement system is to obtain repeated measurements of a designated standard, or a known standard, *and to place these measurements on a process behavior chart*. If this chart displays evidence of unpredictability then the measurement process is not operating, or is not being operated, in a consistent manner, and all further attempts to characterize this measurement process need to be postponed until the reasons for the inconsistent operation are found and removed.

When the process behavior chart for the repeated measurements of the standard displays a reasonable degree of predictability, the measurement process can be said to be consistent, and we can estimate the precision of the measurements produced.

Precision has to do with the degree that a measurement system can reproduce the same value when it is applied to the same object—a property that is variously known as repeatability, test-retest error, measurement error, measurement uncertainty, and replication error. Obviously, without a consistent measurement process the concept of precision will not be well-defined.

5. Quantifying Precision with the Estimated Standard Deviation

The National Institute for Standards and Technology recommends that the precision of a measurement system be reported using the estimated standard deviation of the measurement system. This number is readily found from the *XmR* chart used to establish and maintain the consistency of the measurement process—the average moving range, when divided by the bias correction factor of d_2 = 1.128, will result in an estimate of the standard deviation of the measurement system.

Figure 3 consists the *XmR* chart for 20 measurements of the same reference assembly. The average moving range for the repeated measurements of this reference assembly is 1.32 units and the precision of this gauge can be characterized by the value:

$$\hat{\sigma}_e = \frac{\bar{R}}{d_2} = 1.32 / 1.128 = 1.17 \text{ units}$$

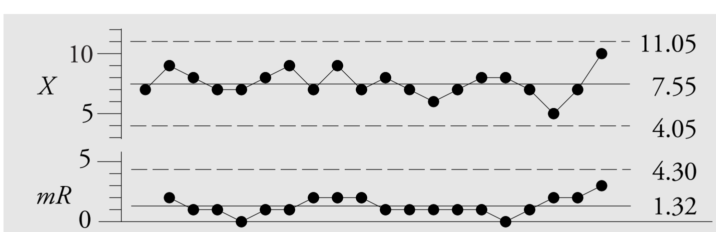

Figure 3: *XmR* Chart for the Stack Height Data

This estimate, like the limits for the *XmR* chart, is said to be based on 0.62 (k–1) ≈ 12 degrees of freedom.

An alternative way of estimating the precision of this gauge would be to compute the standard deviation statistic for the 20 repeated measurements of this reference assembly. For the data of Figure 3:

$$\hat{\sigma}_e = s = 0.8660 \text{ units}$$

This estimate is said to be based on (k–1) = 19 degrees of freedom. While these two numbers are different, they both tell the same story—the standard deviation for the measurement system is approximately equal to one measurement unit.

When a measurement system is operated consistently the two estimates illustrated above will be similar. That is, the estimate based on a global standard deviation statistic, s, will be similar to the range-based estimate. In addition, the s statistic will always have more degrees of freedom than the range-based estimate. While this might suggest that we should always use the global standard deviation statistic, this statistic becomes problematic when the measurement system is not consistent.

When a measurement system is not being operated consistently the estimate based on the global standard deviation statistic will be inflated by the irregular and unpredictable causes of inconsistent operation and thus will become meaningless. However, even when the measurement system is inconsistent, the range-based estimate will still provide an *approximation* of the *hypothetical* precision of the measurement system. Therefore, the only time that we can take advantage of the greater number of degrees of freedom for the global standard deviation statistic, s, is when the *XmR* chart shows that the measurement system has been operated predictably.

6. Quantifying Precision with the Probable Error

While the standard deviation of the measurement system is the yardstick for characterizing precision, there is another, older way to describe precision that is easier to explain to those without advanced degrees in physics or mathematics. It is the notion of Probable Error introduced by F. W. Bessel sometime prior to 1820, and defined as $0.675\,\sigma_e$. Probable Error was used throughout the Nineteenth Century in lieu of the standard deviation because it provided a powerful and simple way of characterizing measurement error.

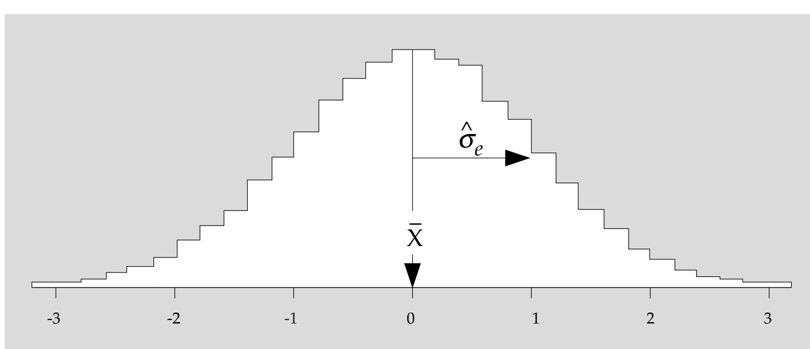

Figure 4: Histogram of Repeated Measurements of the Same Thing

To understand what the Probable Error tells us about the measurement process consider the results of measuring the same thing hundreds, or thousands, of times. If the measurements are not rounded off too much the result will be a histogram like that shown in Figure 4. It is intuitive and natural to use the average for this histogram as the "Estimate" for the value of the item being measured. And, as was noted in the previous section, the dispersion of this histogram will provide an estimate of σ_e .

Ever since 1810, when Laplace published what has come to be called the central limit theorem, the normal probability model has been the primary model for measurement error. Therefore, we can use the normal probability model to approximate the behavior of repeated measurements of the same thing. In particular, since the middle 50 percent of the normal model is defined by the interval:

$$\mu \pm 0.675\,\sigma$$

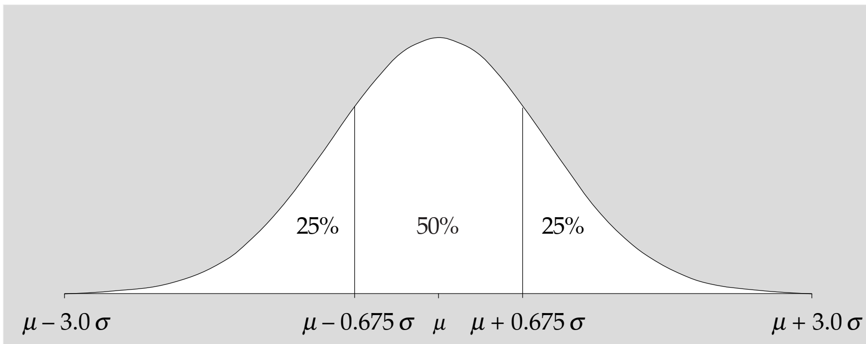

Figure 5: Some Areas Under a Normal Curve

we can say that approximately half of the repeated measurements of the same thing should fall within the interval:

$$\bar{X} \pm 0.675\,\hat{\sigma}_e \quad \text{or} \quad \text{Estimate} \pm \text{One Probable Error}$$

The final step in justifying the use of Probable Error consists of thinking about the "error" of a single measurement. Without knowing the value of the Estimate, we can think about the difference between a single measurement and the Estimate as the error of our measurement. This difference is shown in Figure 6. Comparing Figures 5 and 6, we can say that the error of a single measurement will be less than one Probable Error half the time, and it will exceed one Probable Error half the time.

Thus, the Probable Error is the median amount by which a single measurement will err— it is, in effect, the average error of a single measurement.

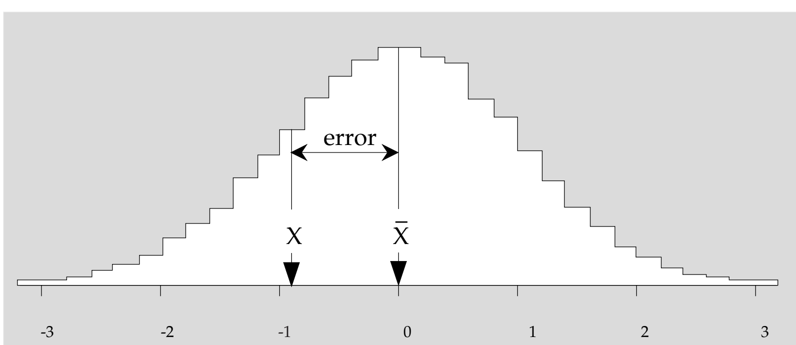

Figure 6: The Error of a Single Measurement

The Probable Error is a very useful quantity. In addition to being the basic unit used in creating the Manufacturing Specifications, it also defines the effective resolution of a single measurement, and it determines how many digits are appropriate for recording a measurement. But it is the fact that it represents the median error of a measurement that makes the Probable Error a more appealing way of characterizing the precision of a measurement system. And since the Probable Error is a multiple of the standard deviation of the measurement system, the use of Probable Error is equivalent to the recommended practice of reporting measurement error by stating the value for the estimate of σ_e.

For the data of Figure 3, the Probable Error would be estimated to be either:

$$PE = 0.675\ \hat{\sigma}_e = 0.675\,(1.17) = 0.8 \text{ units} \qquad \text{or} \qquad PE = 0.675\,(0.866) = 0.6 \text{ units}$$

while the data themselves were recorded to the nearest whole unit (1.0 units).

7. The Effective Resolution of a Measurement

As we can see in the previous example, there are two things that limit the precision of a measurement. The first of these is the round-off operation that is carried out when the measurement is written down. This round-off operation will define the smallest increment in the measurements, which will define the “apparent resolution” of the measurements. In many cases the number of digits recorded will depend upon some superstitious tradition. Recording more digits is thought to be excessive, and recording fewer digits is thought to be careless.

But as was shown in the previous section, the uncertainty in the measurement process also places a limit on the precision of a measurement. This uncertainty, or fuzz, inherent in the measurement system is defined by the Probable Error. Since the Probable Error is the median error of a measurement, you know that half the time your measurement will differ from the Estimate by more than one Probable Error. For this reason there is no point in attempting to interpret any value more precisely than plus or minus one Probable Error.

The larger of these two limiting factors will define the effective resolution of a measurement.

When the Probable Error exceeds the smallest measurement increment
it is the Probable Error that will define the effective resolution of the measurement.

When the measurement increment exceeds the Probable Error
it is the measurement increment that will define the effective resolution.

In the latter case the extra round-off involved in using measurement increments that are larger than the Probable Error will add to the uncertainty of the measurements and will increase the error of the measurements. Figure 7 shows how the size of the measurement increment will affect the median error of a measurement.

Inspection of Figure 7 will show that as long as the measurement increment is less than one Probable Error, there is virtually no inflation in the error of a measurement. However, as the measurement increment exceeds the Probable Error, the inflation begins to increase in a non-linear manner. This non-linear relationship leads to a preference that the measurement increment should be approximately the same size as the Probable Error. When the measurement increment is greater than 2.0 PE you will be loosing information due to the round-off. When it is less than 0.2 PE you will be writing down meaningless digits.

Thus, the Probable Error provides us with an objective way to determine how many digits to record for any measurement. When the number of digits recorded is discretionary, we should seek to have a measurement increment that satisfies the following:

$$0.2 \text{ Probable Error} < \text{measurement increment} < 2 \text{ Probable Errors}$$

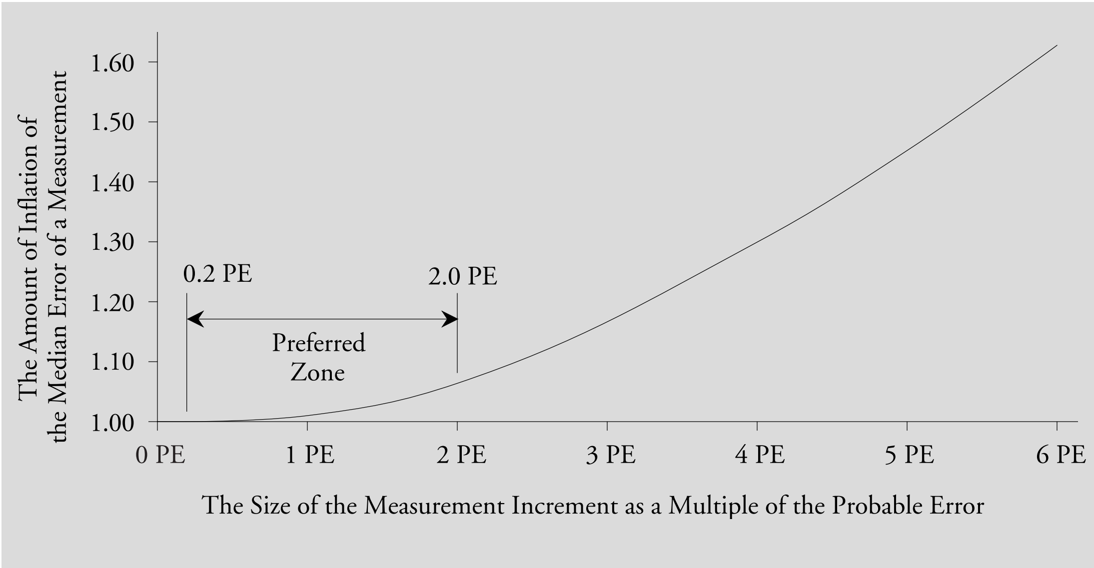

Figure 7: Inflation of Measurement Error as a Function of the Measurement Increment

When you automatically get too many digits, you can use the preferred zone above to determine how many digits are really useful and delete the remainder. When you automatically get too few digits, you do not have to worry about measurement error being intrusive. While such measurements could be better than they are, the effective resolution will be the same as the apparent resolution, and the measurements will be good to the last digit.

For the data of Figure 3, with a Probable Error in the neighborhood of 0.6 to 0.8 units, the measurement increment of 1.0 unit is in the right zone. This means that the data are being recorded to the proper number of digits, and they are basically good to the last digit. To record these values to 0.1 unit would be excessive, and to round them off to the nearest multiple of 10 units would degrade the measurements.

For another example consider the duplicate viscosity values for Product 10F shown in Figure 8.

Lot Number	32	33	34	35	36	37	38
Duplicate	20,480	19,370	20,350	19,870	20,360	19,320	20,580
Viscosities	20,430	19,230	20,390	19,930	20,340	19,300	20,680
Ranges	50	140	40	60	20	20	100

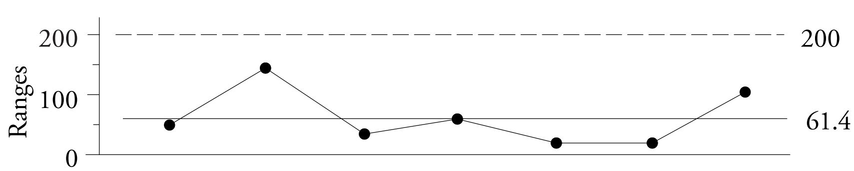

Figure 8: Range Chart for Consistency of Viscosity Measurements.

Each Lot is sampled, the sample is split, and the viscosity is determined twice. The average range is 61.4 centistokes, giving us an estimate for the standard deviation of the measurement process of:

$$\hat{\sigma}_e = 61.4 \text{ cs}/1.128 = 54.4 \text{ centistokes}$$

Thus the Probable Error for a single measurement is:

$$PE = 0.675\,(54.4\text{ cs}) = 37\text{ cs}$$

Based on this estimate of the Probable Error we should aim to have a measurement increment somewhere in the range of 7 cs to 70 cs. In this case the individual measurements are made to the nearest 10 centistokes, which is in this preferred zone. It would be a mistake to record these viscosities to the nearest centistoke, and it would be a mistake to round them off to the nearest 100 centistokes. While this estimated Probable Error has only $0.9\,k(n{-}1) = 6.3$ degrees of freedom, it is sufficient to determine that the data are being recorded to the proper number of digits. The order of magnitude spread on the recommended zone makes it possible to use even soft estimates like this to make useful judgments about the measurement process.

8. How to Verify the Curve of Figure 7

To verify the curve shown in Figure 7 for yourself, without getting bogged down in the mathematics, you can use a spreadsheet program to carry out a simulation study in the following manner.

- Let Y denote the product values. Generate a column of N random values for Y and then fix them so they will not change with subsequent operations. These values can have any distribution, but a normal distribution is commonly used. (You should use $N \geq 1000$ values.)
- Let E denote the measurement errors. Generate a column of N random values for E and then fix them so that they will not change in subsequent operations. These values should be normally distributed with a mean of 0.0. They may have any standard deviation you choose, but it will need to be a known value. (To avoid situations where measurement error dominates the product variation, the standard deviation for E should not exceed half of that for Y.) Denote this known value by σ_e. For later use you will need to compute the median of the absolute values of these N values for E. (This value should end up being close to $0.675\,\sigma_e$.)
- Let X denote the product measurements. The N values for X will be the sum of the corresponding values for Y and E. That is, $X = Y + E$. To simulate the chunkiness of real data you will need to modify these X values. Since the purpose here is to examine the effect of the measurement increment upon the discrepancy between X and Y, you will need to pick a measurement increment. Define this measurement increment to be:

  $$MI = C \,*\, 0.675 * \sigma_e.$$

  For different values of C you can then perform the following operations.
- Round off the X values to the nearest measurement increment by
  - (1) dividing X by MI,
  - (2) ROUND this result to 0 decimals, and
  - (3) multiply the result by MI.
- Find the absolute value of the difference between each rounded value for X and the corresponding value for Y.
- Find the median of these N absolute values. Divide this by the median of the absolute values of E. This ratio shows the amount by which the measurement increment has inflated the median error of a single measurement.

Repeating the last three bullets for different values of C will result in a curve similar to the one shown in Figure 7.

9. The Bivariate Normal Model

To justify the guidelines for manufacturing specifications given earlier we must begin with the model given in the previous section: let Y denote the product value, let E denote the measurement error, and let X denote the product measurement, where $X = Y + E$. We begin by assuming that the product values, Y, may be characterized by a normal distribution having a mean denoted by μ_y and a standard deviation denoted by σ_y.

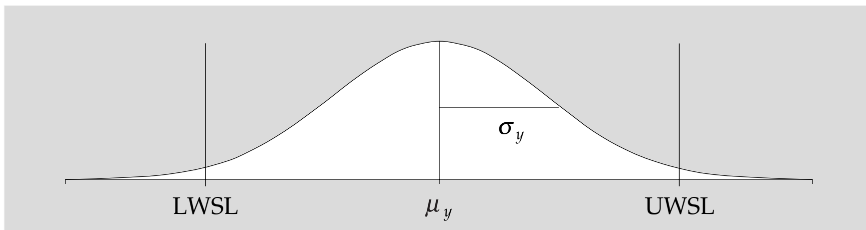

Figure 9: Assumed Distribution and Watershed Specifications for Product Values, ***Y***

The lines denoted by *LWSL* and *UWSL* represent the Watershed Specifications.

There will always be some amount of uncertainty attached to each measurement. For this reason the distribution of the recorded measurements, X, will not be exactly the same as the distribution of the product values, Y, shown in Figure 9. Let μ_x denote the mean of the product measurements, X, and let σ_x denote the standard deviation of these product measurements. Since the classic model for measurement error, E, is a normal distribution with a mean of zero and a variance of σ_e^2, the distribution of X will therefore be a normal distribution with:

$$\mu_x = \mu_y$$

$$\sigma_x^2 = \sigma_y^2 + \sigma_e^2$$

Furthermore, let us assume that X and Y follow a bivariate normal distribution with correlation coefficient ρ where:

$$\rho^2 = 1 - \frac{\sigma_e^2}{\sigma_x^2} = \frac{\sigma_y^2}{\sigma_x^2}$$

Given this formulation of the problem, the relationship between the product values, Y, and the product measurements, X, can be represented by the plot in Figure 10. As may be seen there, a specific product value, $Y = y$, will give rise to a distribution of product measurements, $f(x \mid y)$. This distribution is said to be conditioned upon the value of y, and this condition is shown symbolically by the vertical line followed by y.

This conditional distribution of X, given that $Y = y$, is shown in Figure 10. It is the distribution of repeated measurements of the same thing seen in Figure 4, and it is modeled by a normal distribution with mean:

$$\mu_{x\,|\,y} \quad = \quad \mu_x \;+\; \rho\,\frac{\sigma_x}{\sigma_y}\,(\,y - \mu_y\,) \quad = \quad y$$

and variance :

$$\sigma^2_{x\,|\,y} \quad = \quad \sigma_x^{\,2}\,(\,1 - \rho^2\,) \quad = \quad \sigma_e^{\,2}$$

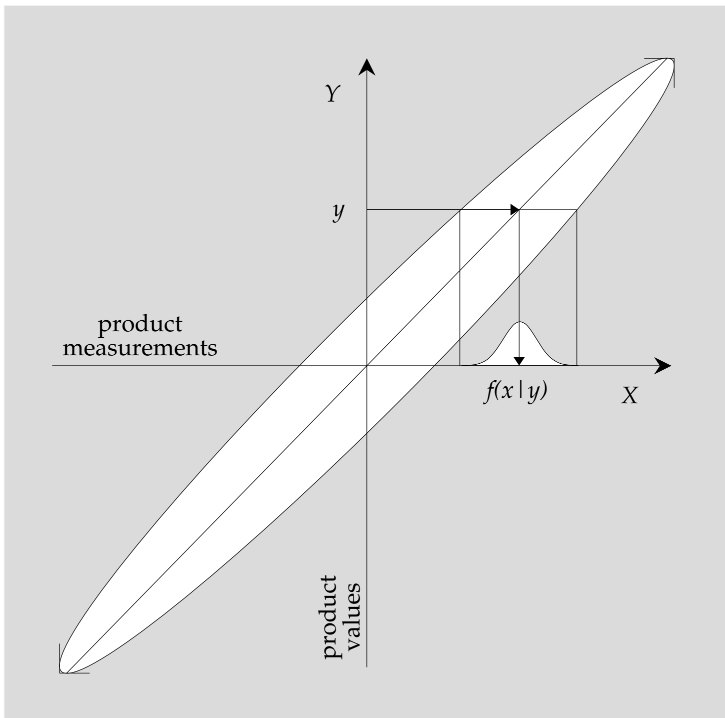

Figure 10: The Relationship Between Product Values and Product Measurements

With this formulation of the problem, the issue of characterizing product relative to specifications becomes one of determining the inverse of the relationship shown in Figure 10. Any one specific measurement, $X = x$, corresponds to a range of product values, $y_1 < Y < y_2$. When this range of values for Y falls essentially inside, or essentially outside the specifications, there is little ambiguity to the measurement. However, when this range of values for Y includes one of the specification limits, there may be considerable ambiguity to the measurement, x. In order to characterize the nature of this ambiguity we shall have to consider the conditional distribution of Y given $X = x$. (Since the values of Y give rise to the values of X, this conditional distribution, $f(y\,|\,x)$, is sometimes known as the inverse probability distribution, or the *a posteriori* distribution.)

This conditional distribution is shown on the vertical axis in Figure 11. It is a normal distribution with mean:

$$\mu_{y\,|\,x} \;=\; \mu_y \;+\; \rho\,\frac{\sigma_y}{\sigma_x}\,(\,x-\mu_x\,) \;=\; \mu_x \;+\; \rho^2\,(\,x-\mu_x\,) \;=\; \rho^2 x \;+\; (\,1-\rho^2)\,\mu_x$$

and variance :

$$\sigma^2_{y\,|\,x} \;=\; \sigma_y{}^2\,(\,1-\rho^2\,) \;=\; \frac{\sigma_y{}^2\;\sigma_e{}^2}{\sigma_x{}^2} \;=\; \rho^2\,\sigma_e{}^2$$

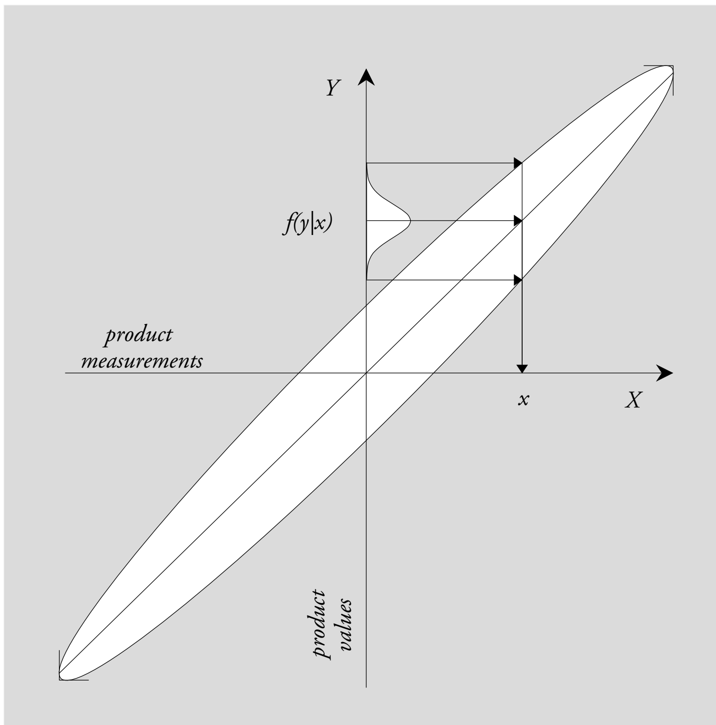

Figure 11: Different Product Values Result in the Same Measurement Value

Therefore, the likelihood that an item is conforming, given a specific measurement $X = x$, would be found by obtaining the integral of $f(y\,|\,x)$ evaluated between the specification limits:

$$P[\;Y \text{ is conforming } \mid X = x\;] \;=\; \int_{LWSL}^{UWSL} f(y\,|\,x)\,dy$$

where $f(y\,|\,x)$ is the normal *a posteriori* distribution described above. This integral may be reasonably approximated using a Burr distribution, which allows the finite integral above to be expressed in closed form as:

$$P[\;Y \text{ is conforming } \mid X = x\;] \;\approx\; \left\{1 + [\;0.644717 + 0.161990\;Z_L\;]^{4.873717}\right\}^{-6.157568}$$

$$-\;\left\{1 + [\;0.644717 + 0.161990\;Z_U\;]^{4.873717}\right\}^{-6.157568}$$

where:

$$Z_L = \text{Max}\left\{ \frac{\text{LWSL} - \mu_{y|x}}{\sigma_{y|x}} \quad \text{and} \quad -3.97 \right\}$$

$$\text{and} \quad Z_U = \frac{\text{UWSL} - \mu_{y|x}}{\sigma_{y|x}}$$

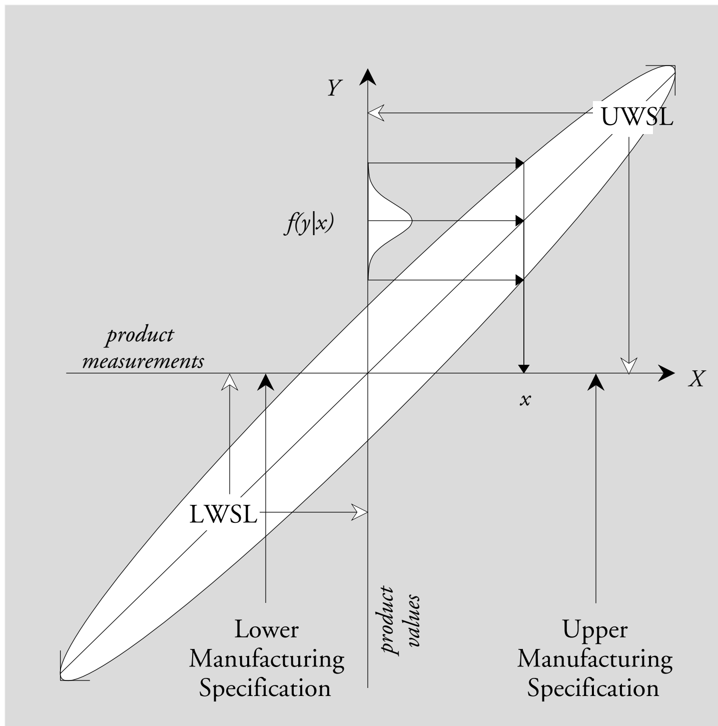

Figure 12: Manufacturing Specifications and Customer Specifications

Figure 12 shows what happens for the important case where the capability is less than 1.0. If $X = x$ is equal to the Upper Watershed Specification Limit value we should expect to have approximately a 50-50 chance of conforming product. However, as X moves down from the $UWSL$ value the probability of conforming product will increase. By evaluating the integral given above for different values of X we will usually find a value for X that will result in a probability that is large enough to make us comfortable in shipping the product. This value will be our Upper Manufacturing Specification. The same argument applied to the lower end will result in a Lower Manufacturing Specification.

10. The Probability of Conforming Product

To evaluate the probability integral we can, without loss of generality, assume that the product values, Y, have a mean of 0.0 and a variance of 1.0. This will simplify the conditional mean and variance of Y given X = x into:

$$\mu_{y|x} = \mu_x + \rho^2 (x - \mu_x) = \rho^2 x \qquad \text{and} \qquad \sigma^2_{y|x} = \sigma_y^2 (1 - \rho^2) = (1 - \rho^2)$$

As a result the values for Z_L and Z_U will depend solely upon (1) the Intraclass Correlation Coefficient, ρ^2, and (2) the distance from the watershed specification limit to the value $X = x$.

To obtain realistic values for the probability integral we will need to take the granularity of the measurements into account. Since the Probable Error defines the essential granularity for a measurement we will assume that the measurement increment is equal to the Probable Error. With the variance for Y fixed at 1.0, the Probable Error becomes a function of the Intraclass Correlation Coefficient, ρ^2:

$$PE = 0.675\sqrt{\frac{1-\rho^2}{\rho^2}} = \text{Measurement Increment}$$

Furthermore, we shall also assume that the Watershed Specification Limits must fall at the midpoint between two possible values. Given the common way that specifications are expressed, this will correctly describe most situations. If we let a denote the maximum acceptable value, then the possible values around the Upper Watershed Specification would be those shown in Figure 13.

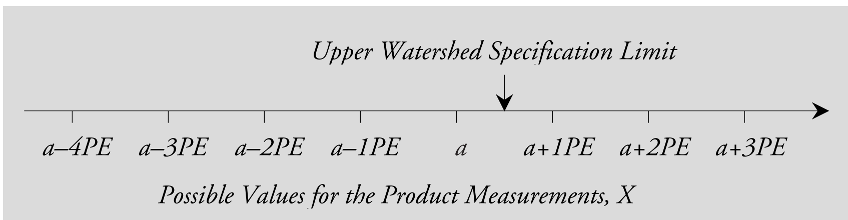

Figure 13: The Relationship Between the *X* Values and the Upper Watershed Specification

If a product measurement should happen to take on the maximum acceptable value, it would, by definition, be one-half measurement increment away from the Upper Watershed Specification Limit. If a product measurement should happen to fall one unit below the maximum acceptable value, it would be 1.5 measurement increments away from the Upper Watershed Specification. The value $a-2PE$ is 2.5 PE away from the spec. All of the values in Figure 13 are some odd multiple of one-half PE away from the spec. Thus, by considering the granularity of our data, we only have a finite number of values for X that we will need to consider.

The first evaluation was for the case of X = the maximum acceptable value, a. Twenty different values of the Intraclass Correlation Coefficient, ranging from 0.10 to 0.99 were used, along with 13 different values for the Capability Index ranging from 0.10 to 2.00.

$$P[\,Y \text{ is conforming} \mid X = a\,] = \int_{LWSL}^{UWSL} f(y \mid a)\, dy$$

$$\approx \left\{ 1 + [\,0.644717 + 0.161990\, Z_L\,]^{4.873717} \right\}^{-6.157568} - \left\{ 1 + [\,0.644717 + 0.161990\, Z_U\,]^{4.873717} \right\}^{-6.157568}$$

where:

$$Z_L = \text{Max}\left\{ \frac{LWSL - \rho^2 a}{\sqrt{1-\rho^2}} \text{ and } -3.97 \right\}$$

$$\text{and } Z_U = \frac{UWSL - \rho^2 a}{\sqrt{1-\rho^2}}$$

We substitute [a + 0.5 PE] for UWSL, and, assuming the specs are symmetric around the mean of zero, we can substitute [– a – 0.5 PE] for LWSL. These substitutions turn Z_L and Z_U into quantities that can be evaluated for a given value of the Intraclass Correlation Coefficient, ρ^2.

The results of these computations are summarized in Figure 14. There the minimum probability of a conforming item is shown as a function of the capability index. This graph shows that, without adjustment, using the Watershed Specifications as the Manufacturing Specifications will result in anywhere from a 64% chance of conforming product to an 83% chance as the capability varies from 0.10 to 2.00.

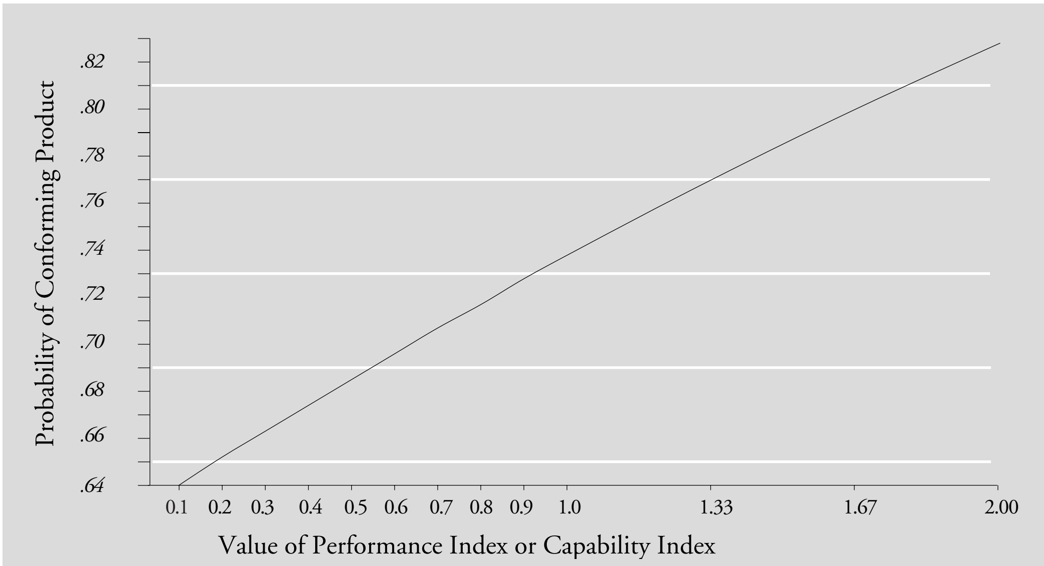

Figure 14: The Minimum Probability of Conforming Product When *X* = Maximum Acceptable Value

The use of capability values here may seem strange since we are interested in the problem of characterizing an item as conforming or nonconforming. In evaluating the probability integral the capability values were used to show the effects of wider or tighter customer specifications. If you actually make use of Figure 14 to characterize the likelihood of conforming product, the width of the specifications can be indicated using any capability type of computation that is appropriate for your situation, e.g. either the capability index, the centered capability index, the performance index, or the centered performance index.

The idea in creating the graph in Figure 14 was to characterize the probability of conforming product when the measurement was at the maximum acceptable value. Specifically I was concerned with the probability of the item being below the Upper Watershed Specification Limit, therefore I did not include the probabilities for those cases where the probability integral was limited by *both* the upper and the lower spec. If a measurement at the maximum acceptable value has a non-zero probability of being below the lower spec, you have a very severe problem. These cases only occurred with a capability value of 0.10 and an Intraclass Correlation below 94%.

Thus, Figure 14 serves to justify the statement that the Watershed Specification Limits themselves provide 64% Manufacturing Specifications.

The second evaluation of the probability integral was for the case where X is one measurement increment below the maximum acceptable value, $X = a - 1PE$. This case represents what would happen if the Watershed Specifications were tightened by one Probable Error and the measurement was just inside these tightened specifications. Once again twenty different values of the Intraclass Correlation Coefficient, ranging from 0.10 to 0.99 were used, along with 13 different values for the Capability Index. The minimum probabilities of conforming product for each different capability is shown in Figure 16.

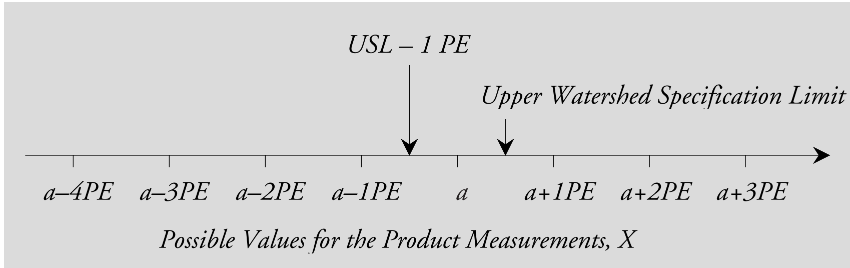

Figure 15: Watershed Specification Tightened by One Probable Error

The probabilities in Figure 16 range from 85% to 95% depending upon the capability. Thus, Figure 16 justifies the rule for obtaining 85% Manufacturing Specifications: If you want to have at least a 85 percent chance that the measured item is in spec when the measurement falls within the Manufacturing Specifications, then you will need to tighten up the Watershed Specification Limits by one Probable Error, or $0.675\,\sigma_e$ on each end, and use these tightened specs as your 85% Manufacturing Specifications. For two-sided specifications this adjustment for measurement error consumes $1.35\,\sigma_e$ of the Watershed Tolerance.

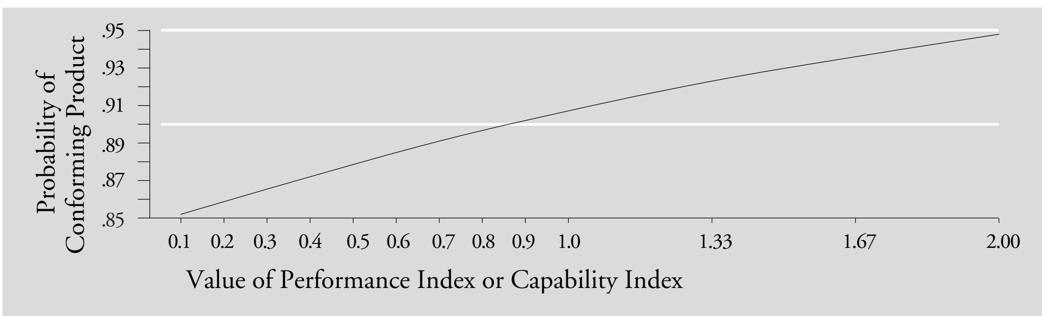

Figure 16: The Probabilities of Conforming Product When Specs are Tightened by One PE

To represent the case where the Watershed Specifications have been tightened by two Probable Errors the probability integral was evaluated at $X = a - 2PE$. Here the minimum probabilities ranged from 0.958 to 0.989. These values appear as the bottom curve in Figure 17. Thus, if you want to have at least a 96 percent chance that the measured item is in spec when the measurement falls within the Manufacturing Specifications, then you will need to tighten up the watershed specification limits by two Probable Errors, or $1.35\,\sigma_e$ on each end, and use these tightened specs as your 96% Manufacturing Specifications. For two-sided specifications this adjustment for measurement error consumes $2.70\,\sigma_e$ of the Watershed Tolerance.

To represent the case where the watershed specifications have been tightened by three Probable Errors the probability integral was evaluated at $X = a - 3PE$. Here the minimum probabilities ranged from 0.991 to 0.998. These values appear as the middle curve in Figure 17. Thus, if you want at least a 99 percent chance that you are shipping good stuff, your 99% Manufacturing Specifications will be the Watershed Specification Limits tightened by three Probable Errors, or $2.02\,\sigma_e$ on each end. For two-sided specifications this adjustment for measurement error consumes $4.05\,\sigma_e$ of the Watershed Tolerance.

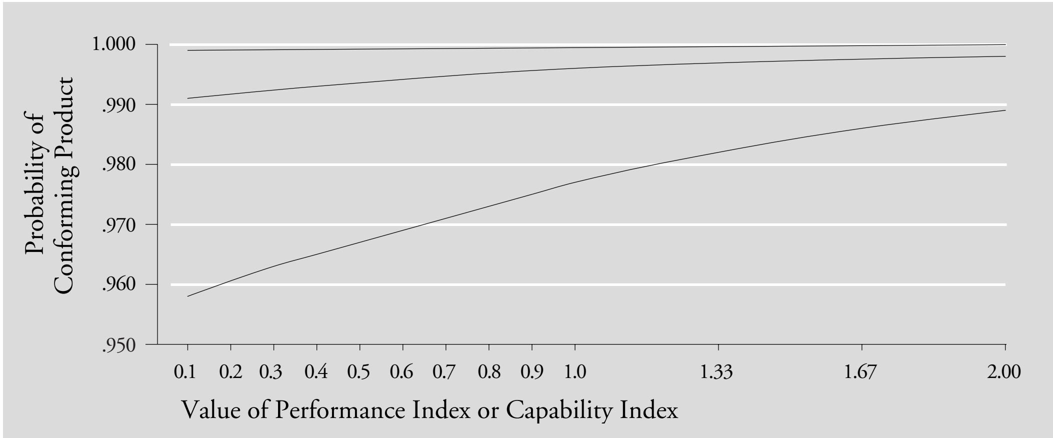

Figure 17: The Probabilities of Conforming Product When Specs are Tightened by 2, 3, and 4 PE

And finally to represent the case where the watershed specifications have been tightened by four Probable Errors the probability integral was evaluated at $X = a - 4PE$. Here there was one case with a probability of 0.998 and 174 cases where the probability was either 0.999 or 1.000. These values appear as the top curve in Figure 17. Thus, if you want at least a 99.9 percent chance that you are shipping good stuff, your 99.9% Manufacturing Specifications will be the Watershed Specification Limits tightened by four Probable Errors, or 2.70 σ_e, on each end. For two-sided specifications this adjustment for measurement error consumes 5.40 σ_e of the Watershed Tolerance.

All mathematical models have limitations, and the bivariate normal model used here is no exception. While this model is appropriate for use when the specifications are within, or just outside, the Natural Process Limits, it is not a satisfactory model for use with large capability values. So while this model is sufficient for justifying the rules for obtaining manufacturing specifications, it will give anomalous results with excessively wide specifications. In particular, the conditional mean of Y given x will begin to fall within the specifications when the value for x is outside the specifications, resulting in large probabilities of conformance for observations that are out-of-spec.

Therefore, the reader is cautioned against the uncritical use of the model when the capability values become excessively large. The model is sufficient to justify the rules for obtaining manufacturing specifications, and the rules work in practice, without anomalous results. Exact probability computations are not required. Simply choose your level of certainty, adjust the specs accordingly, and use the measurements.

11. Summary

Traditional gauge R&R studies have consistently overstated the damage due to measurement error by failing to take in to account the nature of the relationship between X and Y. When this relationship is carefully and rigorously worked out, the rules for obtaining Manufacturing Specifications given here are found to be the correct and appropriate ones for use in practice.