Peer-reviewed version: *Journal of Scientific and Mathematical Research*, Vol. 6 (2026) pp. 109-140
jscimath/pubs/J202606-109TM.pdf

The geometry of 3D statistical tolerancing: Radial errors and non-commuting manufacturing processes

Thad S. Morton

Normally distributed Cartesian errors centered at nominal produce a radial probability density concentrated away from nominal, becoming increasingly shell-like in higher dimensions. Position and runout measurements are examples of this in two dimensions. Relations are given for the variance of a gap in a three-dimensional stack path and for radial error when Cartesian component variances are equal and for when they are unequal, and examples are given of each. The effect of physically ordered, non-commuting manufacturing operations on covariance-based error propagation is examined. Higher-order moment methods are also developed for nonlinear error propagation and compared with Monte Carlo simulations.

1. Introduction

The vector loop used in tolerance stackup analysis is well known, but when stack paths are two-dimensional or three-dimensional, engineers are often unfamiliar with how to apply this principle. When x and y component data from position or runout measurements are processed, they are often treated separately, whereas engineering specifications frequently constrain distance r itself. This paper describes some principles that can be applied to analyze tolerances when stack paths are not confined to one or two dimensions and explains the chi distribution for radial data. For a review of work related to tolerancing, see Hong and Chang (2002) and Morse *et al.* (2018).

2. Tolerance Stackup Analysis

Figure 1 shows a hypothetical one-dimensional stack path, or closed loop, for a tolerance stackup analysis. Since all real constructions must allow for variations, a gap must occur somewhere along a closed path formed by mating hardware in an assembly. This closed path is called the *stack path*. Stack paths forming a 3-D vector loop have been depicted graphically by Gao *et al.* (1998) and Yang and Gong (2011), but an explicit formula for the total variance was not found in the literature. This formula will be given in (17) below.

If the length of the ith component in an assembly varies in the tolerance range $\pm\mathrm{Tol}_Y$, then the allowable tolerance of the last component to be inserted into the stack path is dictated by that of all the others. To avoid the possibility of part-to-part variation causing a "no-build" situation on an assembly line, we determine the tolerances of all parts by assuming that the worst possible dimensional combinations allowed on the engineering drawing arrive together on the assembly floor for a specific build. This method of worst-case tolerancing assumes that the gap in an assembly of N components is the following vector sum

$$\sum_{i=1}^{N} \triangle \mathbf{x}_i \;=\; \mathbf{Y} \tag{1}$$

of the worst dimensions $\triangle x_i$ around a closed loop in space along adjoining parts in the

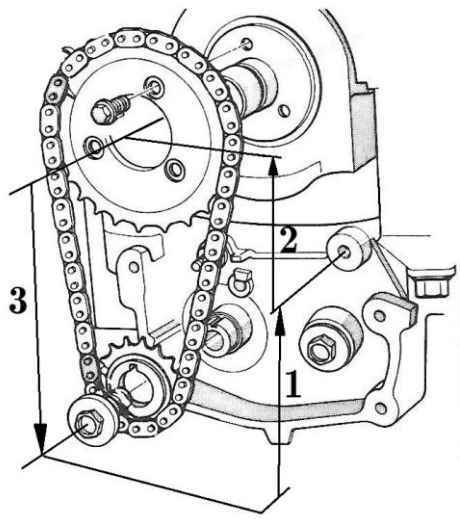


Figure 1. Simplified 1-D stack path in an engine involving the separation distance between sprockets. (Does not include runout of the sprocket or shaft.)

assembly. A tolerance stackup analysis consists of two tolerance stacks—one where the $\triangle \mathbf{x}_i$ are part dimensions at *maximum material conditions*, and another where the $\triangle \mathbf{x}_i$ are part dimensions at *least material condition.*

The way to think of the above summation (the same line integral used in surveying) is with its continuous analogue—the closed line integral:

$$\oint \mathrm{d}\mathbf{x} = 0 \qquad \text{exact construction (impossible)}$$

$$\oint \mathrm{d}\mathbf{x} = \mathbf{Y} \qquad \text{real construction} \tag{2}$$

In a perfect world, tolerances would usually all be zero so that any closed loop traversed vectorially through space along part dimensions called out on the engineering drawing would always yield exactly the zero vector if we returned to where we started. But unless all parts in the stack path are press-fit together (an interference fit), the gap $\mathbf{Y}$ of a stack path will never be exactly zero.

Consider a gap $\mathbf{Y}$ in a stack path containing m part dimensions from some assembly. The gap vector is given by the vector sum of the dimensions in the stack path:

$$\mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_m + \mathbf{Y} = 0\,. \tag{3}$$

This is referred to as a straight stack. When there are more than $m \approx 6$ or so dimensions in a stack path, it is customary to compute the gap statistically with an L^2 norm instead of using the straight stack (the vector sum). The variance V_Y of an output variable (which here is a gap in some subassembly of components) that is dependent on m independent input variables (such as all m dimensions in the stack path) is given by:

$$V_Y = V_1 + V_2 + \ldots + V_m\,, \tag{4}$$

where V_i is the variance of the ith variable. This statement will be generalized to interdependent input variables in the next section (see (17)), but (4) suffices for many cases. If these variables are normally distributed, then we can express the relation in terms of standard deviations, as follows:

$$s_Y{}^2 = s_1{}^2 + s_2{}^2 + \ldots + s_m{}^2 \tag{5}$$

We can use the prediction interval to set the upper and lower tolerance limits of the ith dimension in the stack path:

$$\mathrm{Tol}_i \;=\; \pm\, t_i\, s_i \sqrt{\frac{1}{N_i} \;+\; 1}\,, \tag{6}$$

where t is the number of standard deviations allowed in the tolerance band in order to achieve the desired probability of success P. (It is the student t-distribution, which is just the dimensionless z-value for small sample size). It is computed with a t-table or with the Excel function "=T.INV((1 - P)/Tails, ndf)," where Tails is either 1 or 2, depending upon whether the specified tolerance is one-sided or two-sided, and ndf is the number of degrees of freedom, which is $N-1$. The complement of the probability P is used in order to get the body of the distribution rather than the tails of the distribution. The result Tol_i is the difference $(x_i - \bar{x}_i)$ between the upper tolerance limit specified on the drawing and its nominal value $\bar{x}_i$. Notice from (6) that in an ideal world, the tolerance limits should depend upon the sample size N used to compute the standard deviation.

Knowing the standard deviation of the gap, the gap tolerance can be related to the fallout rate, or defect rate, allowed for the gap by means of a stackup t-score, t_Y:

$$\mathrm{Tol}_Y \;=\; \pm\, t_Y\, s_Y\,. \tag{7}$$

So t_Y is half the number of standard deviations contained in the tolerance band for the gap that will achieve the desired probability of success. By substituting (5) into (7), we have

$$\mathrm{Tol}_Y \;=\; \pm\, t_Y \sqrt{{s_1}^2 \;+\; {s_2}^2 \;+\; \ldots \;+\; {s_m}^2}\,. \tag{8}$$

If all the components in the stack path are designed to the same failure rate as the assembly itself, that is, if each t-value in (6) is equal to t_Y in (8), then (8) can be written as

$$\mathrm{Tol}_Y \;=\; \pm \sqrt{(t_1 s_1)^2 \;+\; (t_2 s_2)^2 \;+\; \ldots \;+\; (t_m s_m)^2}\,. \tag{9}$$

If we assume further that all the component tolerances in (6) are based on sample sizes with large N_i, then (6) becomes $\mathrm{Tol}_i \;\approx\; \pm\, t_i\, s_i$ so that (9) can be written:

$$\mathrm{Tol}_Y \;\approx\; \pm \sqrt{\sum_{i=1}^{m} {\mathrm{Tol}_i}^2}\,. \tag{10}$$

In other words, if the t value for the assembly gap is the same as the t values for all the components, then the t values drop out.

To compare the statistical tolerance stackup to the straight stack, we assume for simplicity that all the component tolerances are identical. Then (10) becomes

$$\mathrm{Tol}_Y \;=\; \sqrt{m}\,\mathrm{Tol}_i\,, \qquad \text{(statistical stack)} \tag{11}$$

whereas for the straight stack, we have simply

$$\mathrm{Tol}_Y \;=\; m\,\mathrm{Tol}_i\,. \qquad \text{(straight stack)} \tag{12}$$

2.1. Bender Safety Factor

To account for the fact that processes drift over time and for other uncertainties, Bender (1968) proposed a "$6 \times 2.5 = 9$" rule of thumb in tolerance stacking. He observed that if there are six independent components, each with a tolerance of ± 2.5, the combined assembly tolerance is not the arithmetic sum of $6(\pm 2.5) = \pm 15$ according to (12), but instead, it is closer to ± 9 (see Figure 2). (Bender pointed out that his lengths 3.5 ± 2.5 were chosen to correspond with the number of sides on the commercially available dice he used in his study.) Meanwhile, the statistical stack according to (11) gives a total tolerance of $\mathrm{Tol}_Y = \sqrt{6}\,(\pm 2.5)$

$\approx \pm 6.12$. This is why Bender introduced the factor $k = 1.5$ on the right side of (10). Bender stated that the factor 1.5 is necessary because without it, the distributions must all remain perfectly centered. And while this is true, the reason the 1.5 factor was needed in Bender's study is because he used uniform probability distributions for the 6 components in his assemblies (see Figure 3). If, instead, a normal probability distribution had been used, a factor of approximately $k = 1.1$ would have been appropriate rather than $k = 1.5$. This is illustrated in Figure 4. Still, component lengths can have non-normal distributions, so the assumption of uniform distribution is conservative and seems warranted.

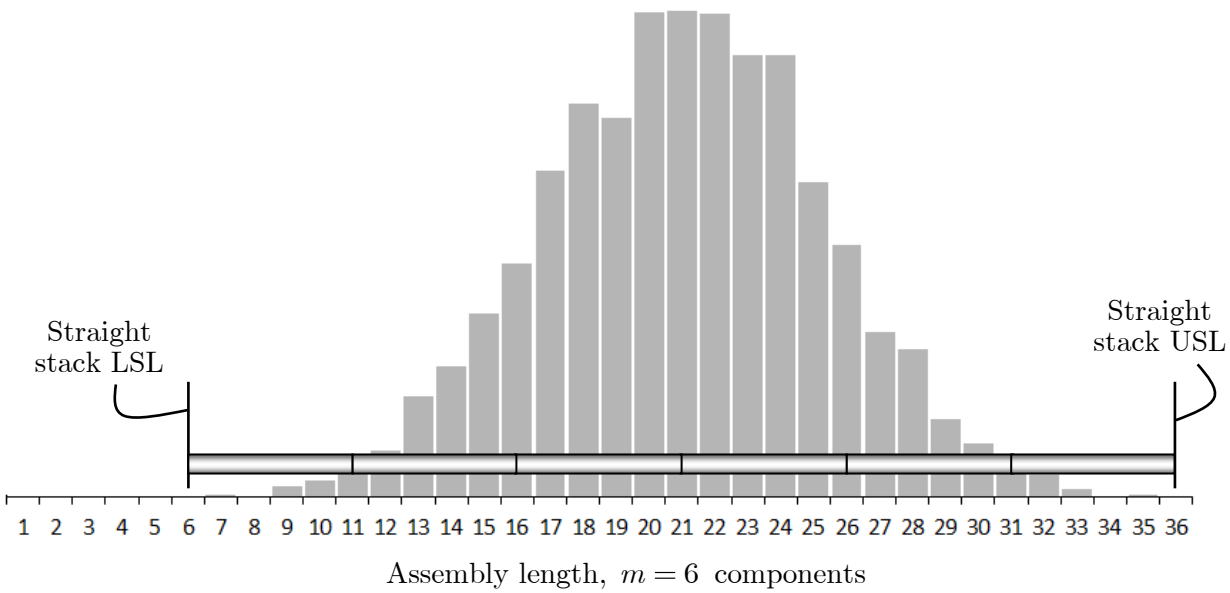


Figure 2. Histogram of the length of a 6-part assembly, with each part of length 3.5 ± 2.5. For comparison, the sum of tolerances of the 6-parts in the assembly are superimposed.

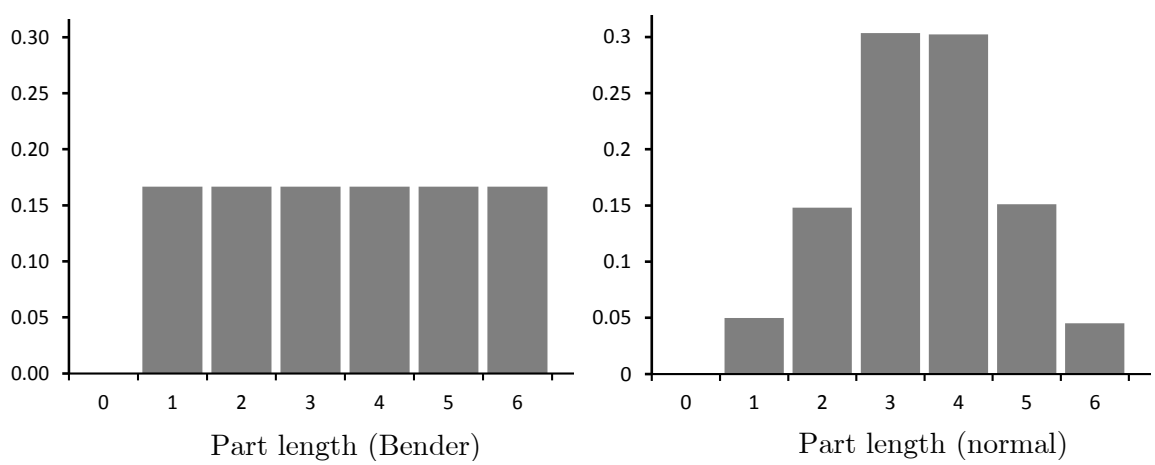


Figure 3. Histogram of Bender's uniformly distributed part (Left) and a normally distributed part (Right) with $\sigma = 1.2$. In both cases (used in Figure 4), the true means were $\mu = 3.5$.

Between the curves for the straight stack (the worst case stack) and the statistical stack shown in Figure 4 are plots from two simulations. Each point in these curves represents a simulation of 3000 points. The reason the defect rate is set to 4.5% (2.25% in each tail) in these simulations is because that is the probability to which Bender's factor $k = 1.5$ effectively corresponds. For comparison purposes, therefore, this probability was preserved for the simulation with normally distributed components.

The 1.5 factor has become the standard in many industries such as the automotive industry. Note that Figure 4 shows that a straight stack involving 3 components yields essentially the same assembly tolerance as the "Benderized" (with $k = 1.5$) statistical stack.

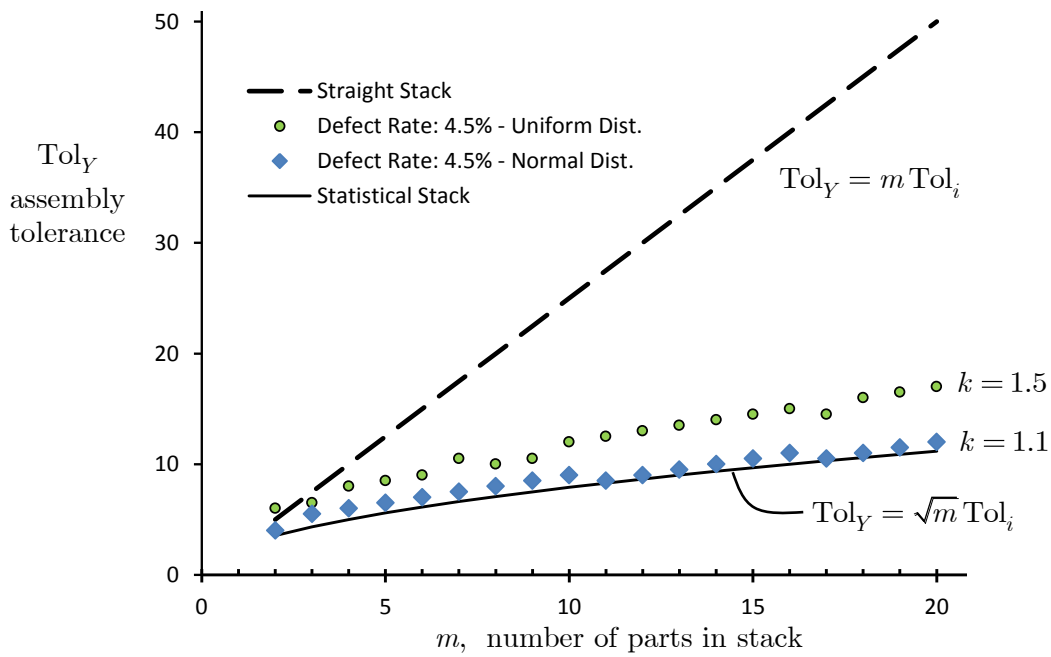


Figure 4. The effect of the $k = 1.5$ factor when component tolerances are all $\mathrm{Tol}_i = \pm 2.5$.

In addition, numerical simulation with components having continuous normal distributions confirms that the defect rate of the assembly will always match that of the components, without the need for the factor k, regardless of how few or how many parts are in the stack. But as noted, since non-normal distributions can arise in practice, the more conservative assumption of uniform distribution is probably prudent. Nevertheless, the crossover point for the two methods is clearly closer to $m = 3$ than $m = 6$. Incidentally, before Bender, Gilson (1951, pp. 45-46) had proposed a proportionality constant of $k = 1.6$.

2.2. Statistical Tolerance Stackup

To derive the relation for the gap in a statistical stack, now regard each component dimension as a random variable X_i, and write the differential of (3) as a gap function $\mathbf{Y}(X_i)$ that spans three-dimensional space ($\mathbf{Y} \in \mathbb{R}^n$ with $n = 3$):

$$\mathrm{d}\mathbf{Y} = -\frac{\partial \mathbf{Y}}{\partial X_1}\mathrm{d}X_1 - \frac{\partial \mathbf{Y}}{\partial X_2}\mathrm{d}X_2 - \ldots - \frac{\partial \mathbf{Y}}{\partial X_m}\mathrm{d}X_m. \tag{13}$$

The square of distance transforms as the square of the vector differential in (13) using Riemann's generalization of the Pythagorean Theorem:

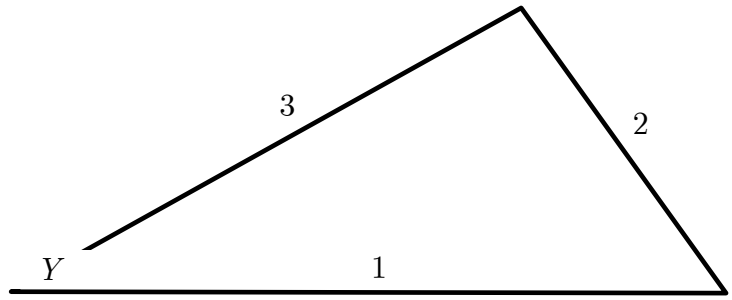


Figure 5. Stack path with 3 mating dimensions in a plane $(n = 2)$.

$$\mathrm{d}\,Y^2 \;=\; \mathrm{d}\mathbf{Y}\cdot\mathrm{d}\mathbf{Y} \;=\; \sum_{j=1}^{n}\sum_{i=1}^{m}\sum_{k=1}^{m}\frac{\partial\,Y_j}{\partial X_i}\frac{\partial\,Y_j}{\partial X_k}\,\mathrm{d}X_i\,\mathrm{d}X_k\,. \tag{14}$$

Variance also transforms as in (14), which is a generalization of a form proposed by Kline and McClintock (1953) for error analysis. The sample variance of variable X_1 is defined by

$$V_{11} = \sum_{s=1}^{N}\frac{[(x_1)_s - \overline{x}_1]^2}{N-1}\,. \tag{15}$$

Here, $\overline{x}_1$ is the mean (and possibly the nominal) of component 1, and $(x_1)_s$ is the sth measurement or observation of dimension 1. The number N is the number of observations, i.e., the sample size used to estimate the variance.

The generalization of (15) for the rare case when a part dimension can be influenced by the dimension of some other part in the assembly is the *sample covariance* V_{ik}, defined as:

$$V_{ik} = \sum_{s=1}^{N}\frac{[(x_i)_s - \overline{x}_i][(x_k)_s - \overline{x}_k]}{N-1}\,. \tag{16}$$

Since variance also transforms as in (14), the total variance $V(Y)$ due to the influence of variation of all components x_i on each other is:

$$V(Y) \;=\; \sum_{j=1}^{n}\sum_{i=1}^{m}\sum_{k=1}^{m}\frac{\partial\,Y_j}{\partial X_i}\frac{\partial\,Y_j}{\partial X_k}V_{ik}\,. \tag{17}$$

If we have an assembly of three parts as shown in Figure 5, the indices i and k both range from 1 to $m = 3$. The variable X_i represents the ith dimension in the stack, and X_k represents the kth dimension in the stack. The random variables X_i in a tolerance stackup will usually be statistically independent of each other so that the off-diagonal components of the covariance matrix will be zero.

If the entire stack path lies in a plane, as does the "assembly" in Figure 5, then $n = 2$, the index j sums only from 1 to 2, and the partial derivatives $\partial Y_1/\partial X_i$ and $\partial Y_2/\partial X_i$ will be the cosine and sine, respectively, of the angle made by dimension i. Since all distance functions in an assembly tolerance stackup are linear, the first derivatives are always constant, and (17) is generally valid. For sensitivity analysis or error propagation with nonlinear functions, see §4.

The relation in (17) says that the variance of an n-dimensional gap in an m-part assembly transforms as the square of distance in an nm-dimensional space.

2.3. Examples

Example 1: 2-D Stackpath, 1 Variable

Suppose that 6 components of equal length ($L = 60\,\text{mm}$) must fit together end-to-end to form a hexagon. The errors in the manufactured lengths of the parts L_1, L_2, L_3,..., L_6 will accumulate as we assemble them end-to-end, as shown in Figure 6. The position of the first end of the first piece (shown as a blue dot) is treated as fixed because we assume that the location of all the other part interfaces can adjust to the location of its starting point. Also, the assembly process is such that there is no variation in the angle at which the parts are assembled. The only variation is in the length L of the parts. Figure 6 shows the results of a thousand assemblies of our hexagon, assuming the lengths are normally distributed with standard deviation in length equal to $\sigma_L = 1\,\text{mm}$. Notice that the scatter of the endpoints of $\mathbf{L}_1$ appears in a linear pattern that follows the direction parallel to $\mathbf{L}_1$, but the two-dimensionality causes this pattern to break down at other positions. The accumulation of scatter in the counterclockwise direction seen in Figure 6 illustrates why it is so important to dimension all of the features from a common datum (coordinate origin) whenever possible.

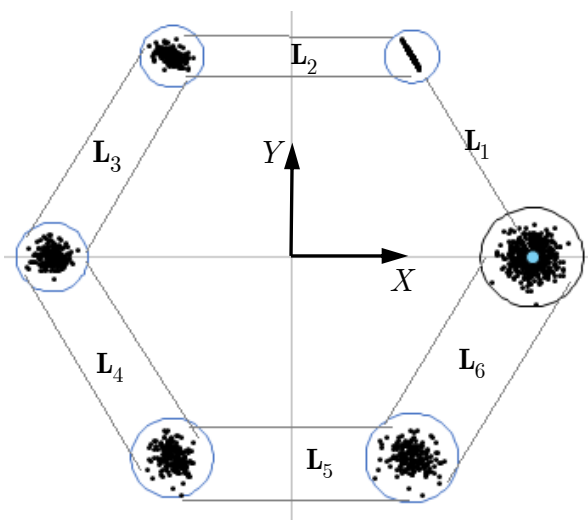


Figure 6. In multi-dimensional problems, variation of a single variable (the length of **L** in this case) can cause variation in multiple output variables (X and Y in this case). $L_i = 60\,\text{mm}$, $\sigma_{L_i} = 1\,\text{mm}$ $(i = 1,2,...,6)$. The circles shown correspond to $2.5\,\sigma_{L_i}$.

Let X and Y be the Cartesian coordinates measured from the origin. Then from

$$R = \sqrt{X^2 + Y^2}\,, \tag{18}$$

so that the expected value $\mathrm{E}[R]$ is the mean distance $\bar{R}$ of the endpoint from the origin (not the mean distance from the segment's starting point). The total Cartesian variance is

$$V_{\text{Cart}} \;=\; {\sigma_X}^2 + {\sigma_Y}^2\,.$$

This differs from the radial variance V_R, which describes the variance of the non-negative distance R from the coordinate origin. V_R is found by substituting $\mathrm{E}[R^2] = \mathrm{E}[X^2] + \mathrm{E}[Y^2]$, obtained from (18), into the well-known variance relation $V_R = \mathrm{E}[R^2] - \bar{R}^2$, yielding $V_R = \mathrm{E}[X^2] + \mathrm{E}[Y^2] - (\mathrm{E}[R])^2$. Then $\mathrm{E}[X^2]$ and $\mathrm{E}[Y^2]$ are replaced by again applying the same variance relation but to V_X and V_Y separately. The result is

$$V_R \;=\; {\sigma_R}^2 \;=\; {\sigma_X}^2 + {\sigma_Y}^2 + (\mathrm{E}[X])^2 + (\mathrm{E}[Y])^2 - (\mathrm{E}[R])^2\,. \tag{19}$$

Represent the length of each part in the assembly as $L + \delta L$, the distance L being the mean length. Then the variation can be simulated in Excel with a normal distribution by:

$$\delta L \text{ = NORMINV(RAND(), 0, } \sigma_L \text{)} \tag{20}$$

Therefore, δL is normally distributed and so can be positive and negative. Then the lengths of the parts are

$$\begin{aligned} L_1 &= L + \delta L_1 , \\ L_2 &= L + \delta L_2 , \\ &\vdots \end{aligned}$$

and so on.

The horizontal and vertical components of the end position of the sixth and final component are expressed as summations of all of the others:

$$\left.\begin{aligned} X_6 &= L_1 \cos\theta_1 + L_2 \cos\theta_2 + \ldots + L_6 \cos\theta_6 \\ Y_6 &= L_1 \sin\theta_1 + L_2 \sin\theta_2 + \ldots + L_6 \sin\theta_6 \end{aligned}\right\} \tag{21}$$

As you can see from these relations, errors in the end position of the final component in the assembly will depend on errors of all the other components.

Since the variables L and θ are independent (in fact, θ has no variability), the only partial derivatives in (17) that are needed are

$$\frac{\partial X}{\partial L_1} = \cos\theta_1 , \quad \frac{\partial X}{\partial L_2} = \cos\theta_2 , \quad \ldots, \frac{\partial X}{\partial L_6} = \cos\theta_6$$

$$\frac{\partial Y}{\partial L_1} = \sin\theta_1 , \quad \frac{\partial Y}{\partial L_2} = \sin\theta_2 , \quad \ldots, \frac{\partial Y}{\partial L_6} = \sin\theta_6$$

To understand the effect that our manufacturing process has on the variance of location of the sixth and final piece in the assembly, we expand (17) and write

$$\begin{aligned} V_{\text{Cart}} = \; & \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial L_1} V_{L_1 L_1} + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial L_2} V_{L_1 L_2} + \ldots + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial L_6} V_{L_1 L_6} \\ & + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial L_1} V_{L_1 L_2} + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial L_2} V_{L_2 L_2} + \ldots + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial L_6} V_{L_2 L_6} \\ & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ & + \frac{\partial X}{\partial L_6}\frac{\partial X}{\partial L_1} V_{L_6 L_1} + \frac{\partial X}{\partial L_6}\frac{\partial X}{\partial L_2} V_{L_6 L_2} + \ldots + \frac{\partial X}{\partial L_6}\frac{\partial X}{\partial L_6} V_{L_6 L_6} \\ + \; & \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial L_1} V_{L_1 L_1} + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial L_2} V_{L_1 L_2} + \ldots + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial L_6} V_{L_1 L_6} \\ & + \frac{\partial Y}{\partial L_2}\frac{\partial Y}{\partial L_1} V_{L_1 L_2} + \frac{\partial Y}{\partial L_2}\frac{\partial Y}{\partial L_2} V_{L_2 L_2} + \ldots + \frac{\partial Y}{\partial L_2}\frac{\partial Y}{\partial L_6} V_{L_2 L_6} \\ & \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ & + \frac{\partial Y}{\partial L_6}\frac{\partial Y}{\partial L_1} V_{L_6 L_1} + \frac{\partial Y}{\partial L_6}\frac{\partial Y}{\partial L_2} V_{L_6 L_2} + \ldots + \frac{\partial Y}{\partial L_6}\frac{\partial Y}{\partial L_6} V_{L_6 L_6} \end{aligned}$$

But since none of the independent variables are correlated with each other, all off-diagonal components of the covariance matrix are zero. Now we input the standard deviations:

$$V_{\text{Cart}} = \left(\frac{\partial X}{\partial L_1}\right)^2 \sigma_{L_1}{}^2 + \left(\frac{\partial X}{\partial L_2}\right)^2 \sigma_{L_2}{}^2 + \ldots + \left(\frac{\partial X}{\partial L_6}\right)^2 \sigma_{L_6}{}^2$$
$$+ \left(\frac{\partial Y}{\partial L_1}\right)^2 \sigma_{L_1}{}^2 + \left(\frac{\partial Y}{\partial L_2}\right)^2 \sigma_{L_2}{}^2 + \ldots + \left(\frac{\partial Y}{\partial L_6}\right)^2 \sigma_{L_6}{}^2$$

Substituting values for the partial derivatives computed above gives

$$\sigma_{\text{Cart}}{}^2 = \cos^2\theta_1\, \sigma_{L_1}{}^2 + \cos^2\theta_2\, \sigma_{L_2}{}^2 + \ldots + \cos^2\theta_6\, \sigma_{L_6}{}^2$$
$$+ \sin^2\theta_1\, \sigma_{L_1}{}^2 + \sin^2\theta_2\, \sigma_{L_2}{}^2 + \ldots + \sin^2\theta_6\, \sigma_{L_6}{}^2$$

Notice that for a linear (in-line or 1D) stack path (where $\cos\theta = \pm 1$ and $\sin\theta = 0$), this simplifies to the conventional one-dimensional tolerance stackup method wherein the cosine functions are replaced by a factor of 1 or -1, depending upon whether they are directed rightward or leftward (or upward or downward). Note that for a straight stack ($m < 7$), these factors do not square to all positive values.

Simplifying further,

$$\sigma_{\text{Cart}} = \sqrt{\sigma_{L_1}{}^2 + \sigma_{L_2}{}^2 + \sigma_{L_3}{}^2 + \sigma_{L_4}{}^2 + \sigma_{L_5}{}^2 + \sigma_{L_6}{}^2} = \sqrt{6}\, \sigma_{L_i} = \sqrt{6}\ \text{mm}.$$

Since the Cartesian errors at the final endpoint are centered and we have equal variances, $\sigma_X{}^2 = \sigma_Y{}^2 = 3\ \text{mm}^2$, the mean radial error is $\text{E}[R] = \sigma_X\sqrt{\pi/2}$, as derived later in (52). Substituting into (19) gives the radial variance of the final gap:

$$V_R = 6 + 0 + 0 - 3\frac{\pi}{2} = 1.2876\ \text{mm}^2 \quad \rightarrow \quad \sigma_R = 1.1347\ \text{mm}$$

Example 2: 2-D Stackpath, 2 Variables

Now suppose that we let both L and θ vary in the manufacture of our product; however, they still vary independently. The length $L + \delta L$ of the part is normally distributed. The distance L is the mean length, and the variation can again be simulated in Excel by:

$$\delta L = \text{NORMINV(RAND(), 0, } \sigma_L \text{)} \tag{22}$$

Therefore, δL is again normally distributed and so can be positive and negative. Then the Cartesian coordinates of the endpoint of the first part, for example, are

$$\left.\begin{aligned} X_1 &= L\cos\theta_1 + \delta L_1 \cos\theta_1 \\ Y_1 &= L\sin\theta_1 + \delta L_1 \sin\theta_1 \end{aligned}\right\} \tag{23}$$

The angles made by the 6 position vectors of the six points are $\theta_1 = 0°$, $\theta_2 = 60°$, $\theta_3 = 120°$, $\theta_4 = 180°$, $\theta_5 = 240°$, and $\theta_6 = 300°$. The angles vary normally such that the arc length $L\,\delta\theta$ has a standard deviation of 1 mm at the end of each segment. Therefore, $\sigma_\theta = 1\,\text{mm}/L$, and data is simulated with the following Excel formula:

$$\delta\theta = \text{NORMINV(RAND(), 0, } \sigma_\theta \text{)}.$$

The partial derivatives called for in (17) are

$$\frac{\partial X}{\partial L_1} = \cos\theta_1, \quad \frac{\partial X}{\partial L_2} = \cos\theta_2, \quad \cdots \quad \frac{\partial X}{\partial L_6} = \cos\theta_6$$
$$\frac{\partial Y}{\partial L_1} = \sin\theta_1, \quad \frac{\partial Y}{\partial L_2} = \sin\theta_2, \quad \cdots \quad \frac{\partial Y}{\partial L_6} = \sin\theta_6$$

$$\frac{\partial X}{\partial \theta_1} = -L_1 \sin\theta_1, \quad \frac{\partial X}{\partial \theta_2} = -L_2 \sin\theta_2, \quad \cdots \quad \frac{\partial X}{\partial \theta_6} = -L_6 \sin\theta_6$$

$$\frac{\partial Y}{\partial \theta_1} = L_1 \cos\theta_1, \quad \frac{\partial Y}{\partial \theta_2} = L_2 \cos\theta_2, \quad \cdots \quad \frac{\partial Y}{\partial \theta_6} = L_6 \cos\theta_6$$

Now we will see the patterns scatter as shown in Figure 7. Expanding (17) gives:

$$\begin{aligned}
V_{\text{Cart}} = {} & \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial L_1} V_{L_1 L_1} + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial L_2} V_{L_1 L_2} + \ldots + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial L_6} V_{L_1 L_6} \\
& + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial \theta_1} V_{L_1 \theta_1} + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial \theta_2} V_{L_1 \theta_2} + \ldots + \frac{\partial X}{\partial L_1}\frac{\partial X}{\partial \theta_6} V_{L_1 \theta_6} \\
& + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial L_1} V_{L_2 L_1} + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial L_2} V_{L_2 L_2} + \ldots + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial L_6} V_{L_2 L_6} \\
& + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial \theta_1} V_{L_2 \theta_1} + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial \theta_2} V_{L_2 \theta_2} + \ldots + \frac{\partial X}{\partial L_2}\frac{\partial X}{\partial \theta_6} V_{L_2 \theta_6} \\
& \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
& + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial L_1} V_{L_1 L_1} + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial L_2} V_{L_1 L_2} + \ldots + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial L_6} V_{L_1 L_6} \\
& + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial \theta_1} V_{L_1 \theta_1} + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial \theta_2} V_{L_1 \theta_2} + \ldots + \frac{\partial Y}{\partial L_1}\frac{\partial Y}{\partial \theta_6} V_{L_1 \theta_6} \\
& \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
& + \frac{\partial Y}{\partial L_6}\frac{\partial Y}{\partial \theta_1} V_{L_6 \theta_1} + \frac{\partial Y}{\partial L_6}\frac{\partial Y}{\partial \theta_2} V_{L_6 \theta_2} + \ldots + \frac{\partial Y}{\partial L_6}\frac{\partial Y}{\partial \theta_6} V_{L_6 \theta_6}.
\end{aligned}$$

But again, since none of the independent variables are correlated with each other, all off-diagonal components of the covariance matrix are zero. Therefore,

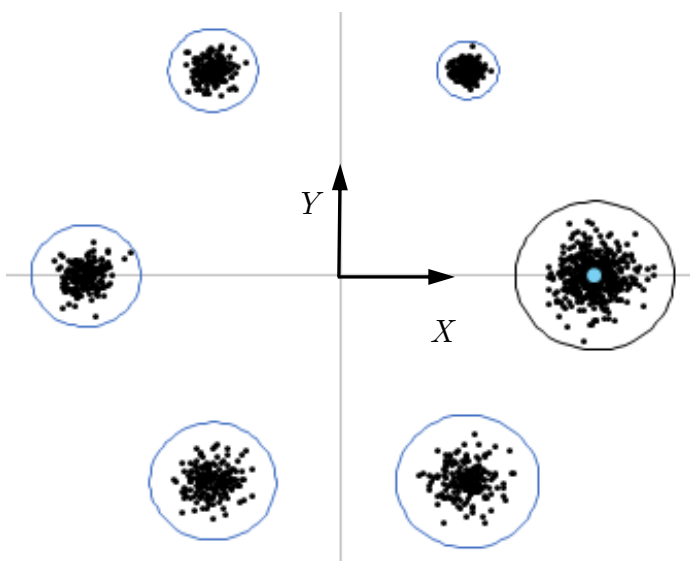


Figure 7. As we start from the Cartesian point $(L, 0)$ and assemble counterclockwise, the variance of each endpoint position accumulates. Values: $L_i = 60\,\text{mm}$, $\sigma_{L_i} = 1\,\text{mm}$, $\sigma_{\theta_i} = 1/60$. The circles shown correspond to $2.5\sigma_{L_i}$ where $(i = 1, 2, \ldots, 6)$.

$$V_{\text{Cart}} = \left(\frac{\partial X}{\partial L_1}\right)^2 \sigma_{L_1}{}^2 + \left(\frac{\partial X}{\partial \theta_1}\right)^2 \sigma_{\theta_1}{}^2 + \ldots + \left(\frac{\partial X}{\partial L_6}\right)^2 \sigma_{L_6}{}^2 + \left(\frac{\partial X}{\partial \theta_6}\right)^2 \sigma_{\theta_6}{}^2$$
$$+ \left(\frac{\partial Y}{\partial L_1}\right)^2 \sigma_{L_1}{}^2 + \left(\frac{\partial Y}{\partial \theta_1}\right)^2 \sigma_{\theta_1}{}^2 + \ldots + \left(\frac{\partial Y}{\partial L_6}\right)^2 \sigma_{L_6}{}^2 + \left(\frac{\partial Y}{\partial \theta_6}\right)^2 \sigma_{\theta_6}{}^2,$$

which, after the appropriate substitutions, becomes

$$V_{\text{Cart}} = \cos^2\theta_1\, \sigma_{L_1}{}^2 + (L_1 \sin\theta_1)^2\, \sigma_{\theta_1}{}^2 + \ldots + \cos^2\theta_6\, \sigma_{L_6}{}^2 + (L_6 \sin\theta_6)^2\, \sigma_{\theta_6}{}^2$$
$$+ \sin^2\theta_1\, \sigma_{L_1}{}^2 + (L_1 \cos\theta_1)^2\, \sigma_{\theta_1}{}^2 + \ldots + \sin^2\theta_6\, \sigma_{L_6}{}^2 + (L_6 \cos\theta_6)^2\, \sigma_{\theta_6}{}^2.$$

Therefore,

$$V_{\text{Cart}} = \sum_{i=1}^{6} \sigma_{L_i}{}^2 + L_i^2 \sigma_{\theta_i}{}^2$$

Substituting values, we find that for each segment, $\sigma_L{}^2 + L^2\sigma_\theta{}^2 = 2\ \text{mm}^2$. Therefore, $V_{\text{Cart}} = 12\ \text{mm}^2$. Since the Cartesian variances are equal, $\sigma_X{}^2 = \sigma_Y{}^2 = 6\ \text{mm}^2$. From (19), $V_R = V_{\text{Cart}} - \bar{R}^2$, so using (52) for $\bar{R}$ again gives

$$V_R = 12 - (6)\pi/2 = 2.5752\ \text{mm}^2 \qquad \rightarrow \qquad \sigma_R = 1.6047\ \text{mm}$$

This provides a first-order estimate of the standard deviation of position error.

Example 3: 2-D Variation, Correlated Dimensions

An example of a situation wherein two dimensions may be correlated is in the stackup of casting tolerances on a casting drawing and machining tolerances on the subsequent machining drawing (or finish drawing). Often holes that must be machined in a cast housing will be first cast with core pin holes in order to minimize the amount of material that must be removed during machining, thereby reducing cycle time and increasing tool life. Core pin holes also reduce the amount of porosity in the casting wall that is exposed during the machining process. Porosity in a thin-walled housing can allow, for example, detrimental oil seepage out of the machine housing over time.

Suppose we wish to determine a reasonable tolerance for a hole that is to be machined into a casting in which a smaller core pin hole is already cast. The machining equipment will introduce some variation into the process when it brings the drill to the desired hole location. But when the tool begins to actually drill the material, the final hole position will be shifted slightly because the drill will "follow" the position of the cast core-pin hole to some extent.

Suppose that a casting process has demonstrated the capability of casting holes to within a position tolerance of 0.5 mm. And the subsequent machining process has demonstrated the capability to hold hole position to a tolerance of 0.2 mm when drilling a hole without the presence of a core pin hole. However, when drilling into a pre-cast hole, the machinist finds that the tool or drill bit follows the cast hole so that the final position (X, Y) of the machined hole is influenced by the variability in the cast hole position (x, y). So the two variables are slightly correlated, and the covariance of the two will be non-zero. In order to apply (17), use the following notation for the input variables:

$x_1 =$ cast position x
$x_2 =$ cast position y
$x_3 =$ machined position x (hole drilled into a fresh (flat) surface)

$x_4 =$ machined position y (hole drilled into a fresh (flat) surface)

The output variables are:

$X =$ final hole x position that has been drilled into a cast hole

$Y =$ final hole y position that has been drilled into a cast hole

and

$$R^2 = X^2 + Y^2 .$$

Applying (17) gives

$$\begin{aligned}
{\sigma_R}^2 &= {\sigma_X}^2 + {\sigma_Y}^2 \\
{\sigma_R}^2 &= \frac{\partial X}{\partial x_1}\frac{\partial X}{\partial x_1}V_{11} + \frac{\partial X}{\partial x_2}\frac{\partial X}{\partial x_2}V_{22} + \frac{\partial X}{\partial x_3}\frac{\partial X}{\partial x_3}V_{33} + \frac{\partial X}{\partial x_4}\frac{\partial X}{\partial x_4}V_{44} \\
&\quad + 2\frac{\partial X}{\partial x_1}\frac{\partial X}{\partial x_2}V_{12} + 2\frac{\partial X}{\partial x_1}\frac{\partial X}{\partial x_3}V_{13} + 2\frac{\partial X}{\partial x_1}\frac{\partial X}{\partial x_4}V_{14} + 2\frac{\partial X}{\partial x_2}\frac{\partial X}{\partial x_3}V_{23} \\
&\quad + 2\frac{\partial X}{\partial x_2}\frac{\partial X}{\partial x_4}V_{24} + 2\frac{\partial X}{\partial x_3}\frac{\partial X}{\partial x_4}V_{34} \\
&\quad + \frac{\partial Y}{\partial x_1}\frac{\partial Y}{\partial x_1}V_{11} + \frac{\partial Y}{\partial x_2}\frac{\partial Y}{\partial x_2}V_{22} + \frac{\partial Y}{\partial x_3}\frac{\partial Y}{\partial x_3}V_{33} + \frac{\partial Y}{\partial x_4}\frac{\partial Y}{\partial x_4}V_{44} \\
&\quad + 2\frac{\partial Y}{\partial x_1}\frac{\partial Y}{\partial x_2}V_{12} + 2\frac{\partial Y}{\partial x_1}\frac{\partial Y}{\partial x_3}V_{13} + 2\frac{\partial Y}{\partial x_1}\frac{\partial Y}{\partial x_4}V_{14} + 2\frac{\partial Y}{\partial x_2}\frac{\partial Y}{\partial x_3}V_{23} \\
&\quad + 2\frac{\partial Y}{\partial x_2}\frac{\partial Y}{\partial x_4}V_{24} + 2\frac{\partial Y}{\partial x_3}\frac{\partial Y}{\partial x_4}V_{34}
\end{aligned}$$

Since X is orthogonal to x_2 and x_4, and Y is orthogonal to x_1 and x_3, many of the partial derivatives above are zero. Those remaining have magnitude 1. We are left with

$$\sigma_R = \sqrt{V_{11} + V_{33} + 2V_{13} + V_{22} + V_{44} + 2V_{24}} \tag{24}$$

Since

$$V_{11} + V_{22} = ({\sigma_x}^2 + {\sigma_y}^2)_{\text{cast}} = (\sigma_{R_{\text{cast}}})^2$$

$$V_{33} + V_{44} = ({\sigma_x}^2 + {\sigma_y}^2)_{\text{machined}} = (\sigma_{R_{\text{machined}}})^2$$

we can write (24) as

$$\sigma_R = \sqrt{(\sigma_{R_{\text{cast}}})^2 + (\sigma_{R_{\text{machined}}})^2 + 2(V_{13} + V_{24})} \tag{25}$$

In order to compute the covariance V_{13} in (24) from measured data using (16), we need *paired data*. That is, we need the position of the drill bit just before it contacts the core pin hole region—before it is influenced by the core pin. This data is, of course, not practical to obtain. However, we could measure all other quantities in (24) besides the covariance terms in order to assess them. But to do this, it will be easier if we separate (24) into two separate equations—one for the horizontal position and one for the vertical position, as follows:

$${\sigma_X}^2 = V_{11} + V_{33} + 2V_{13} \tag{26}$$

$${\sigma_Y}^2 = V_{22} + V_{44} + 2V_{24} \tag{27}$$

Now we could measure V_{33} by measuring hole positions machined without a core pin,

measure V_{11} by measuring core pin hole positions without machining, and finally, measure hole positions machined into pre-cast holes to obtain σ_X. Then we could solve (26) for V_{13}.

If we were to attempt instead to measure V_{13}, it would be imperative that X-data and Y-data be tracked separately, and not simply the total error in the position R alone. When assessing whether the machining process will leave some raw casting surface exposed (non-cleanup) in the hole threads, it is critical to know whether the machined hole has shifted *toward the center* of the core pin hole or *away from it*. This information will be lost if one only analyzes the resultant position R.

We can simulate the situation with the R programming language. We want two vectors, each with 100,000 random numbers. The first vector is to represent the x position of a core pin hole that is cast into a metal housing. The mean of this x position data will be, say, 50.1 mm from some datum, and the variance of the x data is, say, 0.0038 mm^2. A second vector will represent the capability of the machining equipment. It will be the x position obtained by the drill bit just before it begins machining into a surface where a cast hole already exists. Let's say that during this machining trial, the mean of this position was 50 mm and the variance was found to be 0.001 mm^2. Since the machining "follows" the cast core pin hole to some extent, let's assume (again in lieu of test data) that the covariance between the cast hole position and the machining of the hole is $-0.0003\ \text{mm}^2$. For this situation, we can create a hypothetical example with the help of the statistical software "R" to generate random machining and casting data sets that are correlated to each other with a specified correlation coefficient. To do this, we use the command **mvrnorm**, which means "multivariate random numbers that are normally distributed." It is a function from an R library called **MASS**, so in order to use it, we must invoke that library first by typing:

```
library(MASS)
```

Then the command to generate a two-dimensional normal distribution with a specified covariance matrix is:

```
x <- mvrnorm(100000, mu=c(50.1,50), matrix(c(0.0038, -0.0003,
-0.0003, 0.001),2))
```
(28)

This line generates 100,000 random numbers from two normal distributions where variable 1 has a mean of 50.1, and variable 2 has a mean of 50.0. The covariance matrix is specified next (it must be a positive-definite symmetric matrix). We can check to see how close the results of our simulated random data are to our specified input statistics by using the **var()** command on the newly created data **x**. To compute the variance of **x** with the R programming language, type the function **var(x)**. Here is the type of output it provides:

```
var(x)
              [,1]          [,2]
[1,]  0.0037632600 -0.0002961223
[2,] -0.0002961223  0.0009999631
```

The variance and covariance values we supply as input must satisfy the Schwarz inequality:

$$V_{ik} = \sum_{s=1}^{N} \frac{[(x_i)_s - \overline{x}_i][(x_k)_s - \overline{x}_k]}{N-1}$$
$$\leq \sqrt{\sum_{s=1}^{N} \frac{[(x_i)_s - \overline{x}_i]^2}{N-1}} \sqrt{\sum_{s=1}^{N} \frac{[(x_k)_s - \overline{x}_k]^2}{N-1}} = \sigma_i \sigma_k$$

For the present 2×2 case, strict inequality will make the covariance matrix positive definite.

Then applying the conventional symmetric-covariance interpretation of (24), the horizontal variance if a hole is drilled where a smaller cast hole already exists would be

$$\begin{aligned} {\sigma_X}^2 &= V_{11} + V_{33} + 2V_{13} \\ &= 0.0038613 + 0.00093936 + 2(-0.0002822) \\ &= 0.004236 \end{aligned} \tag{29}$$

Therefore, the X-component of the standard deviation of the final machining process will be approximately

$$\sigma_X = 0.065\,\text{mm}.$$

We would follow the same process on the Y position data and obtain σ_Y. This result, however, implicitly counts both V_{13} and V_{31}. For the casting-machining process, that symmetry has a physical consequence that must be examined. Equation (29) is the result obtained from conventional covariance-based propagation. The next section examines whether that treatment is appropriate when the correlated operations possess a fixed physical ordering.

2.4. Non-commuting Processes

Variance propagation by (17) treats covariance conventionally as a symmetric statistical quantity:

$$\mathbf{V} = \begin{bmatrix} \sum_{s=1}^{N} R_{1s}R_{1s} & \sum_{s=1}^{N} R_{1s}R_{2s} \\ \sum_{s=1}^{N} R_{2s}R_{1s} & \sum_{s=1}^{N} R_{2s}R_{2s} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1 \cdot \mathbf{R}_1 & \mathbf{R}_1 \cdot \mathbf{R}_2 \\ \mathbf{R}_2 \cdot \mathbf{R}_1 & \mathbf{R}_2 \cdot \mathbf{R}_2 \end{bmatrix},$$

where $R_{is} = x_{is} - \bar{x}_i$ is the residual. Thus, the off-diagonal V_{ik} and V_{ki} are equal, and the two terms combine to produce the familiar factor of 2 multiplying the covariance in (29). The two cross terms arise algebraically from expanding the square of a sum; they do not represent two physical process sequences. This distinction becomes important when the variables describe sequential manufacturing operations, because manufacturing processes generally need not commute. In the casting-machining example above, machining of a hole can follow casting of a core-pin hole, but the reverse sequence does not physically occur—machining of a hole *never precedes* the casting of a core pin. Equation (29), however, contains both off-diagonal covariance terms and therefore does not retain this physical ordering.

For immediate engineering use, a simple empirical treatment can be employed for a one-way process pair of the type considered here. In manufacturing data analyzed by the author in a case such as this where paired data establish the covariance of a process, the measured variance of the completed process was reproduced only when the cross term was counted once, not twice. Thus, for the present casting-machining data, (29) becomes

$${\sigma_X}^2 = V_{11} + V_{33} + V_{13}. \tag{30}$$

For a one-way process pair of this type, manufacturing data analyzed separately by the author showed that counting the cross term once reproduced the measured completed-process variance, whereas counting it twice did not. That empirical observation suggests that (30) provides a simple empirical correction for this one-way process pair. It should not be interpreted as a redefinition of ordinary covariance, which remains symmetric, or as a general identity for arbitrary correlated random variables. Rather, it is an empirical representation of a physically ordered interaction whose reverse process is absent. This strategy is particularly convenient here because R requires the covariance matrix entered in (28) to be symmetric and therefore cannot represent the proposed one-way interaction by assigning different values

to the two off-diagonal entries, which can be confirmed by checking using the **var(x)** command. As noted, when the reverse sequence in a manufacturing process does not exist, the covariance contribution may be counted once rather than twice. Thus, for this physically ordered process, the variance consistent with the observed data is not (29), but rather:

$$\sigma_X{}^2 = 0.0038613 + 0.00093936 + (-0.0002822)$$

$$\sigma_X{}^2 = 0.004518. \tag{31}$$

Therefore, the X-component of the standard deviation of the final machining process will be

$$\sigma_X = 0.067\,\text{mm}.$$

To the variance in (31) would be added other sources of variation, such as the variation between cast datum tabs (used for fixturing the part in the machining center) and the machined datum to which the machined hole may be referenced.

A more general problem arises when several manufacturing operations can be performed in different orders. Suppose that we are examining surface roughness measurement data and trying to understand its impact on the friction coefficient of a surface. Let us say that the surface can be honed, or polished, or both. Let the standard deviation of the surface roughness value for the honing process be σ_1 and that for a polished surface be σ_2. The variance V of the friction coefficient Y of the surface can be found from (17), where the covariance V_{ik} of honing and polishing is computed from the data by (16). But the variance of surface finish when polishing a honed surface is likely not the same as it is for honing a polished surface. Furthermore, a polishing operation following honing may behave differently from the same polishing operation following grinding. More generally, an operation may retain effects of processes separated from it by several intervening operations. Surface geometry, residual stress, subsurface damage, material hardness, and other properties produced by an earlier operation may survive subsequent material removal and influence a later operation. Process history therefore contains information that an ordinary symmetric covariance matrix cannot retain. The interactions among manufacturing stages need not be specified entirely a priori; Zeng and Zhou (2007), for example, used graphical models together with engineering knowledge to infer interstage interactions in complex manufacturing processes.

One possible framework is to regard each manufacturing operation as an operator acting on the state produced by the preceding operations. Zhou *et al.* (2003) developed a state-space model in which dimensional variation is accumulated and transformed as a workpiece passes through successive machining stages, and Huang and Shi (2004) extended this approach to serial-parallel manufacturing systems having multiple process routes. In a linear approximation,

$$\mathbf{x}_{k+1} = \mathbf{A}_k\,\mathbf{x}_k + \varepsilon_k\,,$$

where $\mathbf{x}_k$ describes the state entering process k, $\mathbf{A}_k$ describes how that process transmits or modifies the existing state, and ε_k represents variation newly introduced by process k. Even deterministic process operators need not commute,

$$\mathbf{A}_j\,\mathbf{A}_i \neq \mathbf{A}_i\,\mathbf{A}_j\,.$$

Manufacturing history can produce a stronger ordering effect, however, because the statistical character of an operation may itself depend upon the operations that preceded it. The process operator may therefore be regarded as

$$\mathbf{A}_k \sim D(P_k \mid P_{k-1} \mid P_{k-2} \mid \ldots)\,,$$

meaning that $\mathbf{A}_k$ is random, with probability distribution D conditioned upon process history. For example, $\mathbf{A}_{P|H|T|G}$ may describe polishing P following honing H following heat treat T following grinding G. There is no general requirement that this operator resemble $\mathbf{A}_{H|P|G|T}$, or that either bear a simple relationship to the operators associated with a reversed process sequence. Thus, changing process order may change not merely the order in which operators act, but the operators and their probability distributions themselves.

If, conditioned upon the specified process history, $\mathbf{A}_k$ is statistically independent of the entering state, and if the zero-mean process error ε_k is uncorrelated with $\mathbf{A}_k \mathbf{x}_k$, then the first two moments can be propagated recursively. For zero-mean process error, define

$$\boldsymbol{\mu}_k = \mathrm{E}[\mathbf{x}_k], \qquad \mathbf{M}_k = \mathrm{E}[\mathbf{x}_k \, \mathbf{x}_k{}^T], \qquad \bar{\mathbf{A}}_k = \mathrm{E}[\mathbf{A}_k],$$

and

$$\mathbf{Q}_k = \mathrm{E}[\varepsilon_k \, \varepsilon_k{}^T].$$

where $\mathrm{E}[\mathbf{x}_k]$ is the expected value of $\mathbf{x}_k$. Then let (Xing *et al.* 2022)

$$\boldsymbol{\mu}_{k+1} = \bar{\mathbf{A}}_k \boldsymbol{\mu}_k, \qquad \mathbf{M}_{k+1} = \mathrm{E}[\mathbf{A}_k \mathbf{M}_k \mathbf{A}_k^T] + \mathbf{Q}_k. \tag{32}$$

The covariance matrix of the resulting state is

$$\mathrm{Cov}(\mathbf{x}_{k+1}) = \mathbf{M}_{k+1} - \boldsymbol{\mu}_{k+1} \boldsymbol{\mu}_{k+1}^T. \tag{33}$$

Successive application of (32) and (33) therefore provides a possible means of propagating variation through an arbitrary sequence of manufacturing operations. (Xing *et al.* 2022 used random state-transition matrices within multiplicative-noise system models.)

Variation-propagation models have been used for process planning. Liu *et al.* (2009), for example, formulated multistage machining setup planning as an optimization problem in which variation propagation is evaluated quantitatively to obtain a cost-effective process plan satisfying final quality requirements. Different process sequences can involve different history-conditioned operators, such as $\mathbf{A}_G$, $\mathbf{A}_{H|G}$, and $\mathbf{A}_{P|H|G}$ rather than merely different permutations of the same matrices. If the required operator statistics can be determined experimentally, alternative manufacturing sequences could be compared and their ordering selected to minimize final variation or probability of exceeding a tolerance, thereby extending established variation-based approaches to process and setup planning (Liu *et al.* 2009; Abellan-Nebot *et al.* 2011).

A complete theory of such history-dependent stochastic process propagation is beyond the scope of the present work. In particular, the conditional operators must be identified experimentally, the amount of process history that must be retained remains to be determined, and additional state variables may be needed when earlier operations influence later ones through persistent material or geometrical properties. The present observations therefore suggest a broader question for future investigation:

Can manufacturing variation be represented as a sequence of history-conditioned random operators whose experimentally determined statistics permit the ordering of otherwise acceptable manufacturing processes to be chosen so as to minimize the final probability of nonconformance?

In such a formulation, ordinary covariance would retain its conventional symmetric meaning, while the history-conditioned process operators would retain the directional information that covariance itself cannot contain.

3. Statistics of Distances

The expected value, or mean, for the discrete case is:

$$\mathrm{E}[q] \;=\; \sum_{i=1}^{N} q_i\, f_i\,, \tag{34}$$

where f_i is the probability of each outcome q_i. In the limit of continuous variables and outcomes, the *probability density function* is used to compute the mean, as follows:

$$\mathrm{E}[q] \;=\; \int q(x) f(x) \mathrm{d}x\,. \tag{35}$$

Table 1 provides some useful integrals involving the standard normal probability density function (36) and several means calculated using it. They are instances of (35) for finding the mean of a function.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{(-z^2/2)}\,\mathrm{d}z \;=\; 1 \qquad (36)$$	
$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\, e^{(-z^2/2)}\,\mathrm{d}z \;=\; 0 \qquad (37)$$	$$\int_{-\infty}^{\infty} \lvert z\rvert \frac{e^{(-z^2/2)}}{\sqrt{2\pi}}\,\mathrm{d}z \;=\; \sqrt{\frac{2}{\pi}} = 0.79788 \qquad (38)$$
$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2\, e^{(-z^2/2)}\,\mathrm{d}z \;=\; 1 \qquad (39)$$	
$$\frac{1}{2\pi} \int_{-\infty}^{\infty} z^3\, e^{(-z^2/2)}\,\mathrm{d}z \;=\; 0 \qquad (40)$$	$$\int_{-\infty}^{\infty} \lvert z\rvert^3 \frac{e^{(-z^2/2)}}{\sqrt{2\pi}}\,\mathrm{d}z \;=\; 2\sqrt{\frac{2}{\pi}} \;=\; 1.5958 \qquad (41)$$

Table 1. The means of various powers of z and $\lvert z\rvert$, where z is standard normal, i.e., it has the distribution given in (36). ($\mu = 0\,,\ \sigma = 1$).

Suppose that the three-dimensional gap between two points is measured in N different assemblies. Each measurement produces three numbers x_{1s}, x_{2s}, x_{3s}, with $s = 1, 2, ..., N$. Suppose this produces the following data:

s	x_{1s}	x_{2s}	x_{3s}	r_s
1	x_{11}	x_{21}	x_{31}	$\sqrt{{x_{11}}^2 + {x_{21}}^2 + {x_{31}}^2}$
2	x_{12}	x_{22}	x_{32}	$\sqrt{{x_{12}}^2 + {x_{22}}^2 + {x_{32}}^2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	x_{1N}	x_{2N}	x_{3N}	$\sqrt{{x_{1N}}^2 + {x_{2N}}^2 + {x_{3N}}^2}$

Table 2. Mutually orthogonal gap coordinates measured in N different assemblies.

The value x_{is} denotes the sth measured value of the random variable X_i in coordinate direction i, where $i = 1,2,3$. Assume that the coordinate errors are independent and normally distributed, $X_i \sim \mathrm{N}(0,\sigma^2)$, with equal component variance σ^2. Define the standardized random variables

$$Z_i = \frac{X_i}{\sigma} \sim \mathrm{N}(0,1) \tag{42}$$

whose measured values are

$$z_{is} = \frac{x_{is}}{\sigma}.$$

The last column in the table is a sample from another variable R given by

$$R = \sqrt{X_1^2 + X_2^2 + X_3^2}, \qquad r_s = \sqrt{\sum_{i=1}^{3} x_{is}^2}.$$

As a generalization, the "gap" could exist mathematically in n mutually orthogonal coordinate directions, with N measurements made in each direction.

3.1. Chi-Squared Variables

Suppose that measurements in the X_i coordinate direction are normally distributed ($X_i \sim \mathrm{N}(0, \sigma^2)$) and that this is true in all three spatial directions $i = 1,2,3$. The sum of squares, normalized using (42), as follows:

$$Q = \chi^2 = \sum_{i=1}^{3} Z_i^2 = Z_1^2 + Z_2^2 + Z_3^2, \tag{43}$$

is a chi-square (χ^2) variable. This equation is written explicitly for the three-dimensional case, so the summation is over the $n = 3$ coordinate directions, not over the N observations. The observation index remains $s = 1,2,...,N$.

The general formula for the probability density function for a χ^2 variable with n degrees of freedom (n dimensions) and unit variance was shown by Johnson & Leone (1977, Vol. 1, p. 166) to be

$$f_Q(Q) = \frac{2^{-n/2}}{\Gamma(n/2)} Q^{n/2-1} e^{-Q/2} \qquad (Q \geq 0) \tag{44}$$

The same chi-square distribution also arises in sample-variance theory, but there the degrees of freedom are determined by sample size (typically $N-1$). In the present problem, Q is instead the squared Euclidean norm of one n-dimensional observation, so its degrees of freedom equal the number n of mutually orthogonal coordinate directions.

3.2. Chi Variables (Norms) with Component Variances All Equal

3.2.1 Definition

We saw that the chi-square distribution is given by (44), and that this is the distribution of, for example, R^2. But for measurements of R, such as runout, it would be helpful to have the distribution of R itself. This is called simply the *chi density function.* (Another example of a chi variable is the standard deviation itself.) The norm R of n-dimensional data is given by the following formula:

$$R = \sqrt{X_1^2 + X_2^2 + ... + X_n^2}. \tag{45}$$

Hole position and shaft runout are two such examples of this (with $n = 2$ since the data lies in a plane).

By (45), the radius R in three dimensions is:

$$R \;=\; \sqrt{\sum_{i=1}^{3} X_i^{\,2}} \;=\; \sigma\sqrt{Q}\,. \tag{46}$$

The density function can be obtained by first noting from (46)

$$Q \;=\; \frac{R^2}{\sigma^2}.$$

The density transformation contains the Jacobian

$$\frac{\mathrm{d}Q}{\mathrm{d}R} \;=\; \frac{2R}{\sigma^2}\,, \tag{47}$$

so the change-of-variable formula gives

$$f_R(R) \;=\; f_Q\!\left(\frac{R}{\sigma^2}\right)\!\left|\frac{\mathrm{d}Q}{\mathrm{d}R}\right|.$$

Substitute this into (44) to obtain (Johnson *et al.* 1994, p. 452):

$$f_R(R) \;=\; 2\frac{2^{-n/2}}{\Gamma(n/2)}\frac{R^{n-1}}{\sigma^n}e^{-R^2/(2\sigma^2)}. \qquad (R>0) \tag{48}$$

This is the general formula for the probability density function of the norm $\bar{R}$ of n-dimensional data centered at the origin when all of its rectangular coordinates X_i are normally distributed. It is the n-dimensional chi distribution. When $n=2$ (data is confined to a plane), (48) is called the Rayleigh probability density function. When $n=3$, it is the Maxwell distribution (Maxwell 1860), and it was first used to describe the distribution of the velocity vector of gases in the kinetic theory of gases.

For convenience, the gamma function in the denominator of (48) is tabulated below at some key points (e.g., Weisstein 2003):

$$\Gamma(1/2)=\sqrt{\pi}\,, \qquad \Gamma(1)=1\,, \qquad \Gamma(3/2)=\sqrt{\pi}/2\,,$$

$$\Gamma(2)=1\,, \qquad \Gamma(5/2)=3\sqrt{\pi}/4\,, \qquad \Gamma(3)=2\,.$$

In most older versions of Excel that do not have a dedicated gamma function such as GAMMA(), it can be implemented by taking the exponential of its GAMMALN() function, which older versions had. So, e.g.,

$\Gamma(n/2)$ = EXP(GAMMALN(n / 2)).

3.2.2 Mean

As noted in (35), the mean of a function can be written

$$\bar{R} = \int_0^\infty R\,f(R)\mathrm{d}R\,. \tag{49}$$

Therefore, weighting (48) by R as stipulated in (49) gives the integral for computing the mean of the n-dimensional distance R from the origin of centered n-dimensional data (Johnson *et al.* 1994, p. 452):

$$\bar{R} \;=\; \frac{2^{1-n/2}}{\Gamma(n/2)\sigma^n}\int_0^\infty r^n\, e^{-r^2/(2\sigma^2)}\,\mathrm{d}r \;=\; \sqrt{2}\,\sigma\,\frac{\Gamma[(n+1)/2]}{2\Gamma(n/2)} \tag{50}$$

Since r is always positive, $\bar{R} > 0$ even though the measurements of x_{is} will be both positive and negative numbers.

Using (50), the mean of R, when R comes from 1, 2, and 3 dimensional data centered at the origin, is computed in Table 3. Here, σ is the common Cartesian component standard deviation.

$(n=1)$	$\bar{R} \;=\; \frac{2^{1/2}}{\Gamma(1/2)\sigma}\int_0^\infty R\, e^{-R^2/(2\sigma^2)}\,\mathrm{d}R \;=\; \sqrt{\frac{2}{\pi}}\,\sigma \;=\; 0.7979\,\sigma$	(51)
$(n=2)$	$\bar{R} \;=\; \frac{1}{\Gamma(1)\sigma^2}\int_0^\infty R^2\, e^{-R^2/(2\sigma^2)}\,\mathrm{d}R \;=\; \sqrt{\frac{\pi}{2}}\,\sigma \;=\; 1.25331\,\sigma$	(52)
$(n=3)$	$\bar{R} \;=\; \frac{1}{\sqrt{2}\Gamma(3/2)\sigma^3}\int_0^\infty R^3\, e^{-R^2/(2\sigma^2)}\,\mathrm{d}R \;=\; 2\sqrt{\frac{2}{\pi}}\,\sigma \;=\; 1.59577\,\sigma$	(53)

Table 3. The mean of the chi variable $R = \sum_{i=1}^{n}\sqrt{x_i^{\,2}}$ in $n = 1$, 2, and 3 dimensions. $(\bar{x}_i = 0)$

Notice that (51) leads to the rule of thumb when working with surface roughness:

$$\frac{R\mathrm{q}}{R\mathrm{a}} \;=\; \frac{\sigma}{\bar{R}} \;=\; \sqrt{\frac{\pi}{2}} \;\approx\; 1.25\,,$$

where Rq is the root mean square of surface profile height deviations, and Ra is the well-known arithmetic mean of their absolute values. It is only a rule of thumb because it assumes normality of surface roughness.

3.2.3 Variance

In order to calculate the variances using the well-known formula

$$V \;=\; \mathrm{E}[R^2] \;-\; \bar{R}^{\,2}\,, \tag{54}$$

we need, in addition to (50), to compute the following means of R^2 by weighting it by the probability in (48) for various dimensions and including the Jacobian:

$$(n=1) \qquad \mathrm{E}[R^2] \;=\; \frac{2^{1/2}}{\Gamma(1/2)\sigma}\int_0^\infty R^2\, e^{-R^2/(2\sigma^2)}\,\mathrm{d}R \;=\; \sigma^2 \tag{55}$$

$$(n=2) \qquad \mathrm{E}[R^2] \;=\; \frac{1}{\Gamma(1)\sigma^2}\int_0^\infty R^3\, e^{-R^2/(2\sigma^2)}\,\mathrm{d}R \;=\; 2\sigma^2 \tag{56}$$

$$(n=3) \qquad \mathrm{E}[R^2] \;=\; \frac{1}{\sqrt{2}\Gamma(3/2)\sigma^3}\int_0^\infty R^4\, e^{-R^2/(2\sigma^2)}\,\mathrm{d}R \;=\; 3\sigma^2 \tag{57}$$

Then the variances of R for 1, 2, and 3 dimensional data can be constructed by substituting (55)–(57) into (54), along with (51)–(53). This gives:

$$(n=1) \qquad \mathrm{Var}(R) \;=\; (\mathrm{E}[R^2]-\bar{R}^2) \;=\; \left(1 \;-\; \frac{2}{\pi}\right)\sigma^2 \;=\; 0.3634\,\sigma^2 \tag{58}$$

$$(n=2) \qquad \mathrm{Var}(R) \;=\; (\mathrm{E}[R^2]-\bar{R}^2) \;=\; \left(2-\frac{\pi}{2}\right)\sigma^2 \;=\; 0.4292\,\sigma^2 \tag{59}$$

$$(n=3) \qquad \mathrm{Var}(R) \;=\; (\mathrm{E}[R^2]-\bar{R}^2) \;=\; \left(3-\frac{8}{\pi}\right)\sigma^2 \;=\; 0.4535\,\sigma^2 \tag{60}$$

The general closed form of the variance for an n-dimensional point cloud is (Faber 2012):

$$\mathrm{Var}(R) \;=\; \sigma^2\left[n-2\frac{[\Gamma((n+1)/2]^2}{[\Gamma(n/2)]^2}\right]. \tag{61}$$

These relations are only valid for data that is centered about zero. But they can be used as a quick sanity check when working with position error of machined holes, for example. The relations can also be used when simulating runout data from the x and y position error. The relations above assume that the variances σ^2 in all 3 directions are equal. The case when this is not true will be addressed in the next section.

3.3. Chi Variables (Norms) with Unequal Variances

Suppose that a random variable R is defined as the following function of three normally distributed random variables X, Y, and Z:

$$R^2 \;=\; X^2 \;+\; Y^2 \;+\; Z^2$$

and that these three have a mean of zero. Writing this relation in terms of expectation,

$$\mathrm{E}[R^2] \;=\; \mathrm{E}[X^2] \;+\; \mathrm{E}[Y^2] \;+\; \mathrm{E}[Z^2]$$

Because each Cartesian component is centered,

$$\mathrm{E}[X^2] \;=\; {\sigma_X}^2\,, \qquad \mathrm{E}[Y^2] \;=\; {\sigma_Y}^2\,, \qquad \mathrm{E}[Z^2] \;=\; {\sigma_Z}^2$$

Therefore, (54) becomes

$$\mathrm{Var}(R) \;=\; {\sigma_X}^2+{\sigma_Y}^2+{\sigma_Z}^2 \;-\; (\mathrm{E}[R])^2\,. \tag{62}$$

The complexity is in the last term. The problem of relating the variance of R to the variances of the three independent random variables when their variances are unequal has been studied by Khatri (1971) and Moschopoulos (1985). Moschopoulos gave an exact infinite gamma-series representation for this problem. So exact representations of the unequal-variance problem are available through the theory of quadratic forms and sums of gamma variates, but they generally involve numerical integration or infinite series. The following compact algebraic approximation was instead obtained empirically from Monte Carlo optimization with 30,000 normally distributed X, Y, and Z data points (90,000 data points in all):

$$\begin{aligned} {\sigma_R}^2 \;=\; & A({\sigma_X}^2+{\sigma_Y}^2+{\sigma_Z}^2) \;+\; B(\sigma_X\,\sigma_Y+\sigma_Y\,\sigma_Z+\sigma_Z\,\sigma_X) \\ & +\; C(\sigma_X\,\sigma_Y\,\sigma_Z)^{2/3} \end{aligned} \tag{63}$$

The leading coefficient A can be determined by letting $\sigma_X=1$ and $\sigma_Y=\sigma_Z=0$ so that $n=1$ (for 1D data), and using (58). This gives

$$A \;=\; 1 \;-\; \frac{2}{\pi} \;\approx\; 0.3634\,.$$

For 2-dimensional data, we have $n = 2$ and set $\sigma_X = \sigma_Y = 1$ in (63). We then find by equating (63) to (59) that

$$B = \frac{4}{\pi} - \frac{\pi}{2} \approx -0.2976 .$$

Finally, for the case when $n = 3$, we set $\sigma_X = \sigma_Y = \sigma_Z = 1$ in (63) and find by equating (63) to (60) that

$$C = \frac{3\pi}{2} - \frac{14}{\pi} \approx 0.25605 .$$

Substituting all these values into (63) yields the following relation between the variances of normally distributed measurements in 3 mutually orthogonal directions and the variance of their resultant distance from the origin (which distance cannot be negative):

$$\begin{aligned} {\sigma_R}^2 = {} & 0.3634({\sigma_X}^2 + {\sigma_Y}^2 + {\sigma_Z}^2) - 0.2976(\sigma_X \sigma_Y + \sigma_Y \sigma_Z + \sigma_Z \sigma_X) \\ & + 0.256(\sigma_X \sigma_Y \sigma_Z)^{2/3} \end{aligned} \tag{64}$$

This shows the interesting result that in two dimensions the standard deviation of R is actually *less than* the standard deviation of its Cartesian components σ_X and σ_Y when $\sigma_X = \sigma_Y$. In 2D, if $\sigma_X = \sigma_Y$, (63) gives:

$$\sigma_R = \sqrt{2 - \frac{\pi}{2}}\, \sigma_X . \qquad \text{(2D)} \tag{65}$$

It is important to keep in mind that (64) and (65) are only valid for data centered about zero. So it is useful for 3-D position tolerancing or in simulating runout measurements.

There is a useful exact check in two dimensions. With unequal zero-mean Gaussian component variances, this is the well-known Hoyt distribution (Hoyt 1947), also called the Nakagami-q. Its mean can be obtained directly. Let

$$a = \max(\sigma_X, \sigma_Y), \qquad b = \min(\sigma_X, \sigma_Y), \qquad m = 1 - \frac{b^2}{a^2} .$$

Then $\mathrm{E}[R]$ in (62) becomes

$$\mathrm{E}[R] = a\sqrt{\frac{2}{\pi}}\, E(m) ,$$

where $E(m)$ is the complete elliptic integral of the second kind given by

$$E(m) = \int_0^{\pi/2} \sqrt{1 - m \sin^2\theta}\, \mathrm{d}\theta .$$

It can be evaluated numerically. Hence

$$\mathrm{Var}(R) = a^2 + b^2 - \frac{2a^2}{\pi}[E(m)]^2 .$$

This allows us to compare the empirical approximation (64), which in this case becomes

$$\mathrm{Var}(R) = \left(1 - \frac{2}{\pi}\right)({\sigma_X}^2 + {\sigma_Y}^2) + \left(\frac{4}{\pi} - \frac{\pi}{2}\right)\sigma_X \sigma_Y ,$$

with the exact value.

Equation (64) is exact at both endpoints—the one-dimensional limit and the isotropic Rayleigh case of $\sigma_X = \sigma_Y$. But between them it is not exactly equal to the elliptic-integral result. In the range $0 \le b/a \le 1$, the largest relative error in the variance is about 5.7%, near

$$\frac{b}{a} \approx 0.23 .$$

For example, with $a = 1$ and $b = 0.23$,

$$\mathrm{Var}(R)_{\text{exact}} \approx 0.3332\,,$$

whereas (64) gives

$$\mathrm{Var}(R)_{(64)} \approx 0.3142\,.$$

In terms of standard deviation,

$$\sigma_{R,\,\text{exact}} \approx 0.5772\,, \qquad \sigma_{R,(64)} \approx 0.5605$$

which is about a 2.9% difference. Therefore, (64) provides a useful, compact three-term approximation for unequal variances.

3.4. Examples

<u>Example 4: Hole Position</u>

Suppose the position tolerance for a hole specified on a drawing is $0.15\,\mathrm{mm}$. The requirement is to create a Monte Carlo simulation of thousands of normally distributed Cartesian position "measurements" that meet this requirement by using

$$\delta X = \text{NORMINV(RAND(), 0, } \sigma_X \text{)} \tag{66}$$

$$\delta Y = \text{NORMINV(RAND(), 0, } \sigma_Y \text{)}. \tag{67}$$

Suppose the manufacturer typically protects for a spread of 5 standard deviations in its drawing callouts (2.5 on either side of nominal). Since position tolerance is a one-sided tolerance, we cannot proceed in that manner. But we still want to protect for the same defect rate in the one-sided position tolerance. Position tolerance is expressed in terms of a diameter around a position, so we will work first with radius

$$R = \sqrt{X^2 + Y^2}\,, \tag{68}$$

which follows a χ distribution. The allowable radius is half the diametrical tolerance

$$r_{\text{Tol}} = \frac{0.15}{2} = 0.075\,\mathrm{mm}\,.$$

The probability density function was given in (48). It was integrated with radius to obtain the mean in (52), but to get the probability, we simply set $n = 2$ and integrate (48) without the radius factor. First, write it for two dimensions:

$$f_R(r) = \frac{r}{\sigma^2} e^{-r^2/(2\sigma^2)}\,.$$

Then integrate from 0 to the allowable radial tolerance r_{Tol}:

$$P(R \le r_{\text{Tol}}) = \int_0^{r_{\text{Tol}}} f_R(r)\mathrm{d}r = \int_0^{r_{\text{Tol}}} \frac{r}{\sigma^2} e^{-r^2/(2\sigma^2)}\mathrm{d}r$$

Fortunately, the integral in two dimensions is much easier than it is in one dimension. There is no need for a spreadsheet function or a table lookup. Since

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[-e^{-r^2/(2\sigma^2)}\right] = \frac{r}{\sigma^2} e^{-r^2/(2\sigma^2)}\,,$$

the integral is simply

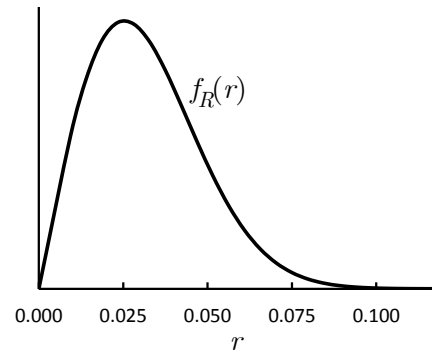
$$P(R \le r_{\text{Tol}}) = 1 - e^{-{r_{\text{Tol}}}^2/(2\sigma^2)}\,.$$

Now let P be the desired probability, and solve for the Cartesian standard deviation σ:

$$\sigma_X = \sigma_Y = \frac{r_{\text{Tol}}}{\sqrt{-2\ln(1-P)}}.$$

The probability associated with $2.5\,\sigma$ on each side of nominal is, by a standard normal probability table, $P = 0.98758$ (with two tails removed). Alternatively, in Excel,

P = 2*NORM.S.DIST(2.5,TRUE) - 1.



Then,

$$\sigma_X = \sigma_Y = \frac{0.075\,\text{mm}}{\sqrt{-2\ln(1-0.98758)}} = 0.0253\,\text{mm}.$$

The standard deviation of R is, by (59),

$$\sigma_R = \sqrt{0.4292}\;0.0253\text{ mm} = 0.01657$$

Therefore, the standard deviation implied by the drawing is $\sigma_R = 0.0166$ mm, and 98.758% of the simulated hole centers will fall inside the ∅0.15-mm position tolerance circle, just as 98.758% of a conventional normal Cartesian dimension lies inside its $\pm 2.5\sigma$ limits.

For comparison, we check to see what the approximation (64) would give us for this case:

$$\sigma_R^{\,2} = 0.3634(\sigma_X^{\,2} + \sigma_Y^{\,2}) - 0.2976\,\sigma_X\,\sigma_Y.$$

If we make the assumption that $\sigma_X = \sigma_Y$ (the next example does not make this assumption), then

$$\sigma_R = \sqrt{0.4292}\;0.0253\text{ mm} = 0.01657 \tag{69}$$

which is the value above.

The Cartesian coordinates can be normal, but the drawing controls their radial resultant, so matching failure probabilities requires transforming standard deviations, which amounts to transforming the probability distribution.

<u>Example 5: Yoke Position / Shaft Runout</u>

Suppose the standard deviations of two features that determine the position of a yoke at the end of a shaft are: $\sigma_X = 0.055$ and $\sigma_Y = 0.040$. Determine the expected standard deviation of shaft runout caused by this position error.

According to (64), the variance of shaft runout will be

$$\begin{aligned}\sigma_R^{\,2} &= 0.3634\,(\sigma_X^{\,2} + \sigma_Y^{\,2}) - 0.2976\,\sigma_X\,\sigma_Y \\ &= 0.001026\end{aligned}$$

So the standard deviation of shaft runout is expected to be:

$$\sigma_R = 0.0320\,\text{mm}.$$

This can again be verified with numerical simulation by creating X and Y data using (66) through (68).

For comparison, the exact Hoyt result discussed above in §3.3 gives $\sigma_R = 0.0322$ mm, so the approximation in (64) differs by only about 0.6% for this example.

3.5. Geometric interpretation of radial probability

Consider a three-dimensional point cloud. The Cartesian coordinates of all the points are centered at zero and normally distributed:

$$X, Y, Z \sim \mathrm{N}(0, \sigma^2)$$

independently. A probability density per unit volume can be assigned to every point in space by writing (36) three times—once for each direction:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/(2\sigma^2)}, \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-y^2/(2\sigma^2)}, \qquad f_Z(z) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-z^2/(2\sigma^2)},$$

and then forming the joint probability. By independence, the joint Cartesian probability density is the product of the three one-dimensional normal densities:

$$p(x, y, z) = f_X(x)\, f_Y(y)\, f_Z(z).$$

Making the substitutions, we have

$$p(x, y, z) = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-(x^2+y^2+z^2)/(2\sigma^2)}$$

or

$$p(r) = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-r^2/(2\sigma^2)}. \tag{70}$$

This has its maximum at $r = 0$. So the probability density per unit volume is greatest at the nominal point. But that is only the density. When it is weighted by volume, as it must be, the most probable radial distance is some positive distance away from the center.

A thin spherical shell between r and $r + \mathrm{d}r$ has volume

$$\mathrm{d}V = 4\pi r^2 \mathrm{d}r.$$

So the probability contained in that shell is

$$\mathrm{d}P = p(r) 4\pi r^2 \mathrm{d}r.$$

The radial probability density is found by substituting (70) into this and dividing by $\mathrm{d}r$:

$$f_R(r) = \frac{\mathrm{d}P}{\mathrm{d}r} = 4\pi r^2 p(r) = \sqrt{\frac{2}{\pi}} \frac{r^2}{\sigma^3} e^{-r^2/(2\sigma^2)},$$

which is just (48) for $n = 3$. Near the origin, the r^2 factor overwhelms the Gaussian distribution.

Differentiating gives the mode

$$r_{\text{mode}} = \sqrt{2}\,\sigma,$$

and the mean is given by (53).

Therefore, the following three statements hold for the spherical Gaussian distribution:

Most probable Cartesian coordinate:	$r = 0$
Most probable radius:	$r = 1.414\,\sigma$
Mean radius:	$r = 1.596\,\sigma$

Analogy of a Hydrogen Atom Electron Shell

Another example of this is the hydrogen atom. Its $1s$ ground-state probability density is spherically symmetric and decreases exponentially away from the nucleus (Condon and Morse 1929, p. 57):

$$p(r) \;=\; \frac{1}{\pi a_0^{\,3}} e^{-2r/a_0} .$$

Its probability density (per unit volume) is greatest at $r = 0$. So in exactly the same sense as the toleranced point cloud, the nucleus is the most probable Cartesian point. But again there is only a tiny volume surrounding $r = 0$. Weighting the spatial density by the spherical-shell volume factor $4\pi r^2$ gives the radial probability density

$$f_R(r) \;=\; \frac{4r^2}{a_0^{\,3}} e^{-2r/a_0} .$$

Its maximum occurs at

$$r_{\text{mode}} \;=\; a_0 .$$

So although the ground-state probability density per unit volume is greatest at the nucleus, the most probable radial distance is the Bohr radius a_0 (Condon and Morse 1929, p. 61-62).

The two cases are summarized in Table 4. Although the underlying spatial distributions differ (Gaussian versus exponential), the same geometric phenomenon occurs in both: The location having the greatest probability density per unit volume corresponds to a radius having zero radial probability density. So even though the tolerance cloud and the electron cloud have entirely different physical origins, both show that the most probable Cartesian point need not lie at the most probable *radius R*. In fact, for a smooth spherically symmetric distribution with finite nonzero density at the origin, the radial density necessarily starts from zero at the origin.

	3-D Gaussian point cloud	Hydrogen $1s$ cloud
Spatial density	$e^{-r^2/(2\sigma^2)}$	e^{-2r/a_0}
Point of max probability density	$r = 0$	$r = 0$
Radial density	$r^2\, e^{-r^2/(2\sigma^2)}$	$r^2\, e^{-2r/a_0}$
Most probable radius	$\sqrt{2\sigma}$	a_0

Table 4. Similarities between functional forms of a point cloud toleranced in three-dimensions and the Hydrogen $1s$ shell.

For the isotropic n-dimensional Gaussian distribution, R has the chi distribution. From the general chi density in (48),

$$f_R(r) \;\propto\; r^{n-1} e^{-r^2/(2\sigma^2)} .$$

Setting its derivative to zero gives

$$r_{\text{mode}} \;=\; \sigma\sqrt{n-1}\,, \qquad\qquad n \geq 1 .$$

The shell containing the greatest radial probability corresponds to the mode of the radial density, not its mean. Its mean is given by (50). Therefore,

$$\mathrm{E}[R] \;=\; \sqrt{2}\,\sigma\,\frac{\Gamma[(n+1)/2]}{2\Gamma(n/2)} \;\sim\; \sigma\sqrt{n}\,, \qquad\qquad r_{\mathrm{mode}} \;=\; \sigma\sqrt{n-1} \;\sim\; \sigma\sqrt{n}\,.$$

So while the mode and mean are not the same quantity, they have the same asymptotic scaling. As dimensionality increases, the point of maximum probability density remains fixed at the origin, while the shell containing the greatest radial probability moves outward as $\sqrt{n-1}\,\sigma$, or approximately as $\sqrt{n}\,\sigma$. At the same time, the mean radius grows without bound approximately as $\sqrt{n}\,\sigma$, while the radial thickness of the shell, as measured by its standard deviation, approaches the fixed value $\sigma/\sqrt{2} \approx 0.707\sigma$. Consequently, the shell becomes vanishingly thin relative to its mean radius; that is,

$$\mathrm{E}[R] \;\sim\; \sigma\sqrt{n}\,, \qquad\qquad \sigma_R \;\to\; \frac{\sigma}{\sqrt{2}} \;\approx\; 0.707\,\sigma\,,$$

so the relative shell thickness behaves roughly as

$$\frac{\sigma_R}{\mathrm{E}[R]} \;\sim\; \frac{1}{\sqrt{2n}} \;\to\; 0\,.$$

Viewed in terms of distance from nominal, therefore, a toleranced point cloud becomes increasingly shell-like as the dimensionality of the problem increases. For a smooth spherically symmetric distribution with finite nonzero density at the origin, the radial density starts from zero.

4. Sensitivity Analysis

Error propagation analysis frequently involves nonlinear functions. The nonlinear function might be of the form of (18) or be something as simple as the volume of a cylinder. Kline and McClintock (1953) proposed the following relation as a means of assessing the influence of uncertainty in measured factors x_i on the uncertainty of a function $F(x_i)$ of those factors:

$$\sigma_F \;=\; \sqrt{\left(\frac{\partial F}{\partial x_1}\right)^2 \sigma_1^{\,2} \;+\; \left(\frac{\partial F}{\partial x_2}\right)^2 \sigma_2^{\,2} \;+\; \ldots \;+\; \left(\frac{\partial F}{\partial x_n}\right)^2 \sigma_n^{\,2}}\,. \tag{71}$$

This has become the standard in textbooks describing sensitivity analysis or error propagation; if $F(x_i)$ is a nonlinear function, however, this relation is generally only a first-order approximation. For nonlinear functions, Greenwood and Chase (1990) discussed using Monte Carlo simulations or the Hasofer-Lind Reliability Index method (Hasofer and Lind 1974). Evans (1975) presented Taylor series terms up to sixth order for the purpose of doing nonlinear error analysis. The discussion below illustrates how the required higher-order terms arise when the variance of a nonlinear function is evaluated. Suppose we have the following nonlinear function F of two random variables X_1 and X_2:

$$F \;=\; X_1^{\,2}\,X_2^{\,2}\,. \tag{72}$$

Since this is a nonlinear function of the random variables, to relate the standard deviation of the independent variables to that of F, we should use higher moments to compute the variance.

Suppose we have a scalar function Y of m component dimensions X_i:

$$Y \;=\; a_1 X_1 \;+\; a_2 X_2 \;+\; \ldots \;+\; a_m X_m \;+\; b\,. \tag{73}$$

The m components could be, for example, parts in a stack path (not repeated observations). Its mean is

$$\mu_Y \;=\; a_1\,\mu_1 \;+\; a_2\,\mu_2 \;+\; \ldots \;+\; a_m\,\mu_m \;+\; b\,,$$

so its centered deviation is

$$\delta Y \;=\; (\,Y \;+\; \mu_1) \;=\; a_1\,\delta x_1 \;+\; ... \;+\; a_m\,\delta x_m\,. \tag{74}$$

The variance is

$$\mathrm{Var}(\,Y) \;=\; \mathrm{E}[(\delta Y)^2] \tag{75}$$

There is nothing further to subtract in (54) because $\mathrm{E}[\delta Y] = 0$. Substituting (74) into the above relation gives

$$\mathrm{Var}(\,Y) \;=\; \mathrm{E}\Big[\big(\textstyle\sum_i a_i\,\delta x_i\big)^2\Big] \;=\; \sum_i\sum_k a_i a_k\,\mathrm{E}[\delta x_i\,\delta x_k] \;=\; \sum_i\sum_k a_i a_k\,\mathrm{Cov}(X_i, X_k)$$

so

$$\mathrm{Var}(\,Y) \;=\; \sum_i\sum_k a_i a_k\,V_{ik}\;,$$

which is (17), but without the additional spatial components indexed by j. So for linear functions,

$$\mathrm{Var}(\,Y) \;=\; \mathrm{E}[(\,Y - \mathrm{E}[\,Y])^2] \;=\; \mathrm{E}[(\delta Y)^2]. \tag{76}$$

Now suppose $F = F(x)$ is nonlinear, and approximate it through second order. And let

$$\delta x = X - \mu\,. \tag{77}$$

Then the Taylor expansion requires at least one more term:

$$F(x) \;\approx\; F(\mu) \;+\; F_{,1}\,\delta x \;+\; \frac{1}{2}F_{,11}\,\delta x^2\,. \tag{78}$$

Here, $\partial F/\partial x_1$ is abbreviated as $F_{,1}$ and so forth.

Now take the expectation of both sides:

$$\mathrm{E}[F] \;\approx\; F(\mu) \;+\; F_{,1}\,\mathrm{E}[\delta x] \;+\; \frac{1}{2}F_{,11}\,\mathrm{E}[\delta x^2] \tag{79}$$

The first-order term has zero mean:

$$\mathrm{E}[F_{,1}\,\delta x] \;=\; F_{,1}\,\mathrm{E}[\delta x] \;=\; 0\;, \tag{80}$$

but the second-order term does not:

$$\mathrm{E}[(\delta x)^2] = \sigma^2\,. \tag{81}$$

This holds for any distribution, not only normal distributions. Therefore, (79) becomes

$$\mathrm{E}[F] \;\approx\; F(\mu) \;+\; \frac{1}{2}F_{,11}\,\sigma^2 \tag{82}$$

Therefore, scatter can shift the mean of a nonlinear function, and the shift is the term on the right.

By similar reasoning, with $\delta x = X - \mu$, the central and raw moments are

$$\mathrm{E}[\delta x] = 0\,, \qquad \mathrm{E}[X] = \mu \tag{83}$$

$$\mathrm{E}[\delta x^2] = \sigma^2\,, \qquad \mathrm{E}[X^2] \;=\; \mu^2 \;+\; \sigma^2 \tag{84}$$

$$\mathrm{E}[\delta x^3] = 0\,, \qquad \mathrm{E}[X^3] \;=\; \mu^3 \;+\; 3\mu\sigma^2 \qquad \text{(Gaussian)} \tag{85}$$

$$\mathrm{E}[\delta x^4] = 3\sigma^4\,, \qquad \mathrm{E}[X^4] \;=\; \mu^4 + 6\,\mu^2\,\sigma^2 + 3\,\sigma^4 \qquad \text{(Gaussian)} \tag{86}$$

The next step is to use (54) to compute the variance. Let's examine the question of how far each realization of F lies from that shifted mean. Subtract (82) from (78) and obtain

$$F \;-\; \mathrm{E}[F] \;\approx\; F_{,1}\,\delta x \;+\; \frac{1}{2}F_{,11}(\delta x^2 - \sigma^2)\,. \tag{87}$$

This is the approximation to the centered output deviation of F from its mean: $\delta F \equiv F - \mathrm{E}[F]$. Consequently, according to (76),

$$\mathrm{Var}(F) \;=\; \mathrm{E}[(F - \mathrm{E}[F])^2]\,.$$

Therefore, squaring (87) and taking its expectation gives

$$\begin{aligned}\mathrm{Var}(F) \;&\approx\; \mathrm{E}\Big[\Big[F_{,1}\,\delta x \;+\; \frac{1}{2}F_{,11}(\delta x^2 - \sigma^2)\Big]^2\Big] \\ &=\; F_{,1}{}^2\,\mathrm{E}[\delta x^2] \;+\; F_{,1}F_{,11}\,\mathrm{E}[\delta x(\delta x^2 - \sigma^2)] \;+\; \frac{1}{4}F_{,11}{}^2\,\mathrm{E}[(\delta x^2 - \sigma^2)^2]\,.\end{aligned} \tag{88}$$

The first term, according to (81), is simply $\mathrm{E}[\delta x^2] = \sigma^2$.

Using (85) and (83), the middle term is, for a Gaussian variable,

$$\mathrm{E}[\delta x^3 - \sigma^2\delta x] \;=\; \mathrm{E}[\delta x^3] \;-\; \sigma^2\mathrm{E}[\delta x] \;=\; 0\,.$$

The last term is

$$\begin{aligned}\mathrm{E}[(\delta x^2 - \sigma^2)^2] \;&=\; \mathrm{E}[\delta x^4] \;-\; 2\sigma^2\mathrm{E}[\delta x^2] \;+\; \sigma^4 \\ &=\; 3\sigma^4 \;-\; 2\sigma^4 \;+\; \sigma^4 \\ &=\; 2\sigma^4\end{aligned}$$

Here, (86) and (84) were used. Therefore, (88) simplifies to

$$\mathrm{Var}(F) \;\approx\; F_{,1}{}^2\sigma^2 \;+\; \frac{1}{2}F_{,11}{}^2\,\sigma^4 \tag{89}$$

Now generalize this for two variables. Instead of (77), we now have

$$\delta x_1 = X_1 - \mu_1\,, \qquad\qquad \delta x_2 = X_2 - \mu_2\,. \tag{90}$$

Instead of (87), the centered second-order Taylor approximation is now

$$F \;-\; \mathrm{E}[F] \;\approx\; F_{,1}\,\delta x_1 \;+\; F_{,2}\,\delta x_2 \;+\; \frac{1}{2}F_{,11}(\delta x_1{}^2 - \sigma_1{}^2) \;+\; \frac{1}{2}F_{,22}(\delta x_2{}^2 - \sigma_2{}^2) \;+\; F_{,12}\,\delta x_1\delta x_2\,.$$

For independent Gaussian variables, the surviving terms give, as before,

$$\mathrm{Var}(F) \;\approx\; F_{,1}{}^2\sigma_1{}^2 \;+\; F_{,2}{}^2\sigma_2{}^2 \;+\; \frac{1}{2}F_{,11}{}^2\,\sigma_1{}^4 \;+\; \frac{1}{2}F_{,22}{}^2\,\sigma_2{}^4 \;+\; F_{,12}{}^2\,\sigma_1{}^2\,\sigma_2{}^2 \tag{91}$$

or

$$\sigma_F \;\approx\; \sqrt{(F_{,1}\sigma_1)^2 \;+\; (F_{,2}\sigma_2)^2 \;+\; \frac{1}{2}(F_{,11}\,\sigma_1{}^2)^2 \;+\; \frac{1}{2}(F_{,22}\,\sigma_2{}^2)^2 \;+\; (F_{,12}\,\sigma_1\,\sigma_2)^2} \tag{92}$$

Figure 8 shows a comparison of the actual (simulated) standard deviation σ_F of F in (72) with the conventional standard deviation estimate (71) and that obtained by the second-order relation (92). Each point in Figure 8 represents a standard deviation computed from 30,000 independently generated data points with a normal distribution. The second-order relation gives better agreement with the simulation than the customary linear relation (71), although its error becomes substantial as the input coefficients of variation increase.

As Figure 8 shows, approximations need the standard deviations to be controlled so that the coefficient of variation CV is much less than 1, as follows:

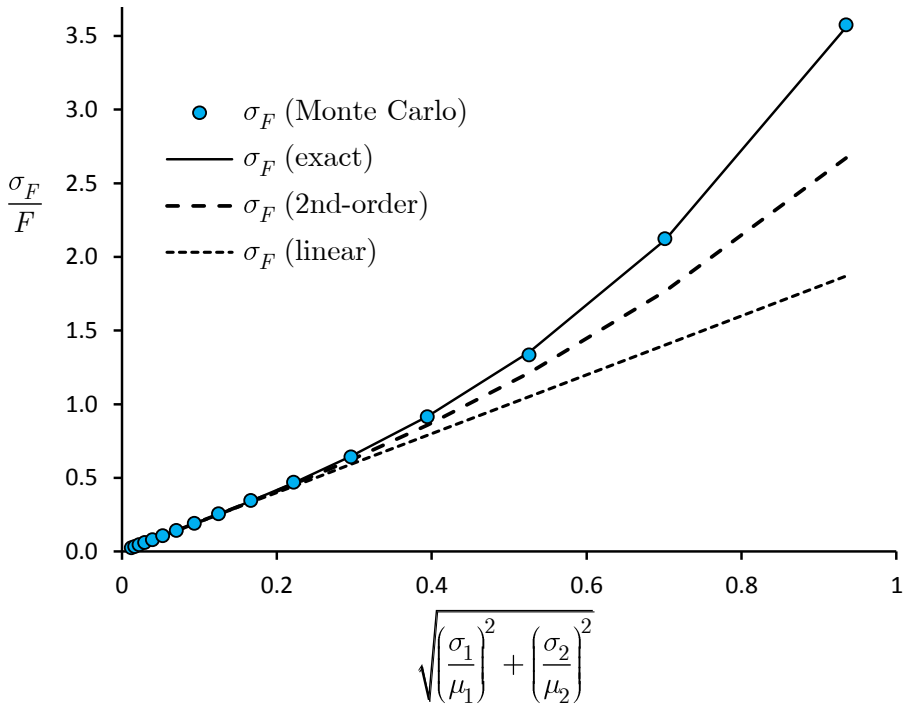


Figure 8. Comparison of standard deviation of F defined in (72) (with normally distributed Monte Carlo simulation data for σ_1 and σ_2) against the linear formula (71), the 2nd-order approximation (92), and exact moment relation (93).

$$CV = \sqrt{\left(\frac{\sigma_1}{\mu_1}\right)^2 + \left(\frac{\sigma_2}{\mu_2}\right)^2 + \ldots + \left(\frac{\sigma_n}{\mu_n}\right)^2} \ll 1 .$$

To obtain the exact relation, take the expected value of (72):

$$\mathrm{E}[F] \;=\; \mathrm{E}[X_1^2]\mathrm{E}[X_2^2] \;=\; (\mu_1^2 + \sigma_1^2)(\mu_2^2 + \sigma_2^2) ,$$

and note that

$$F^2 \;=\; X_1^4 X_2^4 .$$

Therefore,

$$\mathrm{E}[F^2] \;=\; \mathrm{E}[X_1^4]\mathrm{E}[X_2^4] .$$

By (86), $\mathrm{E}[X_i^4] \;=\; \mu_i^4 + 6\mu_i^2\sigma_i^2 + 3\sigma_i^4$. Therefore the exact form is

$$\sigma_F \;\approx\; \sqrt{(\mu_1^4 + 6\mu_1^2\sigma_1^2 + 3\sigma_1^4)(\mu_2^4 + 6\mu_2^2\sigma_2^2 + 3\sigma_2^4) \;-\; (\mu_1^2 + \sigma_1^2)^2(\mu_2^2 + \sigma_2^2)^2} \qquad (93)$$

The exact relation (93) lies almost exactly on the simulation points in Figure 8. For example, for the point on the far-right side of the graph, where

$$\sigma_1 = 7.48029 \qquad \text{and} \qquad \sigma_2 = 11.22043$$

and with the means set to $\mu_1 = 10$ and $\mu_2 = 20$, $F = 10^2\,20^2 = 40{,}000$.

The exact (93) gives

$$\sigma_F^{\text{exact}}/F \;\approx\; 3.5593 ,$$

whereas the Monte Carlo simulation gives

$$\sigma_F^{\mathrm{MC}}/F \approx 3.5755\,.$$

So the error between the simulations and the exact relation (93) is about 0.46%.

5. Conclusion

The statistics of measured Cartesian variables need not be the statistics of the engineering quantity ultimately specified. In multidimensional tolerancing, this distinction arises through spatial reduction (computing distance), process ordering, and nonlinear transformation. For radial quantities such as position error and runout, normally distributed Cartesian errors become chi-distributed when reduced to their distance from nominal. Consequently, a point cloud whose Cartesian probability density is greatest at nominal nevertheless has a positive most-probable radius and a positive mean radius. Examples in two dimensions are radial runout and position. As dimensionality increases, the mean radius grows as $\sqrt{n}\,\sigma$, while the radial standard deviation approaches the fixed value $\sigma/\sqrt{2}$, so that the distribution becomes increasingly shell-like. An exact formula is given that relates the variance of a gap in three dimensions (a centered three-dimensional point cloud) to the variances of three independent normally distributed Cartesian components when those variances are equal (61), and a compact algebraic approximation is given for the case when those variances are unequal (64).

Process order matters for many engineering statistics, as illustrated by the casting-and-machining example. When distinct manufacturing operations are correlated and sequential, a different information loss can occur: conventional symmetric covariance does not retain the physical ordering of the processes. The casting-and-machining example illustrates a case in which the reverse sequence is not physically realized. A history-conditioned stochastic process-operator framework was therefore outlined as a possible means of retaining process-order information while preserving the conventional symmetric definition of covariance.

A second inadvertent loss of information can arise in nonlinear error propagation when the response is represented only by its first-order terms. The nonlinear example shows how higher-order central moments enter error propagation and how a second-order approximation improves upon conventional linear propagation. For polynomial response functions, exact moment relations may also be obtained when the required input moments are known.

Spatial reduction, process ordering, and nonlinear transformation therefore affect statistical tolerancing in fundamentally different ways. Reduction to a radial measure is often intentional because the engineering specification itself concerns distance rather than direction; its altered statistics must nevertheless be recognized. By contrast, loss of process ordering through symmetric covariance and loss of higher-order effects through linearized error propagation can occur inadvertently. In each case, the statistical representation must preserve the information required by the engineering quantity of interest.

References

Abellan-Nebot, J. V., Liu, J., & Romero Subiron, F., "Design of multi-station manufacturing processes by integrating the stream-of-variation model and shop-floor data," *Journal of Manufacturing Systems* **30** (2011) pp. 70-82.

Bender, A., Jr., "Statistical Tolerancing as it Relates to Quality Control and the Designer (6 times 2.5 = 9)," *SAE Technical Paper 680490* (1968) pp. 1965-1971.

Condon, E. U., & Morse, P. M., *Quantum Mechanics*, McGraw-Hill, New York (1929).

Evans, D. H., "Statistical Tolerancing: The State of the Art. Part II. Methods for Estimating Moments," *Journal of Quality Technology* **7** (1975) pp. 1-12.

Faber, M. H., *Statistics and Probability Theory: In Pursuit of Engineering Decision Support*, Springer, Dordrecht (2012).

Gao, J., Chase, K. W., & Magleby, S. P., "Generalized 3-D tolerance analysis of mechanical assemblies with small kinematic adjustments," *IIE Transactions* **30** (1998) pp. 367-377.

Gilson, J., *A New Approach to Engineering Tolerances*, Machinery Publishing Co, London (1951).

Greenwood, W. H. & Chase, K. W., "Root Sum Squares Tolerance Analysis with Nonlinear Problems," *Journal of Engineering for Industry* **112** (1990) pp. 382-384.

Hasofer, A. M., & Lind, N. C., "Exact and Invariant Second-Moment Code Format," *Journal of the Engineering Mechanics Division, Proceedings of the American Society of Civil Engineers* **100** (1974) pp. 111-121.

Hong, Y. S. & Chang, T. C., "A comprehensive review of tolerancing research," *International Journal of Production Research* **40** (2002) pp. 2425-2459.

Hoyt, R. S., "Probability Functions for the Modulus and Angle of the Normal Complex Variate," *Bell System Technical Journal* **26** (1947) pp. 318-359.

Huang, Q., & Shi, J., "Stream of Variation Modeling and Analysis of Serial-Parallel Multistage Manufacturing Systems," *Journal of Manufacturing Science and Engineering* **126** (2004) pp. 611-618.

Johnson, N. L. & Leone, F. C., *Statistics and Experimental Design in Engineering and the Physical Sciences*, Two Volumes, Wiley, New York (1977).

Johnson, N. L., Kotz, S., & Balakrishnan, N., *Continuous Univariate Distributions, Vol. 1*, 2nd ed., Wiley, New York (1994).

Khatri, C. G., "Series representations of distributions of quadratic form in the normal vectors and generalised variance," *Journal of Multivariate Analysis* **1** (1971) pp. 199-214.

Kline, S. J. & McClintock, F. A., "Describing uncertainties in single-sample experiments," *Mechanical Engineering* **75** (1953) pp. 3-8.

Liu, J., Shi, J., & Hu, S. J., "Quality Assured Setup Planning Based on the Stream-of-Variation Model for Multi-Stage Machining Processes," *IIE Transactions* **41** (2009) pp. 323-334.

Maxwell, J. C., "Illustrations of the Dynamical Theory of Gases.—Part I. On the Motions and Collisions of Perfectly Elastic Spheres," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *Series 4*, **19** (1860) pp. 19-32.

Morse, E., Dantan, J.-Y., Anwer, N., Söderberg, R., Moroni, G., Qureshi, A., Jiang, X., & Mathieu, L., "Tolerancing: Managing uncertainty from conceptual design to final product," *CIRP Annals* **67** (2018) pp. 695-717.

Moschopoulos, P. G., "The distribution of the sum of independent gamma random variables," *Annals of the Institute of Statistical Mathematics* **37** (1985) pp. 541-544.

Weisstein, E. W., *CRC Concise Encyclopedia of Mathematics*, 2nd ed., Chapman Hall, Boca Raton (2003).

Xing, Y., Gravell, B., He, X., Johansson, K. H., & Summers, T. H., "Identification of linear systems with multiplicative noise from multiple trajectory data," *Automatica* **144** (2022) 110486.

Yang, D. & Gong, J., "Visualized Equivalent Variational Modeling in Tolerance Analysis of 3D Mechanical Assemblies," *Advanced Materials Research* **201-203** (2011) pp. 229-233.

Zeng, L., & Zhou, S., "Inferring the Interactions in Complex Manufacturing Processes Using Graphical Models," *Technometrics* **49** (2007) pp. 373-381.

Zhou, S., Huang, Q., & Shi, J., "State space modeling of dimensional variation propagation in multistage machining process using differential motion vectors," *IEEE Transactions on Robotics and Automation* **19** (2003) pp. 296-309.