INT. J. PROD. RES., 2002, VOL. 40, NO. 2, 337–352

# Statistical tolerance analysis for non-normal or correlated normal component characteristics

J. P. KHAROUFEH†* and M. J. CHANDRA‡

The paper presents a simple approximation technique for statistical tolerance analysis, namely, the allocation of component tolerances based on a known assembly tolerance. The technique utilizes a discretized, multivariate kernel density estimate and a simple transformation to approximate the probability distribution of the overall assembly characteristic. The data-driven approach is suitable for real-world settings in which components are randomly selected from their respective manufacturing processes to form mechanical assemblies. Demonstrated is the numerical approach in two dimensions for two distinct cases: first, when component characteristics are non-normal, independent random variables, and second, when they are highly correlated, normal random variables. The results are promising in initial test problems.

## 1. Introduction

Many mechanical assemblies consist of two or more components or subassemblies. Owing to variation in machine performance, operator performance, or properties of the work piece, it is impossible to avoid variation in component characteristics completely. Naturally, variation in the components leads to variation in the overall assembly; thus, the component characteristics (and the assembly characteristic) may be considered as continuous random variables. When the assembly characteristic is a linear function of the component characteristics, normality and independence assumptions allow for a simple calculation of the component tolerances to yield a desirable acceptance rate for the assemblies. However, in many situations, neither the component nor assembly characteristic probability distributions are known. Furthermore, underlying dependencies may exist between component characteristics that distort the results of the aforementioned analysis.

This paper illustrates an approximation technique for statistically allocating component tolerances, based on a prespecified assembly tolerance, when the probability distributions of component and assembly characteristics are not known or when the component characteristics are normally distributed and correlated. It is further assumed that the assembly characteristic is a linear function of the component characteristics or a form that may be linearized. The next section reviews some of the past and current literature in the area of statistical tolerancing analysis; section 3 presents the approach to the problem, while section 4 gives a review of

Revision received May 2001.

† Department of Operational Sciences, Air Force Institute of Technology, AFIT/ENS, 2950 P Street, Wright-Patterson AFB, OH 45433-7765, USA.

‡ Department of Industrial and Manufacturing Engineering, Pennsylvania State University, 205 Leonhard Building, University Park, PA 16802, USA.

* To whom correspondence should be addressed. e-mail: Jeffrey.Kharoufeh@afit.edu

*International Journal of Production Research* ISSN 0020–7543 print/ISSN 1366–588X online 
http://www.tandf.co.uk/journals
DOI: 10.1080/00207540110079112

kernel density estimation; section 5 presents numerical examples demonstrating the technique when the assembly is comprised of two components, and section 6 gives concluding remarks and directions for future research.

## 2. Review of past work and problem statement

The focus of the work is on statistical tolerancing analysis. Some other broadly defined tolerancing techniques include manual tolerancing based on the dimensional tolerancing chain, computer-aided tolerancing techniques (Bjorke 1989), and tolerancing-based geometrical modelling. As noted by Zhang and Huq (1992), software implementations of these techniques have been developed. In the usual way, it is assumed that the overall assembly characteristic, $X$, is given as a function of the component characteristics, $X_i$, $i = 1, 2, \ldots, k$ of the form

$$X = \psi(X_1, X_2, \ldots, X_k), \tag{1}$$

where $X$ is assumed to have density $f$, while $X_i$ has density $f_i$ for $i = 1, 2, \ldots, k$. In most analyses, it is assumed that the component characteristics are independent random variables with known densities. However, characterization of the density, $f$, is difficult in the absence of these assumptions. For this reason, techniques for finding the moments of the distribution have been proposed.

Many authors, such as Burr (1958), Mansoor (1963), Bender (1968) and Evans (1974), have considered the problem when $\psi$ is linear. Evans (1974), Nigam and Turner (1995) and Barker (1989) discussed the linear propagation of errors or stack tolerancing in which the first two moments of the distribution of $X$ are determined by partial differentiation of $\psi$ with respect to each component characteristic. Mansoor (1963) considered the tolerance problem in engineering design and assumed that the overall assembly tolerance is prespecified by customer or engineering design requirements. His objective was to allocate the component tolerances so that the assembly tolerance objective was met, yielding an acceptable rejection rate. The technique resulted in a probabilistic allocation of component tolerances, but was heavily contingent upon the assumption of normality and independence of component characteristics. Parkinson (1982) extended Mansoor's work by estimating the probability that the component characteristics would fail to meet the dimensional requirements of the assembly characteristic. Parkinson (1984) optimized component tolerances so that, for a given manufacturing cost, the probability of components failing to meet the assembly specification limits was below some acceptable level. The technique requires the component characteristics to follow either a normal, truncated normal or mixed normal distribution.

When the function $\psi$ is non-linear, it is necessary to employ the non-linear propagation of errors or extended Taylor series approximation. Evans (1974) showed that if the derivatives of $\psi$ were tractable, then the first four moments of the distribution of $X$ may be found analytically without too much difficulty. However, when $\psi$ is unknown analytically, one must resort to numerical integration to evaluate the moments of the distribution. Evans (1974) and Nigam and Turner (1995) gave the appropriate quadrature expression for the mean of $X$; however, the technique assumes independence of the component characteristics. In using the quadrature method, one advantage is that the distribution of component characteristics need not be known, only their lower moments. D'Errico and Zaiano (1988) described a modification to Taguchi's method for determining the moments of the distribution of $X$. Specifically, component distributions were assumed to be indepen-

dent and normal, but were estimated by a discrete distribution at three points. The method uses a modified (Gaussian) quadrature approach that outperforms Taguchi's method, at least for the linear case, by correctly specifying the fifth moment of the distribution. More recently, the concept of distribution function (DF) zones has been introduced into the statistical tolerancing literature (cf. O'Connor and Srinivasan 1997). In this approach, a population of components is deemed acceptable if its distribution function is bounded above and below by some specific distribution functions. This part-level technique has been extended by Zhang *et al.* (1999) to a statistical tolerance synthesis model that additionally minimizes the manufacturing cost. The DF zone approach requires that component characteristics are statistically independent.

The present paper contributes to the literature on statistical tolerancing analysis by considering two separate cases in the allocation of component tolerances when the assembly tolerance is fixed. First, it considers the case in which a small number of component characteristics are independent and follow non-normal distributions. In the second case, it considers normally distributed component characteristics that are highly correlated. We present an approximation technique to allocate the tolerance of the assembly characteristic among $k$ components so that an overall desired assembly tolerance may be achieved when $\psi$ is a linear function of the $k$ component characteristics. The data-driven technique uses a non-parametric estimate of the joint density of component characteristics to approximate the distribution function of $X$ rather than assuming a parametric distribution.

## 3. Approximation technique

The following assumptions are made.

- The assembly characteristic, $X$, and component characteristics, $X_i$, $i = 1, 2, \ldots, k$ are of the Nominal-the-Better type (N-type).
- $X_i$ and $X_j$, $i \neq j$ are, in general, correlated random variables for $i$, $j = 1, 2, \ldots, k$.
- Components are randomly assembled.
- Component characteristics follow some unknown probability distribution.
- The process that generates $X_i$ is adjusted so that the mean of $X_i$, $\mu_i$, equals the nominal size of $X_i$, $i = 1, 2, \ldots, k$.
- The proportion of $X_i$ falling within the specification limits is $1 - p_i$ for $i = 1, 2, \ldots, k$.
- The proportion of $X$ falling within the specification limits is $1 - p$.
- The assembly tolerance $T$ is specified by the customer or by engineering requirements.
- The specification limits need not be equidistant from the nominal dimension of the characteristic.

The natural process capability of the process that generates the components and the assembly is a proportion of the range of all possible values that the component or assembly characteristics may assume. Let $t_i$ denote the natural process capability of the process that generates component characteristic $X_i$ and let $t$ denote the natural process capability for the process that generates the assembly characteristic $X$. Now

for any component or assembly characteristic distribution, the natural process capability may be given as a function of the process standard deviation by

$$t_i = c_i\sigma_i, \quad i = 1, 2, \ldots, k \tag{2}$$

and

$$t = c\sigma, \tag{3}$$

where $c_i$ and $c$ are constants, $\sigma_i$ is the standard deviation for the process that generates component $i$, and $\sigma$ is the standard deviation for the process that generates the assembly characteristic. The values of $c_i$ and $c$ dictate the proportion of the total distribution range that is covered by the natural process capability and depend upon the distribution shape. For instance, if the component characteristics are normally distributed, then the assembly characteristic is also normally distributed and $c = 6$ yields a range of values for the assembly characteristic that covers 99.73% of the entire distribution range (0.135% on each tail). If $c = 8$, then 99.994% of the distribution is covered. Let $T_i$ denote the tolerance of component $i$ and let $T$ denote the assembly tolerance. The process capability ratio, denoted by $C_p$, for either a component or assembly characteristic is the ratio of the width of the tolerance interval to the natural process capability (e.g. $C_p = T/t$). It is assumed here that the manufacturer specifies a target process capability ratio, $C_p^*$, for the component and the assembly characteristics. Using the substitution $g = cC_p^*$ and $g_i = c_iC_p^*$, $i = 1, 2, \ldots, k$ it is seen that

$$\begin{aligned} T &= U - L = g\sigma \\ T_i &= U_i - L_i = g_i\sigma_i \quad i = 1, 2, \ldots, k, \end{aligned} \tag{4}$$

where $U$ and $L$ are the upper and lower specification limits of the assembly, respectively, and $U_i$ and $L_i$ are the upper and lower specification limits for component $i$, respectively. It can be seen by (4) that the component tolerances are not equal in general, but rather, each is uniquely determined by $g$ or $g_i$, which directly depends on the distribution of the component or assembly characteristic. The following subsection relates the prespecified assembly tolerance to the $k$ component tolerances while the means for computing the constant $g$ and $g_i$, $i = 1, 2, \ldots, k$ will be presented in section 3.2.

### 3.1. *Relating assembly and component tolerances*

Let $\mu$ and $\sigma^2$ denote the mean and variance of the assembly characteristic $X$, respectively. For any set of component characteristics, whether correlated or uncorrelated, it is well known that

$$\mu = \sum_{i=1}^{k} \mu_i. \tag{5}$$

The variance of the assembly characteristic is given by (Freund and Walpole 1987)

$$\sigma^2 = \sum_{i=1}^{k} a_i^2\sigma_i^2 + 2\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} a_ia_j\sigma_{i,j}, \tag{6}$$

where $a_i$ is a constant, $\sigma_i^2$ is the variance of the $i$th component characteristic and $\sigma_{i,j}$ is the covariance of $X_i$ and $X_j, i \neq j$. In case the component characteristics are

independent, $a_i = 1$ and $\sigma_{i,j} = 0$ for all $i,\ j = 1, 2, \ldots, k$. The fifth assumption of the problem asserts that the process generating component $i$ is adjusted so that the mean $\mu_i$ is equal to the nominal dimension, say $d_i$. However, it is likely that during process setting, the operator may not be able to adjust the process mean to the nominal setting $d_i$. To account for this phenomenon, one may adopt the approach of Mansoor (1963) who assumed a 'worst-case' tolerance wherein the $i$th mean can shift no more than a fixed distance from the nominal value. This mathematically imposes the conditions

$$|\mu_i - d_i| \leq s_i, \quad i = 1, 2, \ldots, k, \tag{7}$$

where $s_i$ denotes the maximum allowable deviation from the nominal dimension of the $i$th component. It can easily be shown (Mansoor 1963) that the assembly tolerance specified by design engineers to reflect these modifications would be given by

$$T' = T - 2\sum_{i=1}^{k} s_i. \tag{8}$$

Equation (8), in effect, would reduce the amount of tolerance that may be allocated among the $k$ component characteristics. Now, substituting (6) into (4), we obtain

$$T = g\left(\sum_{i=1}^{k} a_i^2 \frac{T_i^2}{g_i^2} + 2\sum_{i=1}^{k-1}\sum_{j=i+1}^{k} a_i a_j \sigma_{i,j}\right)^{1/2} \tag{9}$$

Equation (9) provides one equation relating the tolerance of the assembly characteristic to the component characteristics, $T_i,\ i = 1, 2, \ldots, k$. Thus, we require $k - 1$ additional equations to solve for the component tolerances, and these are obtained by the assumption that the target process capability ratio, $C_p^*$, is equal for each of the processes generating the $k$ component characteristics. By (2) and (3) and the substitutions $g = cC_p^*$ and $g_i = c_i C_p^*,\ i = 1, 2, \ldots, k$, this assumption implies

$$\frac{T_1}{g_1\sigma_1} = \frac{T_2}{g_2\sigma_2} = \cdots = \frac{T_k}{g_k\sigma_k} = 1. \tag{10}$$

Now, (9) combined with any $k - 1$ equations in (10) constitutes a set of $k$ equations that may be solved simultaneously to allocate probabilistically the assembly tolerance $T$ among the component tolerances, $T_i,\ i = 1, 2, \ldots, k$. However, this allocation must take into account the desired proportion of rejects for the components and the final assembly. The following subsection describes the procedure for finding $g$ and $g_i$, $i = 1, 2, \ldots, k$, for a desired proportion of rejects and for allocating the assembly tolerance among the component tolerances.

### 3.2. *Tolerance allocation procedure*

Let the cumulative distribution functions (CDFs) of $X_i,\ i = 1, 2, \ldots, k$ and $X$ be $F_{X_i},\ i = 1, 2, \ldots, k$ and $F_X$, respectively. Define $p_u^{(i)}$ as the proportion of undersized components of type $i$ and define $p_o^{(i)}$ as the proportion of oversized components of the same type. The total proportion of rejects for the process that generates the $i$th component is given by $p_i = p_u^{(i)} + p_o^{(i)}$. Similarly, denote the proportion of oversized assemblies by $p_o$, and the proportion of undersized assemblies by $p_u$ such that $p = p_u + p_o$, where $p$ denotes the total proportion of rejects for the assembly. The upper and lower specification limits need not be equidistant from the mean and the distribution may not be symmetric; hence, the oversized and undersized proportions

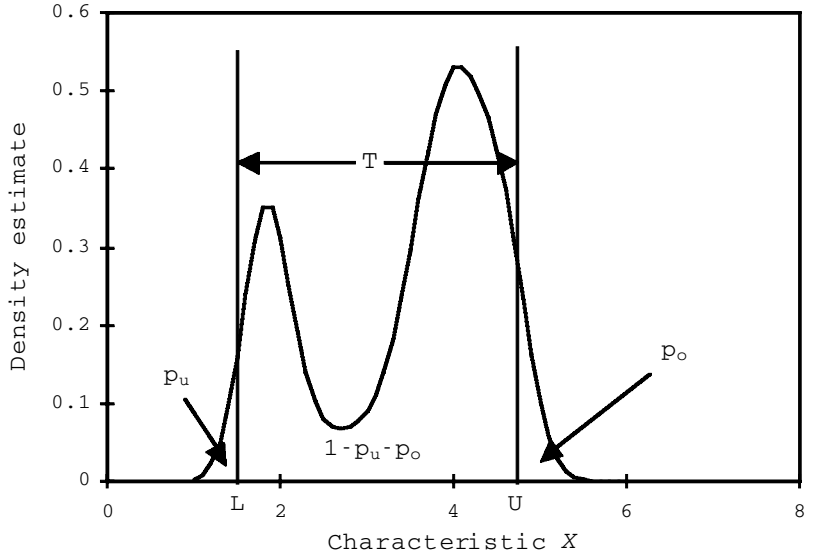


Figure 1. Depiction of the variables for an arbitrary distribution.

need not be equal. Figure 1 depicts the variable definitions for an arbitrary assembly characteristic distribution.

Equations (9) and (10) indicate that $g_i$, $i = 1, 2, \ldots, k$ and $g$ are needed to allocate the assembly tolerance $T$ among the $k$ component tolerances. The value of $g_i$, $i = 1, 2, \ldots, k$ is obtained from the cumulative distribution function of component characteristic $i$, namely $F_{X_i}$, $i = 1, 2, \ldots, k$, while $g$ is obtained from the assembly characteristic CDF, $F_X$, as outlined in the following procedure description.

*Step* 0. Choose acceptable values $p_u^{(i)}$, $p_o^{(i)}$, $p_u$, $p_o$ and $C_p^*$.
*Step* 1. Compute $c$ and $c_i$ for the assembly and each component characteristic by

$$c = \frac{t_u - t_l}{\sigma} \quad \text{and} \quad c_i = \frac{t_u^{(i)} - t_l^{(i)}}{\sigma_i}, \tag{11}$$

where

$$t_l = \sup\{x : F_X(x) \le p_u\} \tag{12}$$

$$t_l^{(i)} = \sup\{x_i : F_{X_i}(x_i) \le p_u^{(i)}\} \tag{13}$$

$$t_u = \inf\{x : 1 - F_X(x) \le p_o\} \tag{14}$$

$$t_u^{(i)} = \inf\{x_i : 1 - f_{x_i}(X_i) \le p_o^{(i)}\}. \tag{15}$$

*Step* 2. Calculate $g$ and $g_i$ for the assembly and each component characteristic by

$$g = cC_p^* \quad \text{and} \quad g_i = c_i C_p^*. \tag{16}$$

*Step* 3. Obtain $T_i$, $i = 1, 2, \ldots, k$ by simultaneously solving (9) and (10).

The procedure here assumes the cumulative distribution functions of components and the assembly are known *a priori*; however, this is seldom the case in reality. One

alternative to the normality assumption is to utilize data-driven, non-parametric distribution approximations, and one well-known technique is reviewed in section 4.

## 4. Distribution approximations

In the absence of normality (and independence) assumptions of the component characteristics, it is difficult to find an analytical solution for the cumulative distribution function of the assembly characteristic, $X$. Thus, numerical approximations will be employed to carry out this task. In this section, a well-known, non-parametric density estimation technique and its multivariate extension are reviewed to find the distribution of $X$.

### 4.1. *Univariate kernel density estimator*

A brief review of the kernel density estimate (KDE) is presented (cf. Silverman 1986). Suppose a continuous random variable $Y$ has probability density function $f(y)$ and let $Y_i$, $i = 1, 2, \ldots, n$, denote $n$ independent, identically distributed observations of $Y$. The kernel density estimate of $f$ at a point $y$ is given by

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y - Y_i}{h}\right), \tag{17}$$

where $K$ is the kernel function, $n$ is the number of real observations and $h$ is the smoothing parameter or bandwidth. The density estimate at a fixed point ($y$) is the sum of $n$ kernels of mass, where the magnitude of each kernel is dictated by the distance between the fixed point and the observed data point (i.e. observations that are within the $h$-radius of the fixed point contribute more to the density estimate at that point while those outside the radius contribute little). The kernel function is usually (but not necessarily) a probability density function itself, being positive valued on the real line and integrating to unity. Its shape dictates the shape of each kernel while the smoothing parameter determines its width. A large smoothing parameter yields a smoother density estimate while a smaller value reveals fine details in the density estimate. One common kernel function for univariate data is the Gaussian kernel given by

$$K_G(u) = \frac{1}{\sqrt{2\pi}} \exp(-0.5u^2). \tag{18}$$

As is the case with most statistical estimators, there exists a trade-off between the variance and bias of a KDE. This trade-off is dictated by adjustment of the smoothing parameter $h$, and when the Gaussian kernel function is used, the optimal smoothing parameter may be computed directly by (Silverman 1986)

$$h_{\text{opt}} = 0.9 \cdot \min\{\sigma,\ \text{interquartile range}/1.34\} \cdot n^{-1/5}, \tag{19}$$

where $\sigma$ may be approximated with the sample standard deviation, $s$. Silverman (1986) stated that this value works well for a wide range of univariate distributions, including those with a bimodal structure. Furthermore, the author demonstrated through simulation that for $n = 100$, skewness or bimodality will be revealed using the smoothing parameter of (19). However, one drawback of the kernel estimator is its tendency to show spurious noise in long-tailed distributions. This is due to the fact that a fixed smoothing parameter is used over the range of the data. By employing

adaptive kernel methods which vary the smoothing parameter over the range of the data, this problem can be alleviated.

We apply the univariate KDE of (17) in the following manner. Let $P$ be a finite set of fixed points at which to estimate the density. The probability mass associated with the point $x_0 \in P$ is given by

$$\hat{p}(x_0) = \delta^{-1}\hat{f}(x_0), \tag{20}$$

where $\delta = \sum_{x \in P}\hat{f}(x)$. The cumulative distribution function approximator is given by

$$\hat{F}(x_0) = \sum_{\{x \in P : x \leq x_0\}} \hat{p}(x). \tag{21}$$

Assuming independence of the component characteristics, it is possible to obtain the distribution of $X$ by means of the convolution of the component characteristic density functions, or by the central limit theorem if there are a large number of components. However, since the component characteristics are not necessarily independent, the distribution must be obtained by means of the joint density. Next, the multivariate kernel estimator will be reviewed for approximating the joint density function of the component characteristics.

### 4.2. *Multivariate kernel density estimator*

The distribution of the assembly characteristic is not easily obtained, particularly when the assumption of independence is removed. In this section, the multivariate form of the kernel density estimator is reviewed (Silverman 1986, Scott 1992).

Let $X_1, X_2, \ldots, X_k$ be $k$ continuous random variables having joint density function $f(x_1, x_2, \ldots, x_k)$. Further suppose that the random variables are a function of $X$ such that

$$X = \psi(X_1, X_2, \ldots, X_k). \tag{22}$$

The density function of $X$ may be obtained from the joint density of the components by the moment-generating function (MGF) technique or the transformation of variables technique. The transformation of variables technique is employed since the MGF technique requires independence of the random variables. The multivariate form of the kernel estimator provides the means by which the joint density of component characteristics, $f(x_1, x_2, \ldots, x_k)$ may be approximated and is given by

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^k}\sum_{j=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}}{h}\right). \tag{23}$$

where $\mathbf{x}$ is a point in $k$-dimensional space. Thus, $\mathbf{x}_j = (x_{1j}, x_{2j}, \ldots, x_{kj})$, $j = 1, 2, \ldots, n$ constitutes a set of $n$ observations from the manufacturing process producing component $X_i$, $i = 1, 2, \ldots, k$. It is seen that the kernel estimate for multivariate data is analogous to the univariate case in that we consider the 'distance' from our point of consideration, $\mathbf{x}$ and the observations, $\mathbf{x}_j$, $j = 1, 2, \ldots, n$. Furthermore, the kernel function is positively valued and integrates to unity on $\mathbf{R}^k$.

If there exist extreme differences in variance between the component characteristic observations, (23) may not yield a good estimate for the overall joint density. Silverman (1986) suggested a method for prescaling the data so that the complication may be alleviated. More specifically, the author suggested a transformation of the form (Fukunaga 1972)

$$\hat{f}(\mathbf{x}) = \frac{|\mathbf{S}|^{-1/2}}{nh^k} \sum_{i=1}^{n} K\{h^{-2}(\mathbf{x} - \mathbf{x}_i)^{\mathrm{T}} \mathbf{S}^{-1} (\mathbf{x} - \mathbf{x}_i)\}, \tag{24}$$

where $\mathbf{S}$ represents the covariance matrix for the observed data, $|\mathbf{S}|$ is its determinant and $\mathbf{S}^{-1}$ is its inverse. Using a substitution of variable, let

$$u_i = h^{-2}(\mathbf{x} - \mathbf{x}_i)^{\mathrm{T}} S^{-1} (\mathbf{x} - \mathbf{x}_i), \quad i = 1, 2, \ldots, n. \tag{25}$$

Each kernel mass is obtained by evaluating $K$ at the point $u_i$. A common kernel function in the multivariate case is the multivariate Epanechnikov given by Silverman (1986):

$$K_e(\mathbf{x}) = \begin{cases} \dfrac{1}{2v_k}(k+2)(1 - \mathbf{x}^{\mathrm{T}}\mathbf{x}) & \mathbf{x}^{\mathrm{T}}\mathbf{x} < 1 \\ 0 & \text{elsewhere,} \end{cases} \tag{26}$$

where $k$ is the number of variables and $v_k$ is the volume of the $k$-dimensional unit sphere. For example, in case $k = 2$, the value of the kernel function is

$$K_e(u_i) = \begin{cases} \dfrac{2}{\pi}(1 - u_i) & u_i < 1 \\ 0 & \text{otherwise.} \end{cases} \qquad i = 1, 2, \ldots, n \tag{27}$$

Silverman (1986) and Scott (1992) have stated that selection of the kernel function is not a crucial issue in constructing the density estimate, thus, the Epanechnikov kernel is used throughout this work due to its computational simplicity.

Selection of the smoothing parameter for the multivariate kernel is analogous to that of the univariate case. The objective again is to optimize the trade-off between variance and bias in the estimator. The optimum smoothing parameter is given by

$$h_{\text{opt}} = A(K) n^{-((k+4)^{-1})} \tag{28}$$

where $A(K)$ is a constant that depends on the kernel function. For example, $A(K) = 2.40$ when using the Epanechnikov kernel for the case $k = 2$ (cf. Silverman 1986). The link between the joint density of the component characteristics and the density of the assembly characteristic, $X$, is made through the transformation of variables technique. Implementation of the technique requires the continuous characteristic distributions to be approximated by discrete distributions; therefore, the discrete form of the transformation should be applied.

## 5. Validation and numerical examples

Distributions of the component and assembly characteristics will be estimated on a finite support; thus, the transformation of variables technique for discrete variables is reviewed for the case $k = 2$. Let $X_1$ and $X_2$ be two discrete random variables with joint probability mass function (pmf), $\phi$. The transformation of variables technique transforms the joint pmf of $X_1$ and $X_2$ to the joint pmf of $X$ and $X_1$. The procedure for obtaining the pmf of the random variable $X = X_1 + X_2$ is as follows.

*Step* 1. Transform $\phi(x_1, x_2)$ to $\pi(x_1, x)$ by a substitution of variables $(x_2 = x - x_1)$ so that

$$\pi(x_1, x) \equiv \phi(x_1, x - x_1). \tag{29}$$

*Step* 2. Sum $\pi(x_1, x)$ over all $x_1$ to obtain $h(x)$, the pmf for $X$ by

$$h(x) = P\{X = x\} = \sum_{x_1} \pi(x_1, x). \tag{30}$$

Let $S_x \equiv \{\mathbf{x} : x = x_1 + x_2\}$ where $\mathbf{x} = (x_1, x_2) \in \mathbf{R}^2$. The probability that $X$ assumes the value $x$ in its sample space is given by

$$h(x) = P\{X = x\} = \sum_{\mathbf{x} \in S_x} \phi(\mathbf{x}). \tag{31}$$

Using this approach, the distribution of the sum of two random variables can easily be approximated, even if the component variables are not independent. The pmf approximation is obtained from the multivariate kernel estimate by converting the joint densities into point masses over a large (but finite) set of points as described by (20) in the univariate case.

### 5.1. *Validation of univariate estimates*

To demonstrate the viability of the univariate kernel estimator in approximating the component characteristic distributions and $g_i$, $i = 1, 2$, samples of 100 observations were taken from normal, symmetric triangular and uniform populations (parameters are given in table 1). These populations were generated using the inverse transform method. The Kolmogorov–Smirnov (K–S) goodness-of-fit test uses the maximum absolute deviation between the estimated and true cumulative distribution functions (CDFs) for a fixed number of points ($m$) as the test statistic. If the calculated test statistic exceeds the critical value, then the null hypothesis that the distributions are equal is rejected. In particular, for $m > 40$, the K-S critical value is (cf. Conover, 1980)

$$\theta = 1.36m^{-1/2}. \tag{32}$$

In the univariate experiments, the kernel and actual CDFs were compared at 151 points which, at the 0.05 level, yielded a maximum allowable deviation in probability of $\theta = 0.111$. For each distribution type, five random samples (of 100 observations each) were drawn and kernel estimates constructed for each case. Table 1 summarizes the cases and reports the calculated K–S test statistic. It is seen that, in each case, we fail to reject the hypothesis that the distributions are equal.

Figure 2 depicts the distribution function of the kernel estimate and the actual CDF of a symmetric triangular random variable with minimum value of 50, mode of 75 and maximum of 100. With a maximum absolute deviation of 0.0411 in probability, it is seen that the kernel estimate closely matches the true triangular distribution function.

| Sample no. | Normal (75, 69.44) | Triangular (50, 75, 1000) | Uniform (50, 100) |
|---|---|---|---|
| 1 | 0.0300 | 0.0601 | 0.0522 |
| 2 | 0.0638 | 0.0322 | 0.0479 |
| 3 | 0.0486 | 0.0634 | 0.0298 |
| 4 | 0.0181 | 0.0425 | 0.0801 |
| 5 | 0.0824 | 0.0497 | 0.0404 |

Table 1. Maximum absolute deviation in probability of actual versus kernel CDFs.

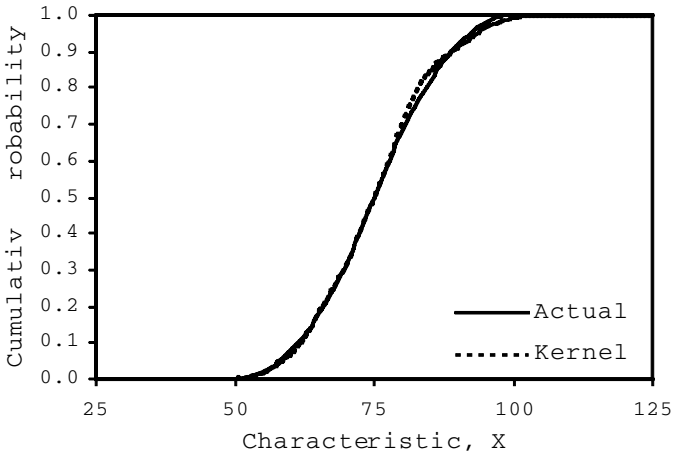


Figure 2. Sample distribution function $X \sim T(50, 75, 100)$.

It was further desired to test the ability of the univariate CDFs to predict the appropriate $g$ $(g_1, g_2)$ as per the technique presented in section 3. However, in each case, the true population standard deviation was replaced by the sample standard deviation, which is based on 100 observations. In this comparison, the approximated $g$ was compared with the known theoretical $g$ for the normal, triangular and uniform distributions assuming $C_p^* = 1$. It is known that for the normal distribution, $g = 6.0$ will cover 99.73% of the distribution with 0.135% on each tail. For the triangular distribution, $g = 4.899$ will yield 100% coverage of the distribution, while it is 3.464 for the uniform distribution with 100% coverage. Table 2 summarizes the results from the univariate experiment.

### 5.2. *Non-normal component characteristics*

To demonstrate the efficacy of the kernel estimator for determining the distribution of the assembly characteristic $X$ when the component characteristics are non-normal and independent, three distributions were assumed for the component characteristics and all combinations tested. A summary of the distribution parameters is given in table 3.

The cumulative distribution function of the assembly characteristic was obtained via Monte Carlo simulation (population size 12 000). The K–S one-sample test was again applied to assess statistical equality of the simulated and kernel distributions. Cumulative distribution functions were compared at 50 distinct points which, by

| | Normal (99.73%) | | Triangular (100%) | | Uniform (100%) | |
|---|---|---|---|---|---|---|
| Sample no. | Estimate | Actual | Estimate | Actual | Estimate | Actual |
| 1 | 5.777 | 6.000 | 5.172 | 4.899 | 3.815 | 3.464 |
| 2 | 5.933 | 6.000 | 4.804 | 4.899 | 3.779 | 3.464 |
| 3 | 6.231 | 6.000 | 5.047 | 4.899 | 4.986 | 3.464 |
| 4 | 6.071 | 6.000 | 4.713 | 4.899 | 3.870 | 3.464 |
| 5 | 5.644 | 6.000 | 4.738 | 4.899 | 3.749 | 3.464 |

Table 2. Estimated versus actual $g$ for component characteristics.

| Case no. | $X_1$ distribution | $X_2$ distribution |
|---|---|---|
| 1 | triangular (50, 75, 100) | uniform (50, 100) |
| 2 | triangular (50, 75, 100) | exponential ($75^{-1}$) |
| 3 | uniform (50, 100) | exponential ($75^{-1}$) |

Table 3. Component characteristic distribution parameters.

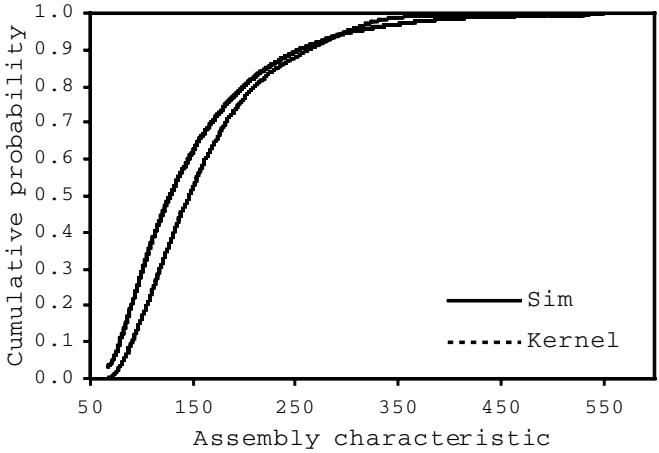


Figure 3. Sample distribution function estimate with 100 observations.

(32), gives a critical value of $\theta = 0.192$ for the K–S test. Figure 3 shows a typical example of the performance of the kernel technique in estimating the distribution function of the assembly characteristic when the first component is symmetric triangular and the second is exponentially distributed.

Table 4 summarizes the results of all goodness-of-fit tests for 10 trials under each case and it is observed that, in every instance, we fail to reject the null hypothesis that the estimated distribution is equal to the true distribution at the 0.05 level. Hence, we may be confident that the distribution functions used for the tolerance allocation technique will yield reliable estimates. The next section considers the case where the

| Sample no. | Triangular + uniform | Triangular + exponential | Uniform + exponential |
|---|---|---|---|
| 1 | 0.1043 | 0.1264 | 0.0818 |
| 2 | 0.0789 | 0.0796 | 0.1370 |
| 3 | 0.0692 | 0.0611 | 0.1074 |
| 4 | 0.0791 | 0.1068 | 0.1038 |
| 5 | 0.1203 | 0.0826 | 0.0886 |
| 6 | 0.1264 | 0.1364 | 0.0892 |
| 7 | 0.1387 | 0.1082 | 0.1099 |
| 8 | 0.0847 | 0.1017 | 0.1175 |
| 9 | 0.0775 | 0.0998 | 0.1137 |
| 10 | 0.0798 | 0.0906 | 0.1349 |

Table 4. Maximum absolute deviation in probability, simulated versus estimated CDF. The critical value is 0.192 at the 0.05 level.

two component characteristics are normally distributed and correlated random variables.

### 5.3. *Correlated, normal component characteristics*

In this experiment, we attempt to characterize the CDF of the assembly characteristic, $X$ when the two component characteristics are identically distributed normal random variables which are correlated. Recall that the sum of two correlated random variables is again normal with mean equal to the sum of the individual means and variance given by (6). We again compare our discretized kernel estimates with the true distribution of $X_1 + X_2$ by means of the K–S test. In each case, the kernel estimate of the joint density function was constructed using a random sample of 50 observations of $(X_1, X_2)$. The normal random variables have a mean of 75 and standard deviation of 10. The resulting CDFs were compared at 50 distinct points (not the observations) to test the hypothesis that the distributions are equal at the 0.05 level. Table 5 gives a summary of the empirical results.

Figure 4 depicts the estimated distribution function ('Kernel') of the assembly characteristic $X$ versus the actual population CDF ('Actual'). It is important to note that each of these was constructed using only 50 sampled components from the populations of $X_1$ and $X_2$ This is significant because a relatively small sample size is used to characterize the distribution, even when the components are correlated.

## 6. Conclusions and future directions

This paper has presented an approximation technique for statistical tolerancing analysis that allocates component tolerances for a given desired assembly tolerance. The main advantage of the simple approach is that it is a data-driven, non-parametric approach that allows the process data to dictate the appropriateness of component characteristic distributions and the assembly characteristic distribution. Thus, the typical assumptions of independence and normality can possibly be relaxed when the number of components being combined is small (i.e. $k < 10$).

To validate the technique's ability to characterize the assembly characteristic distribution, we examined two cases for two components: (1) the two component characteristics are independent and *not* normally distributed, and (2) the component characteristics are normally distributed but highly correlated. In the first case, the results indicate that kennel density estimates *can* characterize the component distributions and the overall assembly characteristic distribution with reasonable precision. The maximum absolute deviations (in probability) are not statistically significant at the 0.05 level for three different distribution types. In the second case, it was found that the kernel estimate performs well when estimating the assembly characteristic distribution even when the component characteristics are highly correlated normal random variables. A logical extension is to consider the case of a small number of highly correlated, non-normal component characteristics. This work is currently underway.

Both experiments indicate that by using actual process data, it may be possible to relax the typical assumptions made in statistical tolerancing analysis, namely the independence and normality assumptions for the component characteristics. Of course, as with any statistical estimator, there are pros and cons for using the kernel density estimates. One disadvantage is the process of selecting the smoothing parameter, $h$. Although this process may be automated to some degree, it is instructive to 'fine tune' the parameter to achieve a better fit of the data, as a different

| Test no. | Correlation | Maximum deviation in probability | Critical value |
|---|---|---|---|
| 1 | 0.5000 | 0.0731 | 0.1920 |
| 2 | 0.7070 | 0.0713 | 0.1920 |
| 3 | 0.9480 | 0.0790 | 0.1920 |
| 4 | 0.9950 | 0.0837 | 0.1920 |

Table 5. Test results for $X = X_1 + X_2$, where $X_1$ and $X_2$ normal and correlated.

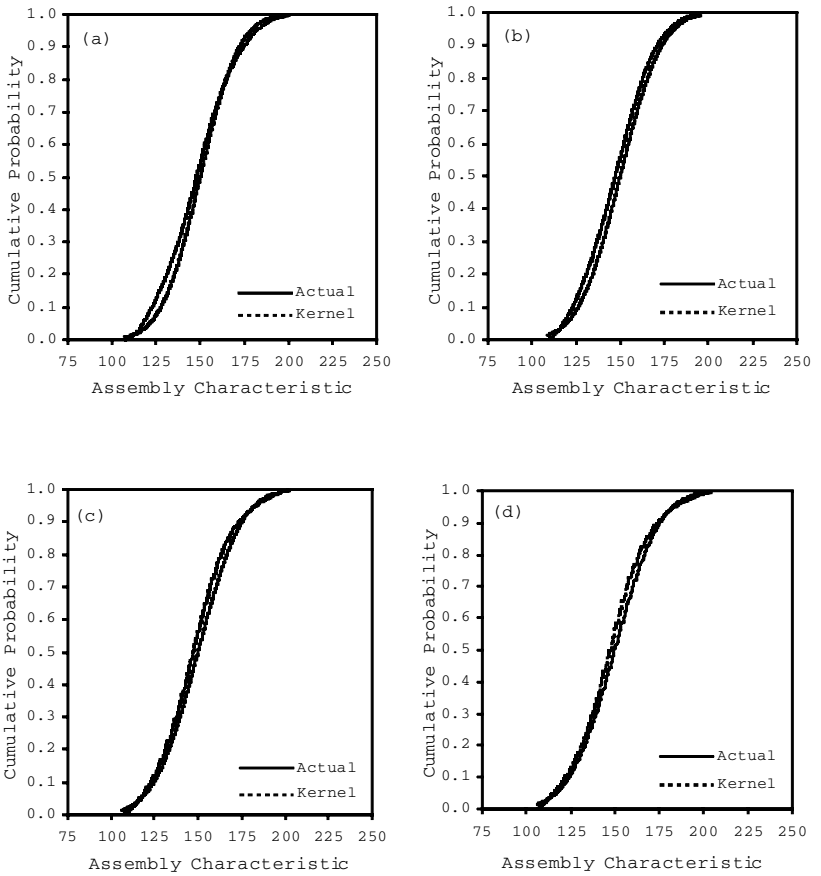


Figure 4. Assembly characteristic distribution when components have a correlation coefficient: (a) 0.500, (b) 0.707, (c) 0.950, (d) 0.995.

smoothing parameter yields possibly dramatically different distributions. One obvious advantage of using the kernel estimator is that we make no distributional assumptions with regard to the component or assembly characteristics, and do not lose much in the discretization of the distributions so long as the smoothing parameter is chosen appropriately. Hence, with some good historical process data, it may be possible to determine quickly the appropriate component tolerances to achieve the overall desired assembly tolerance. Even with a small number of samples, it is possible to find reasonable values for the distribution functions.

As indicated here, the concept of process capability ratios (capability indices) is very closely related to the statistical allocation of component tolerances. In particular, it was seen that the target $C_p$ index was used explicitly to find equations needed to allocate the known assembly tolerance among $k$ components. This procedure can be easily extended to incorporate other target capability indices such as $C_{pk}$ or $C_{pm}$ by using the maximum deviation allowable ($s_i$) between the process mean and nominal dimension $d_i$. These extensions are applicable even for non-normal or correlated component characteristics since the approach is driven by the actual process data to compute the needed values, $c$ and $c_i$, as indicated in section 3.2.

In the future, it would be instructive to investigate more sophisticated kernel estimators that are adaptive. In this paper, the smoothing parameter was fixed in both coordinate directions, but it could be varied by choosing a vector of smoothing parameters in each direction. Such an adaptive method may lead to better estimates of the joint density, and, hence, of the assembly characteristic distribution. The kernel does lend itself to extension to higher dimensions, and we are currently exploring this to include more than two characteristics. Ultimately, the technique will be incorporated into a cost-minimization scheme and compared against existing techniques.

**Acknowledgements**

The authors acknowledge the helpful comments of two anonymous referees whose suggestions helped to clarify several concepts in the paper.

**References**

Barker, A. M., 1989, Statistical tolerancing: methods to improve quality. Proceedings of the 43rd Annual Quality Congress Transactions, 8–10 May, Vol. 43, Toronto, pp. 490–495.

Bender, A., Jr, 1968, Statistical tolerancing as it relates to quality control and the designer. *SAE Transactions*, **77**, 1965–1971.

Bjorke, O., 1989, *Computer-aided Tolerancing*, 2nd edn (New York: ASME Press).

Burr, I. W., 1958, Some theoretical and practical aspects of tolerances for mating parts. *Industrial Quality Control*, **15**, 18–22.

Conover, W. J., 1980, *Practical Nonparametric Statistics* (New York: Wiley).

D'Errico, J. R. and Ziano, N. A., 1988, Statistical tolerancing using a modification of Taguchi's method. *Technometrics*, **30**, 397–405.

Evans, D. H., 1974, Statistical tolerancing: the state of the art, Part I. *Journal of Quality Technology*, **6**, 188–195.

Freund, J. and Walpole, R., 1987, *Mathematical Statistics*, 4th edn (Englewood Cliffs: Prentice-Hall).

Fukunaga, K., 1972, *Introduction to Statistical Pattern Recognition* (New York: Academic Press).

Mansoor, E. M., 1963, The application of probability to tolerances used in engineering design. *Proceedings of the Institution of Mechanical Engineers*, **178**, 29–44.

Nigam, S. D. and Turner, J. U., 1995, Review of statistical approaches to tolerance analysis. *Computer-Aided Design*, **27**, 6–15.

O'Connor, M. A. and Srinivasan, V., 1997, Composing distribution function zones for statistical tolerance analysis. Proceedings of the 5th CIRP International Seminar on Computer-Aided Tolerancing (*CAT*), pp. 13–24.

Parkinson, D. B., 1982, The application of reliability methods to tolerancing. *Transactions of the ASME Journal of Mechanical Design*, **104**, 612–618.

Parkinson, D. B., 1984, Tolerancing of component dimension in CAD. *Computer-Aided Design*, **16**, 25–31.

Scott, D. W., 1992, *Multivariate Density Estimation, Theory, Practice, and Visualization* (New York: Wiley).

SILVERMAN, B. W., 1986, *Density Estimation for Statistics and Data Analysis* (New York: Chapman & Hall).

ZHANG, H. C. and HUQ, M. E., 1992, Tolerancing techniques: the state-of-the-art. *International Journal of Production Research*, **30**, 2111–2135.

ZHANG, C., LUO, J. and WANG, B., 1999, Statistical tolerance synthesis using distribution function zones. *International Journal of Production Research*, **37**, 3995–4006.