Research Journal of Recent Sciences ISSN 2277-2502
Vol. **2(11),** 55-64, November **(2013)** *Res.J.Recent Sci.*

A Statistical Method for Designing and analyzing tolerances of Unidentified Distributions

Movahedi M.M.
Department of Management, Firoozkooh Branch, Islamic Azad University, Firoozkooh, IRAN

Available online at: www.isca.in, www.isca.me
Received 30th April 2013, revised 15th May 2013, accepted 14th July 2013

Abstract

The mechanical tolerances are set to restrict too large dimensional and geometrical variation in a product. Tolerances have to be set in such a manner that functionality, manufacturability, costs and interchangeability are optimized and balanced between each other. The tolerances and available tolerance design techniques are represented in this text. Statistical tolerance design is emphasized because statistical behavior describes the nature of the manufacturing processes more realistically than worst-case methods. To this end, the Generalized Lambda Distribution (GLD) has been used for design of tolerance. This distribution is highly flexible and based on the available data, can identify and present the related probability distribution function and their statistics. After recognizing the underlying probability distribution function, the results can be employed for the design of tolerance.

Keywords: Tolerance, generalized lambda distribution, quality control.

Introduction

Ever-increasing competition in global markets forces companies to manufacture higher-quality products at a faster rate and with lower costs than their competitors. The customer needs and wants are on the rise as well. Thus, product development teams have to rapidly design specifications for complex assemblies. In mass production, mechanical variation is a significant contributor to poor quality, increased costs, and wasted time. Mechanical variation in a product's characteristics is caused by dimensional and geometrical variation in the components and by assembly variation. Further, component variation is brought about by manufacturing variation. Apart from influencing a product's characteristics, component variation can considerably complicate the assembly of the components. In today's competitive and global business environment, every component of a product must be individually replaceable. That is, a great number of parts can be made independent of mating parts, and any one part can be expected to mate with any other and still function properly. Complicated assembly produces scrap, consumes time, and deteriorates the ability to deliver, which in turn involves extra costs and decreases revenue[1].

The goal of tolerance design is to produce designs that could be assembled and function correctly despite variation. In parallel with tolerance design, a careful optimization of the nominal dimensions has to be emphasized in order to make the design as insensitive as possible to variation. By means of tolerance design techniques, the features and their tolerances that mostly affect the assembly requirements can be identified as tighter as possible, which improves performance and quality. On the other hand, the tolerances of the non-critical features can be loosened, which reduces costs and saves time. In addition, inexpensive tolerances can be tightened and expensive ones can be loosened. To design realistic tolerances, an active collaboration between design and manufacturing must take place sufficiently early in the design phase of a product. The acceptable and achievable tolerances have to be discussed then.

We will consider assemblies of k components ($k \geq 2$). The quality of the characteristic of component *i* that is of interest to the designer and is denoted by X_i. This characteristic is assumed to be of the Nominal-the-Better type. The upper and lower specification limits of X_i are $U_i(USL_i)$ and $L_i(LSL_i)$,respectively. The assembly quality characteristic of interest to the designer depended by X is function of $X_i, i = 1, 2, \ldots, k$. That is,

$$X = f(X_1, X_2, \ldots, X_k) \quad (1)$$

At first, we will consider linear functions of X_i only:

$$X = X_1 \pm X_2 \pm \ldots \pm X_k \quad (2)$$

The upper and lower specifications are assumed to be given by the customer or determined by the designer based on the functional requirements specified by the customer.

Tolerance is the difference between the upper and lower specification limits. Let the tolerance of X_i be $T_i, i = 1, 2, \ldots, k,$ and let the tolerance of the assembly characteristic X be T. then,

$$T_i = U_i - L_i, i = 1,2,...,k \quad (3)$$

where L_i and U_i are the lower and upper specification limits of characteristic X_i, respectively. In general, for any linear function $X = X_1 \pm X_2 \pm ... \pm X_k$, we have

$$T_a = T_{a1} + T_{a2} + ... + T_{ak} \quad (4)$$

This is called an *additive relationship*. The design engineer can allocate tolerances $T_1, T_2, ..., T_k$ among the *k* components, for a given specified *T*, using this additive relationship.

Probabilistic Relationship: Tolerance can be defined as being concerned either with physical and chemical properties, including size, weight, hardness, and composition of a part, or with the geometric characteristics, including dimension, shape, position, and surface finish of some part features. As it is impossible to produce many parts each to have exactly the same nominal value of a feature, deviations from the design nominal are unavoidable and hence allowed or tolerated. When a part deviates too much from the nominal, it fails to perform the intended function. To ward off possible functional failures, design engineers usually determine a maximum allowable deviation known as the tolerance, with upper and/or lower limits specified for each quality feature.

As this relationship relies on the probabilistic properties of component and assembly feature, it is essential to making certain *assumptions* regarding these characteristics: i. X_i Are independent of each other. ii. Components are randomly assembled. iii. $X \approx N(\mu_i, \sigma_i^2)$; That is, the characteristic X_i is normally distributed with a mean μ_i and a variance σ_i^2 (this assumption will be relaxed later on). iv. The process that generate characteristic X_i is adjusted and controlled so that the mean of the distribution X_i, μ_i, is equal to the normal size of X_i, denoted by B_i, which is the point of the tolerance region of X_i. That is

$$\mu_i = \frac{(U_i - L_i)}{2} \quad (5)$$

The standard deviation of the distribution of the characteristic X_i, generated by the process, is such that 99.73% of the characteristic X_i falls within the specification limits for X_i. Based upon the property of normal distribution, this is represented as

$$U_i - L_i = T_i = 6\sigma_i, i = 1,2,...,k \quad (6)$$

Let X_i and X_i be the mean and variance of X respectively. As $X = X_1 \pm X_2 \pm ... \pm X_k$,

$$\mu = \mu_1 \pm \mu_2 \pm ... \pm \mu_k \quad (7)$$

the X_i 's are independent of each other,

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_k^2 \quad (8)$$

and considering assumption 2 (above), the assembly characteristic *X* is also normally distributed.

Let us assume that the 99.73% of all assemblies have characteristic *X* within the specification limits *U* and *L*. This yields an equation similar to equation (6). From equation (6) and (8), it can be derived that:

$$\sigma_i^2 = \left(\frac{T_i}{6}\right)^2, i = 1,2,...,k \quad (9)$$

and

$$\left(\frac{T_i}{6}\right)^2 = \left(\frac{T_1}{6}\right)^2 + \left(\frac{T_2}{6}\right)^2 + ... + \left(\frac{T_k}{6}\right)^2 \quad (10)$$

or

$$T_p = \sqrt{T_{p1}^2 + T_{p2}^2 + ... + T_{pk}^2} \quad (11)$$

The relation given in equation (11) is called a *probabilistic relationship* and provides a different means for allocating tolerance among components for a given assembly tolerance, *T*.

For example, let us consider the assembly as having two components with the characteristics X_1 and X_2 respectively. If we assemble this two components, then assembly characteristic can be denoted by *X*, which is equal to: $X=X_1+X_2$ and $T_a=T_{a1}+T_{a2}$. Let's presume now that the tolerance on X, which is T_a, is 0.001 inch, so $T_{a1}+T_{a2}=0.001$.

There are two unknowns, T_{a1}, and T_{a2}, and only one equation. Let us assume that, in one example, the difficulty levels of maintaining both T_{a1} and T_{a2} are the same, hence the designer would like these tolerances to be equal. That is, $T_{a1}=T_{a2}=0.0005$.

Now setting T=0.001 in equation (11) yields: $\sqrt{T_{p1}^2 + T_{p2}^2} = 0.001$. If we introduce the same first relation used earlier $T_{p1}=T_{p2}$, then we have $\sqrt{2T_{p1}^2} = 0.001 \Rightarrow T_{p1} = T_{p2} = \frac{0.001}{\sqrt{2}} = 0.00071$. In this example, we set $T_a=T_p=0.001$ and solved for $T_{a1}=T_{a2}=0.0005$ and $T_{p1}=T_{p2}=0.00071$. We saw that $T_{p1}>T_{a1}$ and $T_{p2}>T_{a2}$.

Now in general we could have two relations between T and $(T_1, T_2, ..., T_k)$ as below:

$$T_a = T_{a1} + T_{a2} + ... + T_{ak} \quad (12)$$

$$T_p = \sqrt{T_{p1}^2 + T_{p2}^2 + ... + T_{pk}^2} \quad (13)$$

Now let us go through the advantages and disadvantages of using the probabilistic relationship to allocate tolerances among the components[2].

Advantage of using a probabilistic relationship: It is a well-established fact that manufacturing cost drops as the tolerance on the quality characteristic increases. Hence, the manufacturing cost of the components will decrease as a result of using the probabilistic relationship.

Disadvantage of using a probabilistic relationship: If the probabilistic relationship is used, the actual maximum range of the clearance of the assemblies using these components will be:

$$T_1 + T_2 = 0.00071 + 0.00071 = 0.000142$$

The allowable range of the clearance of the assemblies, T, is 0.001. This will obviously lead to rejection of the assemblies. In order to estimate the actual proportion of rejection, we need the probability distribution of the assembly characteristic, X, along with its mean and standard deviation.

If the component characteristics are normally distributed, then the assembly characteristics is also normally distributed. Then by using equation 8 and 9, we can calculate the standard deviation and illustrate that the percentage rejection of the assemblies is less than 0.27%. So the percentage rejection of probabilistic relationship is greater than additive relationship.

Probabilistic Relationship for non-normal component characteristics

Two approaches can be basically considered in statistical tolerance design while dealing with conditions that process output holds an abnormal distribution. First, this issue is not that sensitive to cause trouble and consequently, tolerance design is being carried on as before. Second, this is not the case and an alternative should be taken into account. In many conditions, the above-mentioned subject does not have main concerns and is only taken into consideration to improve the quality control. But the distribution of output is the most leading indication in all organizations[2].

Yourstone and Zimmer[3], studied the Skewed and Rocky pattern and found out that when the process output has a normal distribution but the Skewness is not fit, the efficiency of traditional control charts should be considered. Kittlitz[4], used exponential distribution instead of normal distribution for some too skewed abnormal processes, with the fifth root in place of main data.

Peam, Kotz and Johnson[5], proposed a method highly applicable in an immense extent of distributions. This method does not need to know the skewness or rockiness of the distribution, but this method assumes that the output distribution is Gamma, that in many conditions are not true.

In literature, several imputation techniques are described. Thakur and et al[6] present the estimation of mean in presence of missing data under two-phase sampling scheme while the numbers of available observations are considered as random variable. Rekha R. C. and Vikas S[7], have formulated an Inventory model for deteriorating items with Weibull distribution deterioration rate with two parameters. Roman, and et al[8], have used Goodness-of-Fittest such as Anderson-Darling, Chi-square and Kolmogorov-Smirnov to judge the applicability of the distributions for modeling recorded Annual 1-Day Maximum Rainfall (ADMR) data.

It is not appropriate to adopt traditional methods in abnormal distribution cases. Even when the normal test has been done from a distribution point of view and the result is confirmatory, the problem below still exists.

In the above-mentioned test, when H_0 (which is the assumption of being a normal distribution) is rejected, it implies that the above distribution is not normal, while if H_0 is not rejected, it does not necessarily mean that H_0 is correct. Furthermore, due to the likelihood test, plenty of data is required for a certain judgment about H_0. Because of cost limitation or the lack of data as much required, the already mentioned tests are performed with less amount of data[9].

Gunter[10] found some new cases in their research that in spite of having the same mean and standard deviation as well as close distributions, the nature of distributions differed from each other.

Let the probability density function of X_i be $f(x_i)$ with a mean μ_i and a variance σ_i^2. We assume that the range that contains 100% or close to 100% of all possible values of X_i is $g_i\sigma_i$. It is still assumed that:

$$T_i = g_i\sigma_i \quad (14)$$

(ideally $T_i >>>> g_i\sigma_i$). This can be written as:

$$\sigma_i = \frac{T_i}{g_i} \quad (15)$$

Now, given that $X = X_1 \pm X_2 \pm ... \pm X_k$, the distribution of X is approximately normal, because of the Central Limit Theorem. So,

$$T_p = 6\sigma \Rightarrow \sigma = \frac{T_p}{6} \tag{16}$$

assuming 99.73% coverage. Using the formula $\sigma^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_k^2$,

$$\left(\frac{T_p}{6}\right)^2 = \left(\frac{T_1}{g_1}\right)^2 + \left(\frac{T_2}{g_2}\right)^2 + ... + \left(\frac{T_k}{g_k}\right)^2$$

$$T_p = 6 \times \sqrt{\left(\frac{T_1}{g_1}\right)^2 + \left(\frac{T_2}{g_2}\right)^2 + ... + \left(\frac{T_k}{g_k}\right)^2} \tag{17}$$

The Generalized Lambda Distribution

This distribution was first advanced by Tukey[11], and later on was developed by Junior and Rosenblatt[12]. This distribution can precisely fit the ordinary distribution like normal, lognormal, Weibull, etc. The flexibility of this distribution exerts influences on estimating continual distributions and matching on histogram data and estimating the distribution type. As a matter of fact, it serves as a powerful device for research in different areas like estimating parameters, adjusting distributions on data and simulating research based on data production. For example, it is deployed in operational research, Ganeshan[13], psychology meteorology, Ozturk, and Dale[14], Delaney, and Vargha[15], process statistic control, Fournier, and et al[9], safety and fault tolerance Gawand, and et al[16], and queue systems, Dengiz[17]. Zaven and et al[18] studied generalized lambda family of distributions, generalized bootstrap and Monte Carlo, and fitted these distributions with the data.

May researcher have been done their studies on tolerance design and/or GID[19-24]. Bigerelle, and et al[25], for example, use generalized lambda distribution and Bootstrap analysis to the prediction of fatigue lifetime and confidence intervals. In this research, the lambda distributions associated with the Bootstrap technique were first employed to model the Paris coefficients PDF and turned out to be able to estimate accurately the experimental values. Then, lambda distributions were used to model the PDF lifetime of a basic structure under fatigue loading. Acar, and et al[26] applied Estimation using Dimension Reduction and Extended Generalized Lambda Distribution to estimate reliability. They presented an analytical approach for systems reliability. Given an N-dimensional, differentiable, uni-modal performance function along with the statistical properties of the underlying random variables, the proposed approach applies the univariate dimension-reduction technique to the estimation of the five primary statistical moments, which are in turn used for figuring out the unknown parameters in the extended generalized lambda distribution for probability distribution fitting of the performance function.

The characterizing of generalized lambda distribution has been studied by Karvanen, and Nuutinen[27]. It has been introduced as a reversed Probability Cumulative Distribution Function Q as below:

$$x = Q(y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2} \tag{18}$$

where y implies the relative Density Probability in point x and it is obvious that its extent would fall between zero and one. λ_1 and λ_2 are the co efficiencies related to the measurement and place, respectively and λ_3, λ_4 referred to the prominence and suspense of the distribution. Some of the capabilities of this distribution for different distributions have been displayed in figures 1-4[28].

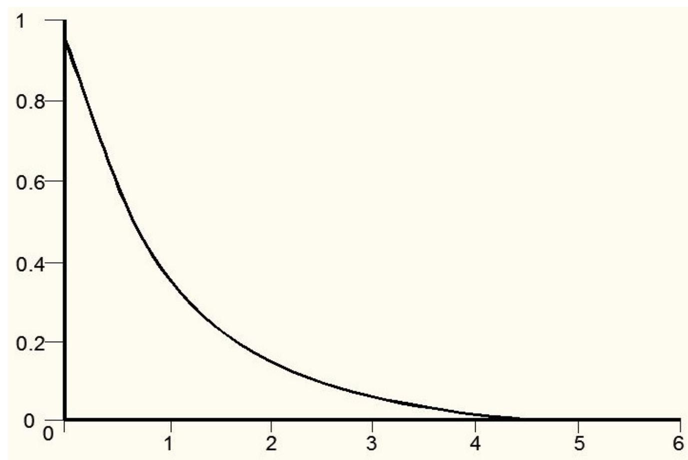

Figure-1
GLD (0.0069, -0.0011, -0.0000, -0.0011)
Negative exponential distribution with parameter 1

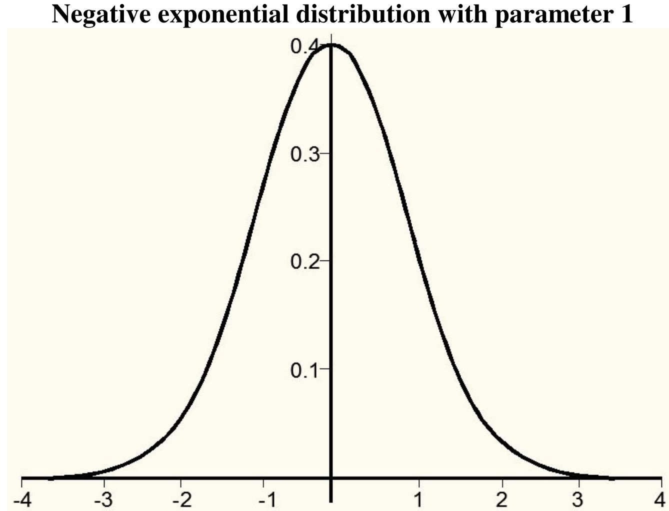

Figure-2
GLD (0, 0.1975, 0.1349, 0.1349)
Standard normal distribution

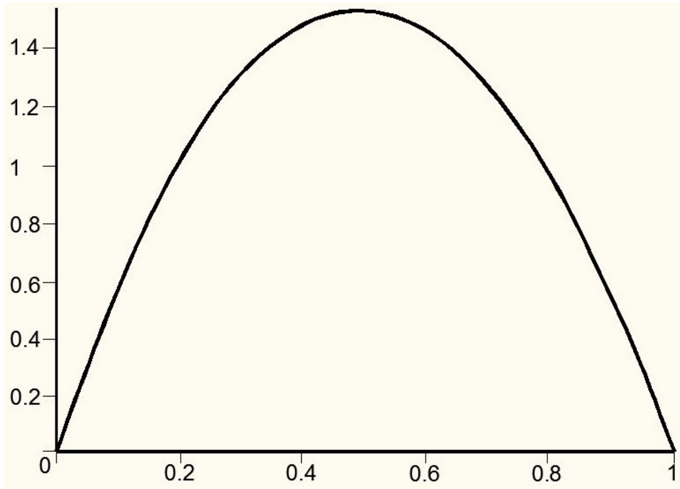

Figure-3
GLD (0, 5, 1.9693, 0.4495, 0.4495)
Beta distribution with parameters (1, 1)

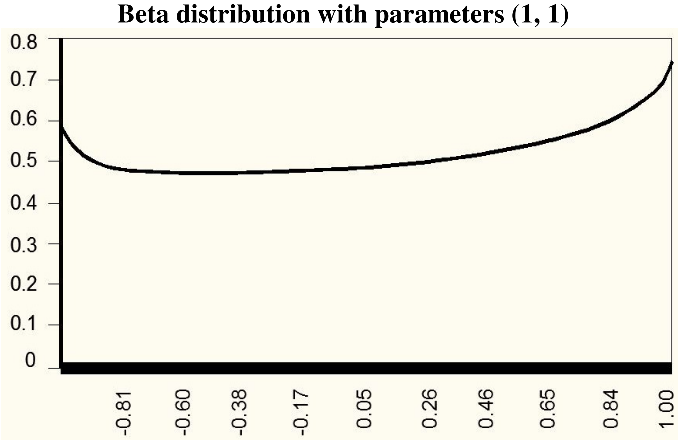

Figure-4
GLD (0, 1, 1.4, 1.6)
d: A distribution similar to a special U distribution

The generalized lambda distribution is a distribution that can organize, simulate and estimate all distributions through changing the parameters. It is flexible enough to exactly simulate and accordingly the quality control operation can be carefully done. This task is done by access to 100 data.

Fitting a probability distribution to data is an important task in any statistical data analysis. The data to be modeled may consist of observed events, such as quality characteristic of components. When fitting data, one typically first selects a general class, or family, of distributions and then finds values for the distributional parameters that best match the observed data.

As it was seen, by changing the amounts of λ, the GLD has been in distinct forms and matched on the distributions. The details of the specifications of this distribution, the applications and the way of computing the parameters have been precisely explained by Karian and Dudewicz[28].

Eventually, Tarsitano[29], raised the number of the parameters of this distribution up to 5 and studied the characteristics. This is done to increase the capability and exactness to fit the panel data distribution. The distribution of a new five parameters is an accumulated opposite as below:

$$X(y,\lambda) = \lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5}, 0 \le y \le 1 \qquad (19)$$

In this distribution, λ_1 is a place indicator, λ_2, λ_3 are measuring indicators and λ_4 and λ_5 are concerned with the distribution feature. The two measuring indicators display different weights for the distribution extent and provide this place for a new distribution to be well- adjusted with the data without symmetric extents are well – matched.
The PDF of this distribution is:

$$f(y) = \frac{dx}{dy} = \lambda_2 \lambda_4 y^{\lambda_4 - 1} + \lambda_3 \lambda_5 (1-y)^{\lambda_5 - 1} \qquad (20)$$

The m^{th} moment of this distribution can be computed by using the relations (19) and (20) and also the total computations of the moments are as follow:

$$\begin{aligned} M_m &= E(X^m) = \int_{-\infty}^{+\infty} X^m f(x) dx \\ &= \int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5})^m dy \end{aligned} \qquad (21)$$

Estimation the GLD parameters

Generalized lambda distribution (GLD) is a distribution that can be used for testing and fitting the data to well-known distributions. Since the GLD is defined by its quintile function, it can provide a simple and effective algorithm for generating random variations.

Fitting a probability distribution to data is an important task in any statistical data analysis. Several methods for estimating the parameters of the GLD, such as: Percentile Matching (PM), the moment matching (MM), Probability-Weighted Moment (PWM), Minimum Cramér-Von Mises (MCM), Maximum Likelihood (ML), Pseudo Least Squares (PLS), Downhill simplex method, and starship methods have been presented in the literature (Tarsitano[29]). Fournier and et al[9], for example, developed a new method for estimating the parameters of a GLD based on the minimization of the Kolmogorov–Smirnov distance in a two-dimension space.

In this research, the moment matching method is being briefly reviewed. The moment-matching method, described in this paper, was proposed in Ramberg and Schmeiser[30]. The method can be described in a straightforward manner as follows: given the GLD distribution with quartile function Q(*u*), find parameters λ_1, λ_2, λ_3, λ_4 and λ_5 so that the mean μ and variance σ^2 of the GLD match the corresponding mean μ^*, and variance $(\sigma^*)^2$ of the sample (i.e., the first five moments of the theoretical GLD match those of the data). More formally, if such a method denotes the probability density function of the random variable X with distribution 4, we compute the parameters λ such that satisfying equations 21. Finally, after determining the 5 parameters of the distribution, it can be demonstrated that the mean and variance of GLD can be calculated as below:

$$\mu = E(X) = \int_{-\infty}^{+\infty} x f(x)dx$$
$$= \int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5})dy \quad (22)$$
$$= \lambda_1 + \frac{\lambda_2}{(\lambda_4+1)} - \frac{\lambda_3}{(\lambda_5+1)}$$

$$\sigma^2 \equiv E(X^2) - [E(X)]^2$$
$$= \int_{-\infty}^{+\infty} x^2 f(x)dx - \left[\int_{-\infty}^{+\infty} x f(x)dx\right]^2 \quad (23)$$
$$= \int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4(1-y)^{\lambda_5})^2 dy - \left(\int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5})dy\right)^2$$
$$= \int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4(1-y)^{\lambda_5})^2 dy - (\lambda_1 + \frac{\lambda_2}{(\lambda_4+1)} - \frac{\lambda_3}{(\lambda_5+1)})^2$$

By using this method, when we can collect data from production line, and then calculate the mean, variance, and standard deviation of the underlying distribution of the qualitative specification, and finally we can design the tolerance.

Methodology and Results

In this section, we consider the assembly s having two components with the qualitative specification of X_1 and X_2, respectively. Since the real data is not available, and data gathering requires time and cost, in this research we produce and use 99 stochastic numbers for each part. We assume that each number is the qualitative specification of parts 1 and 2, respectively. Then, we classified the derived numbers in the frequency tables, and calculate 1 to 5th experimental moment of GLD by using the following equation (tables 1 and 2).

$$M_m = \sum_{i=1}^{k} x_i f_i \quad (24)$$

Where, M_m = m^{th} moment, x_i = midpoint or mean of i^{th} cell interval, i=1, 2, …, k, f_i = frequency of i^{th} cell interval, i=1, 2, …, k, k = number of cell interval,

Table-1
Frequency table and experimental moment for data of pare 1

Cell interval	f_i	x_i	$f_i x_i$	$f_i x_i^2$	$f_i x_i^3$	$f_i x_i^4$	$f_i x_i^5$
11.950 -- 11.965	16	11.958	191.32	2287.7089	27355.27917	327100.7507	3911307.226
11.965 --11.979	15	11.972	179.58	2149.93176	25738.98303	308147.1048	3689137.139
11.979 -- 11.993	13	11.986	155.818	1867.634548	22385.46769	268312.2158	3215990.218
11.993 -- 12.007	13	12.000	156	1872	22464	269568	3234816
12.007 -- 12.021	20	12.014	240.28	2886.72392	34681.10117	416658.7495	5005738.217
12.021 -- 12.035	13	12.028	156.364	1880.746192	22621.6152	272092.7876	3272732.049
12.035 -- 12.050	9	12.043	108.3825	1305.196256	15717.82592	189281.9186	2279427.505
Sum	99		1187.7445	14249.94158	170964.2722	2051161.527	24609148.35

Table-2
Frequency table and experimental moment for data of pare 2

Category	f_i	x_i	$f_i x_i$	$f_i x_i^2$	$f_i x_i^3$	$f_i x_i^4$	$f_i x_i^5$
8.012 -- 7.964	13	7.988	103.844	829.505872	6626.092906	52929.23013	422798.6903
7.964 -- 7.978	17	7.971	135.507	1080.126297	8609.686713	68627.81279	547032.2958
7.978 -- 7.992	15	7.985	119.775	956.403375	7636.880949	60980.49438	486929.2476
7.992 -- 8.006	12	7.999	95.988	767.808012	6141.696288	49127.42861	392970.3014
8.006 -- 8.02	18	8.013	144.234	1155.747042	9261.001048	74208.40139	594631.9204
8.02 -- 8.034	11	8.027	88.297	708.760019	5689.216673	45667.34223	366571.7561
8.034 -- 8.048	13	8.041	104.533	840.549853	6758.861368	54348.00426	437012.3023
Sum	99		792.178	6338.90047	50723.43594	405888.7138	3247946.514

Now we can use the equation (21) and establish 5 equations and 5 variables λ_1 to λ_5. These equations are then equal to the respective sum of relative column in tables 1 and 2. The equations are classified as equations 25 and 26 as bellow:

$$\begin{aligned}
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})dy=1187.7445\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^2 dy=14249.94158\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^3 dy=170964.2722\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^4 dy=2051161.527\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^5 dy=24609148.35
\end{aligned} \tag{25}$$

$$\begin{aligned}
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})dy=792.178\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^2 dy=6338.90047\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^3 dy=50723.43594\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^4 dy=405888.7138\\
&\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^5 dy=3247946.514
\end{aligned} \tag{26}$$

Now the Excell and MATLAB software programs can be deployed to calculate GLD parameters. The results of equations (25) are as follow:

$$\lambda_1=10.7436, \lambda_2=4.0029, \lambda_3=2.1856, \lambda_4=2.0851, \lambda_5=52.1962 \tag{27}$$

The results of equations (26) are as below:

$$\lambda_1=1.2427, \lambda_2=6.7886, \lambda_3=0.2331, \lambda_4=-4.3945, \lambda_5=2.5099 \tag{28}$$

Consequently, the quartile cumulative distribution function of the specification of components 1 and 2 in the study equals:

For part No. 1

$$X(\lambda,y)=10.7436+4.0029y^{2.1856}-2.0851(1-y)^{52.1962},$$

For part No. 2

$$X(\lambda,y)=1.2427+6.7886y^{0.2331}+4.3945(1-y)^{2.5099}$$

Goodness of fitness test: Solving non-linear equations usually yields more than one series of answers, so it is quite necessary to run goodness of fitness test to attain the acceptable answer. To this end, Chi-square statistics was deployed:

$$\chi_0^2=\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i} \tag{29}$$

H_0: GLD with the obtained parameters fits the data. H_1: GLD with the obtained parameters does not fit the data. H_0 is rejected if $\chi_0^2>\chi^2_{\alpha,k-i-1}$ where α is the level of significance of the test, K, the number of sets and I, the number of distribution parameters.

To carry out goodness of fit test for each part of components, first the expected values (E_i) are obtained. For each set i^{th}, the cumulative amount of relative frequency is positioned in the distribution relationship; hence, the value of the mean of the set in question is calculated, which is the very E_i. Similarly, the observed value is the mean of the data set. Chi-square statistic can test with K-6 degrees of freedom, where K is the number of class intervals. Table 5 presents the hypothesis testing of the research data.

Table-5
The goodness of fit test for part 1

Class Interval	Observation value O_i	Cumulative Relative Frequency Y_i	Expected Observation Value E_i	$(O_i-E_i)^2$	$\frac{(O_i-E_i)^2}{E_i}$
1	11.958	0.1616	10.818	1.140	0.1201
2	11.972	0.1515	10.808	1.150	0.1224
3	11.986	0.1313	10.790	1.196	0.1326
4	12.000	0.1313	10.790	1.210	0.1357
5	12.014	0.2020	10.865	1.149	0.1215
6	12.028	0.1313	10.790	1.238	0.1420
7	12.043	0.0909	10.750	1.293	0.1552
Total					0.9295

Table-5

The goodness of fit test for part 2

Class Interval	Observation value O_i	Cumulative Relative Frequency Y_i	Expected Observation Value E_i	$(O_i\text{-}E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	7.988	0.1313	8.558	0.570	0.0380
2	7.971	0.1717	8.484	0.513	0.0310
3	7.985	0.1515	8.525	0.540	0.0342
4	7.999	0.1212	8.571	0.572	0.0382
5	8.013	0.1819	8.461	0.448	0.0237
6	8.027	0.1111	8.580	0.553	0.0356
7	8.041	0.1313	8.558	0.517	0.0312
Total					0.2319

With respect to the relationship for part 1 $\chi_0^2(0.9295) < \chi_{0.05,\,1}^2(3.84),$ and for part 2 $\chi_0^2(0.2319) < \chi_{0.05,\,1}^2(3.84),$ we can say that for both components the derived Generalized Lambda Distribution fits the data.

Design of tolerance

In this section, for designing tolerance we must first determine the mean and standard deviation for both components.
For part 1 from equations 22 and 23 we have:

$$\mu_1 = \lambda_1 + \frac{\lambda_2}{(\lambda_4+1)} - \frac{\lambda_3}{(\lambda_5+1)}$$

$$= 10.7436 + \frac{4.0029}{(2.0851+1)} - \frac{2.1856}{(52.1962+1)}$$

$$= 12.00$$

and

$$\sigma_1^2 \equiv E(X^2) - [E(X)]^2$$

$$= \int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5})^2 dy - \left(\lambda_1 + \frac{\lambda_2}{(\lambda_4+1)} - \frac{\lambda_3}{(\lambda_5+1)}\right)^2$$

$$= \int_0^1 (10.7436 + 4.0029 y^{2.1856} - 2.0851(1-y)^{52.1962})^2 dy - (11.82)^2$$

$$= 0.00049$$

so we have

$$\sigma_1 = \sqrt{0.00049} = 0.022$$

and for part 2 from equations 22 and 23 we have:

$$\mu_2 = \lambda_1 + \frac{\lambda_2}{(\lambda_4+1)} - \frac{\lambda_3}{(\lambda_5+1)}$$

$$= 1.2427 + \frac{6.7886}{(0.2331+1)} + \frac{4.3945}{(2.5099+1)}$$

$$= 8.00$$

and

$$\sigma_2^2 \equiv E(X^2) - [E(X)]^2$$

$$= \int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5})^2 dy - \left(\lambda_1 + \frac{\lambda_2}{(\lambda_4+1)} - \frac{\lambda_3}{(\lambda_5+1)}\right)^2$$

$$= \int_0^1 (10.2427 + 6.7886 y^{0.2331} + 4.3945(1-y)^{2.5099})^2 dy - (8.00)^2$$

$$= 0.000442$$

so we have

$$\sigma_2 = \sqrt{0.000442} = 0.021$$

Now we can determine tolerance of components, using equation 15 as bellow:

$$T_i = g_i \sigma_i$$

For part 1

$$T_1 = 6 \times 0.022$$
$$= 0.132$$

And for part 2

$$T_2 = 6 \times 0.021$$
$$= 0.126$$

By considering this tolerance, now to specify part 1 we can have:

Specification part 1=12 ± 0.066 andd specification part 2 =8 ± 0.052

If it is assumed that these two parts are assembled together, by applying the Equation (11), the sum of their tolerance equals:

$$T_p = \sqrt{T_{p1}^2 + T_{p2}^2}$$

$$= \sqrt{(0.066)^2 + (0.052)^2}$$

$$= 0.084$$

While if this is the sum of tolerance in question and is allocated to each part with an assumed proportion of 50%, for instance, the tolerance of each part amounts to 0.042, the proposed method yields tolerance of 0.066 and 0.052, respectively. However, if the assumptions 1, 2, 3, and 5 of the probabilistic relationship section are true, which is often the case, it can be expected that by considering the proposed tolerance, it is easier to produce such parts.

Conclusion

Tolerance is one of the critical parameters in designing components specifications and also of crucial importance for the customer satisfaction. Therefore, it is of special significance to determine it. In some cases the linear dimension does not have to be tight tolerance but the form does. The most common reason is to ensure the functionality. This way the dimensional tolerance does not have to be unnecessarily tight, which would increase costs. The form tolerances may be achieved more easily or at lower costs.

In this research we introduced an alternative to the design of tolerance using a statistical method. When the distribution of products is unknown, predicting the distribution of specification and their parameters need cost and time. Instead, the generalized lambda distribution can be deployed, which applies to every known or unknown distribution. After estimating the parameters of the distribution, tolerance can be calculated in the next step. So, we deployed a procedure that allows us to compute parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and λ_5 of GLD. While approximation errors may have an impact on the quality of the fitted distribution to some degree, the fact remains that even if the five moments are matched exactly, one cannot be assured that the resulting theoretical distribution will perfectly match the empirical distribution[9]. The quality of the fit can be ascertained only through a goodness-of-fit test.

To compute and estimate the parameters, various methods can be thought of such as: Percentile Matching (PM), the moment matching (MM), Probability-Weighted Moment (PWM), Minimum Cramér-Von Mises (MCM), Maximum Likelihood (ML), Pseudo Least Squares (PLS), Downhill simplex method, and starship methods, In this study, due to the limited access to an appropriate software program, Moment Matching (MM) Estimates were employed without considering the target function. But for the future research, Maximum likelihood can be utilized. Also, Simplex method can be used for solving equations and estimating parameters. Since the Generalized Lambda Distribution is still an innovative approach, it is necessary to investigate different aspects of this distribution.

Reference

1. Syrjala T., Tolerance Design and Coordinate Measurement in Product Development, Helsinki university of technology, Department of Mechanical Engineering, Thesis submitted in partial fulfillment of the requirements for the degree of M.S. in Engineering, (**2004**)

2. Chandra M. Jeya, Quality control, CRC Press LLC, 5-22 (**2001**)

3. Yourstone S. and Zimmer W., Non-normality and the design of control charts for average, *Decision sciences*, (**23**), 1099-1113 (**1992**)

4. Kittlitz R.G., Transforming the exponential for SPC applications, *J. of quality technology*, (**31**), 301-308 (**1999**)

5. Peam W.L., Kotz S. and Johnson N.L., Distribution and inferential properties of process capability indices, *J. of quality technology*, **24**(**4**), 216-231 (**1992**)

6. Thakur N.S., Yadav K. and Pathak S., On Mean Estimation with Imputation in Two- Phase Sampling Design, *Res. J. of Mathematical and Statistical Sci.*, **1**(**3**), 1-9 (**2013**)

7. Rekha R.C. and Vikas S., Retailer's profit maximization Model for Weibull deteriorating items with Permissible Delay on Payments and Shortages, *Res. J. of Mathematical and Statistical Sci.*, **1**(**3**), 16-20 **(2013)**

8. Roman U.C., Porey P.D., Patel P.L. and Vivekanandan N., Assessing Adequacy of Probability Distributional Model for Estimation of Design Storm, *ISCA J. of Engineering Sci.*, **1(1),** 19-25 **(2012)**

9. Fournier B., Rupin N., Bigerelle M., Najjar D., Iost A., Application of the generalized lambda distribution in a statistical process control methodology, *J. of Process control*, (**16**), 1087-1098 **(2006)**

10. Gunter B., The use and Abuse of C chart 1-4. *Quality progress*, part 1: **22**(**1**), 72-73, part 2: **22**(**3**) 108-109, part 3: **22**(**5**), 79-80, and part 4: **22**(**7**), 86-87, (**1989** a-d)

11. Tukey J.W., The future of data analysis, *annals J. of mathematical statistics,* **33**(**1**), 1-67 **(1962)**

12. Joiner, B. L., Rosenblatt, J. R., Some properties of the range in samples from Tukey's symmetric lambda distribution, *J. of the American statistical association,* (**66**), 394 (**1971**)

13. Ganeshan R., Are more supplier better? generating the Gau and Ganeshan procedure, *J. of Oper. Res. Soc.*, (**52**), 122-123 (**2001**)

14. Ozturk A. and Dale R.F., A study of fitting the generalized lambda distribution to solar radiation data, *J. Appl. Meteorol*., (**21**), 995-1004 (**1982**)

15. Delaney H.D. and Vargha A., The effect on non-normality on student's two-sample t-test the annual meeting of the American educational research association, New Orlean, (**2000**)

16. Gawand H., Mundada R.S. and Swaminathan P., Design Patterns to Implement Safety and Fault Tolerance, *Int. J. of Computer Applications* (**18**)**2**, (**2011**)

17. Dengiz B., The generalized lambda distribution in simulation of m/m/1 queue systems, *J. Fac. Engng. Arch. Gazi univ.,* (**3**), 161-171 (**1988**)

18. Zaven A., Karian and Edvard J., Dudewiz, Fitting statistical distributions The generalized lambda distribution and generalized bootstrap methods, *CRC press*, (**2000**)

19. Chowdhury D., and Arbabian M.A., Design of Robust CMOS Circuits for Soft Error Tolerance, Department of EECS, Univ. of California, Berkeley, CA 94720, (**2011**)

20. Shweta Ms., Meshram S., and Ujwala Ms., A. Belorkar, Design Approach For Fault Tolerance in FPGA Architecture, International *J. of VLSI design & Communication Systems (VLSICS)*, **2**(**1**), (**2011**)

21. Kuang W., Xiao E., Ibarra C.M. and Zhao P., Design Asynchronous Circuits for Soft Error Tolerance, University of Texas - Pan American, Edinburg, (**2007**)

22. Aljazar A-N. L., Generalized Lambda Distribution and Estimation Parameters, The Islamic University of Gaza, M.S. Theses, (**2005**)

23. Tarsitano A., Fitting The Generalized Lambda Distribution to Income data, COMPSTAT'2004 Symposium, Physica-Verlag, (**2004**)

24. Yi X., and Jerome, Y., Continuous Setting and Gaussian Generalized Lambda Distribution Model for Synthetic CDO Pricing, Hong Kong University of Science and Technology, (**2008**)

25. Bigerelle M., Najjar D., Fournier B., Rupin N., Iost A., Application of lambda distribution and bootstrap analysis to the prediction of fatigue lifetime and confidence intervals, *Int. J. Fatigue*, (**28**), 223-236(**2006**)

26. Acar A., Rais-Rohani M. and Eamon C.D., Reliability Estimation using Dimension Reduction and Extended Generalized Lambda Distribution, 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Schaumburg, IL, (**2008**)

27. Karvanen J., and Nuutinen A., Characterizing the generalized lambda distribution by L-moments, *Math. ST*, (**2007**)

28. Karian Z.A., and Dudewicz E.J., Fitting statistical distributions: the generalized lambda distribution and generalized bootstrap method, CRC press, (**2000**)

29. Tarsitano A., Estimation of the generalized lambda distribution parameters for grouped data, *J. of Communication in statistics theory and methods*, (**34**), 1689-1709 (**2005**)

30. Ramberg J., and Schmeiser B., An approximate method for generating asymmetric random variables, *communications of the ACM*, **7**(**2**) 78-82 (**1974**)