Qual Quant (2012) 46:1097–1111
DOI 10.1007/s11135-011-9550-x

A process control method based on five-parameter generalized lambda distribution

Majid Nili Ahmadabadi · Yaghub Farjami · Mohammad Bameni Moghadam

Published online: 3 August 2011

Abstract The use of control charts in statistical quality control, which are statistical measures of quality limits, is based on several assumptions. For instance, the process output distribution is assumed to follow a specified probability distribution (normal for continuous measurements and binomial or Poisson for attribute data) and the process supposed to be for large production runs. These assumptions are not always fulfilled in practice. This paper focuses on the problem when the process monitored has an output which has unknown distribution, or/and when the production run is short. The five-parameter generalized lambda distributions (GLD) which are subject to estimating data distributions, as a very flexible family of statistical distributions is presented and proposed as the base of control parameters estimation. The proposed chart is of the Shewhart type and simple equations are proposed for calculating the lower and upper control limits (LCL and UCL) for unknown distribution type of data. When the underlying distribution cannot be modeled sufficiently accurately, the presented control chart comes into the picture. We develop a computationally efficient method for accurate calculations of the control limits. As the vital measure of performance of SPC methods, we compute ARL's and compare them to show the explicit excellence of the proposed method.

Keywords Control chart · GLD · ARL · Short run production · LCL · UCL

M. N. Ahmadabadi (✉)
Department of Management, Najafabad Branch of Islamic Azad University, Najafabad, Iran
e-mail: Nili2536@Gmail.com

Y. Farjami
Department of Technology, Qom University, Qom, Iran
e-mail: Farjami@ut.ac.ir

M. Bameni Moghadam
Department of Statistics, Allameh Tabataba'i University, Tehran, Iran
e-mail: BameniMoghadam@Gmail.com

1 Introduction

The statistical process control (SPC) methodology is mainly used as a tool to understand, model and reduce the variability of an industrial process over time. In conventional SPC, parts are sampled during the production process. Relevant characteristics are checked and the mean, as well as deviations, monitored with the aid of control charts. SPC mainly includes two functions: a control chart and process capability analysis. The control chart is used to monitor whether the process is in control or not. There is a life cycle in the application of control charts. In the preparatory stage, an investigation of the process is required to determine critical variable(s) and probability distribution function about it. To sustain interest, charts must be changed over the life of the application. Eventually, of course, when continued control is assured, the charts should be withdrawn in favor of spot checks as appropriate. Control charts show when action should be taken and, equally important, when no action should be taken. Most control charts are distribution-based procedures in the sense that the observations made on the process output are assumed to follow a specified probability distribution (usually, normal for continuous measurements and binomial or Poisson for attribute data) (Bakir 2004). For a thorough reference on distribution-based control charts, the reader is referred to Montgomery (2001). The Shewhart $\bar{X}$ chart (1931) is one of the most widely used charts in quality control. However, it has two main disadvantages: (a) it is not sensitive to small shifts in the process mean, and (b) for small sample sizes, there is a valid concern that the performance of this chart can be adversely affected if the normality assumption is violated.

It has long been realized that the variability associated with many engineering processes does not have a Gaussian distribution (Chang and Bai 2001; Pearson 2001; Pyzdek 1995). If the measurements are not normally distributed, the statistic $\bar{X}$ will be approximately normally distributed only when the sample size n is sufficiently large (based on the central limit theorem). Unfortunately, when a control chart is applied to monitor the process, the sample size n is always not sufficiently large due to the sampling cost. For heavily skewed distributions, the Type I risk probability (i.e., false alarm rate) of a control chart grows larger when the skewness increases (Yourstone and Zimmer 1992). Excessive false alarms can lead to unnecessary process adjustment and loss of confidence in the control chart as a monitoring tool. Therefore, if the measurements are not normally distributed, the traditional way may not be appropriate for designing the control chart. Several practical examples of non-normal processes were often reported.

Gunter (1989) discusses the problem that processes with process outputs distributed in very different ways actually can give the same value of some of the measures of quality control, as illustrated in Fig. 1. Because all process outputs in Fig. 1 have the same μ and the same σ, they will all have the same value of UCL and LCL. Four differently shaped distributions describing different process outputs. The four distributions have the same expected value equal to 0 and the same standard deviation equal to one. The output of process 1 follows a skew distribution with a finite lower boundary, namely χ^2-distribution with 4.5 degrees of freedom. The output from process 2 follows a heavy-tailed distribution, namely a t-distribution with 8 degrees of freedom. Finally, the output from process 3 follows a uniform distribution. A normal distribution has also been included for comparison purposes. From Katz and Johnson (1995), who reproduced the figure from Gunter.

From a practitioners point of view, it is most desirable that a given value received when estimating the quality control measure (for example, control limits), derived from real distribution of process and then shows the real behavior of system. The example in Fig. 1 shows that existing control limits do not fulfill this requirement.

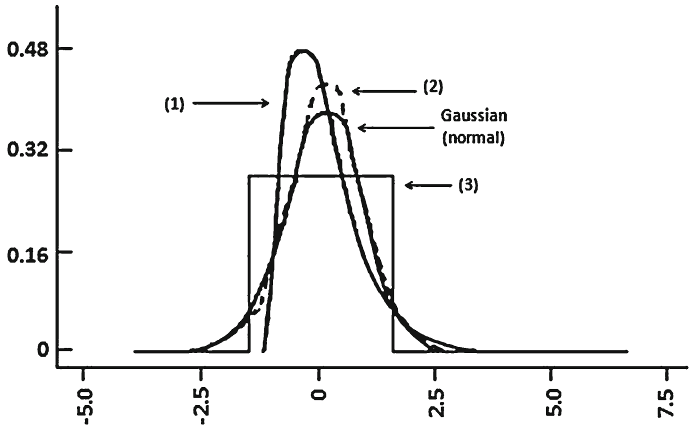

Fig. 1 Four different distribution that have similar averages, variances and shapes

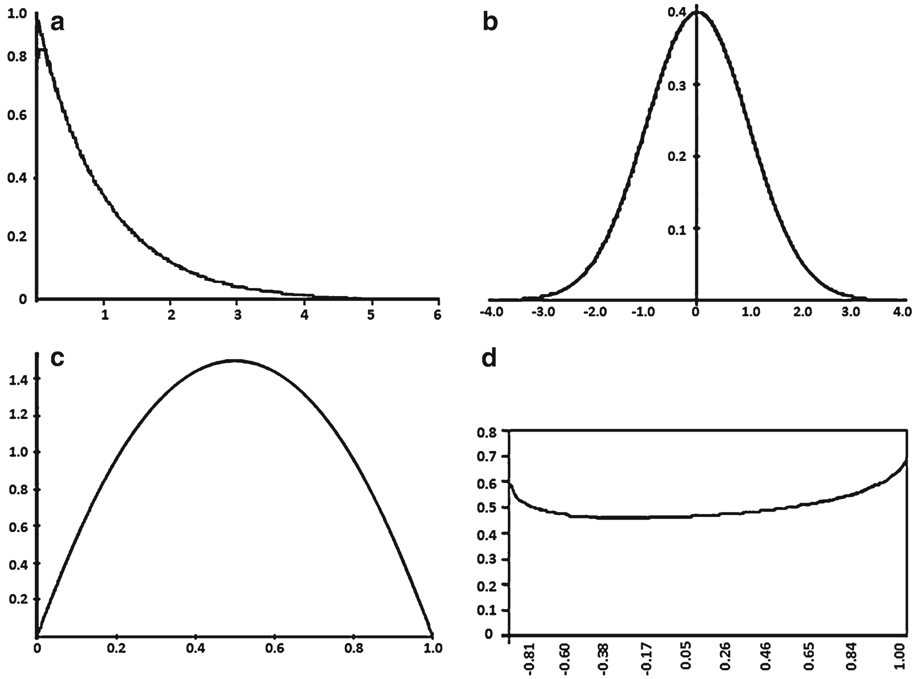

Fig. 2 Some of GLD distributions that conform to known distributions. **a** Distribution of GLD (0.0069, −0.0011, −0.0000, −0.0011) conforms to distribution of exponential (1). **b** Distribution of GLD (0, 0.1975, 0.1349, 0.1349) conforms to distribution of normal (0, 1). **c** Distribution of GLD (0.5, 1.9693, 0.4495, 0.4495) conforms to distribution of beta (1, 1). **d** Distribution of on GLD (0, 1, 1.4, 1.6)

Due to the complexity of the data, it is sometimes difficult for people to interpret the various approaches of statistical process control. Even today, the use of the classical control charts is not widespread in the industry (Woodall 2000). when average (μ) and standard deviation (σ) of the normal distribution are unknown, UCL cannot be calculated, and often estimates of μ and σ are simply plugged in without further adjustment, although the dangers have been pointed out from time to time in the literature; Chakraborti (2000), Nedumaran and Pignatiello (2001), and Albers and Kallenberg (2004). These charts, also, do not provide a method to signal improvement when the lower control limit is equal to zero. When the Shewhart user encounters a lower control limit that is negative or zero there is insufficient information in a single low observation to signal that improvement has occurred.

The traditional statistical inference, based on normal and resulting normal based distributions (such as, Chi Square and Fisher) is almost automated by ready methods and formulas. A departure from this atmosphere has a payoff on computational complexity. An efficient algorithm is devised to manage the relevant computations, and does the job fairly good.

2 Literature review

Basically, there are two solutions to the problem of how to use quality control charts when the output of the process monitored is non-normally distributed: (a) do something about it; (b) do not do anything about it. If the aim is to improve one specific process within an organization it would be sufficient to check the distribution of the process output for a specific characteristic for several processes.

Yourstone and Zimmer (1992) examined the effect of departures from normality (as measured by skewness and kurtosis) on the design and performance of the traditional Shewhart X-bar chart. Kittlitz (1999) suggested that the long-tailed positively skewed exponential distribution could be made into an almost symmetric distribution by taking the fourth root of the data. The transformation data can then be plotted conveniently on an individual's chart, EWMA chart, or CUSUM chart for statistical process control. When the assumption of normality is violated, the average run length (ARL) of the individual control chart is adversely averted. The most important measure for performance of SPC methods is ARL as stated by many authors (Maravelakis et al. 2002; Michael and Khoo 2004).

Peam et al. (1992) proposed a method which is quite robust over a wide range of distributions. In the Peam–Kotz–Johnson method, equal tail areas are no longer guaranteed, and the skewness and kurtosis do not need to be estimated. Still, the Peam–Kotz–Johnson method relies on the assumption that the output from the process has un-normal shape relatively close to a gamma distribution. In practice, not all process outputs will fulfill this requirement though.

Pyzdek (1993) also proposed the factors to calculate control limits for thirteen different samples and five different samples of size (small samples) by choosing the probabilities of type I and type II errors. Somerville and Montgomery (1996) study the errors that can occur in calculating some of the quality control measures for a non-normal distribution and making inferences about the PPM (parts per million) non-conforming with the normality assumption. Using the gamma, lognormal, Weibull and r-distribution they find that although the magnitude of the error can vary substantially depending on the true distribution parameters, the error is meaningful in nearly all cases. Yang (1999) proposed the design of the R control chart for short production runs. Lee and Amin (2000) proposed to obtain long-term and short-term estimates of the process tolerances and to use them in capability analysis as an alternative method to the use of capability ratios.

A number of papers have proposed the use of the "bootstrap method" for obtaining control limits in SPC when the process distribution is unknown (Teyarachakul et al. 2007). Riaz (2008) proposed a dispersion control chart namely Q chart (a threshold control chart according to Farnum (1994) classification of control charts) based on inter-quartile range and estimate of σ, for monitoring changes in process dispersion. For all Gaussian, uniform, lognormal and Weibull distributions, the GLD allow more flexible modeling of shapes than the classical Gaussian law while keeping a very accurate approximation of the predicted control limits considered in SPC methodology. In addition, even in an unfavorable case, the use of GLD-based modeling requires utmost a hundred of sample data to estimate the SPC limits with an acceptable accuracy.

This study aims to present an alternative statistical approach to calculate control limits, which avoids the previously cited drawback related to statistical inference. In this approach, a very flexible family of distributions called GLD is used to model accurately a series of data about the nature and the shape of the unknown underlying distribution. If a good estimate of the population distribution were possible, control limits could be prescribed more accurately. In this paper, the five-parameter GLD is used for modeling unknown distributed data in order to predict accurately the control limits considered in SPC methodology. A numerical example is illustrated and a performance computation is performed, using ARL's. We should mention here that the computations for ARL's are usually a tedious and difficult task in SPC and even there are some controversies on the subjects for some SPC methods. We will overcome this obstacle by using a consistent simulation method for computing the relevant ARL's.

3 Generalized lambda distribution

The GLD firstly introduced by Tukey (1962) and further described by others (Joiner and Rosenblatt 1971). It have already been shown to fit accurately most of the usual statistical distributions (like Gaussian, lognormal, Weibull). Since the GLD model has a great flexibility in fitting to statistical data and estimating distributions proved to be a valuable tool for researchers and scientists, thus GLD method has been used in the most of scientific fields such as parameter estimation, fitting distributions to data and in simulation researches that primarily include univalent data generation. For example, the GLD has been used in studies that include such topics or techniques as: independent component analysis (Karvanen 2003), operations research (Ganeshan 2001), psychometrics (Elaney and Vargha 2000), Engineering (Upadhyay and Ezekoye 2008), corrosion (Najjar et al. 2003), meteorology (Öztürk and Dale 1982), fatigue of materials (Bigerelle et al. 2005), statistical process control (Fournier et al. 2006) and simulation of queue systems (Dengiz 1988). Excellent literature of the generalized lambda distribution, its applications, and parameter estimation methods appear in Karian and Dudewicz (2000).

A generalized lambda distribution was described in terms of the function Q, the inverse of the cdf (cumulative distribution function) as follows:

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

That $y \in [0, 1]$, λ_1, λ_2 are the position and scale parameters respectively, and λ_3, λ_4 are related to the skewness, and the kurtosis of the distribution (Karian and Dudewicz 2000). Capability of this distribution to estimating different distributions is shown as following figures:

As one can see, this distribution is capable of estimating types of well-known (and even unknown) distributions for any different values of λ. Recently, Tarsitano (2005) introduced the properties of a five-parameter generalization of the lambda distribution and studied about it, in order to obtain smoothed analytic representations for grouped data. The five-parameter GLD is specified by the quantile function

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

where λ_1 is a location parameter, λ_2 and λ_3 are linear parameters prevalently related to the scale of the variable, and λ_4 and λ_5 are exponential parameters determining the shape of the quantile function. The two linear parameters contribute to specify the relative weights of the tails and to avoid imposing upon the fitted curves constraints not present in the data. We can expect this model to assume a wider variety of shapes than the four shape parameters GLD. The probability density function (p.d.f.) of this newest version of GLD is

$$f(x)=\frac{dy}{dx}=\left(\frac{dx}{dy}\right)^{-1}=\frac{1}{\lambda_2\lambda_4 y^{\lambda_4-1}+\lambda_3\lambda_5(1-y)^{\lambda_5-1}} \tag{2}$$

Finally, using (1) and (2), calculation of the *m-th* moment (M_m) is possible through

$$M_m=E(x^m)=\int_{-\infty}^{+\infty}x^m f(x)dx=\int_0^1(\lambda_1+\lambda_2 y^{\lambda_3}-\lambda_4(1-y)^{\lambda_5})^m dy \tag{3}$$

3.1 Parameter estimation and distribution fitting

Although there are several methods for estimating the parameters of the GLD, we will follow the moment-matching method that was proposed in Ramberg and Schmeiser (1974). The principal objective in this method is finding the parameters of fitting distribution, which moments of distribution match closely with the moments of the empirical data or distribution. For estimating parameters of GLD, the method of moment can be described briefly as follows:

Given the GLD distribution, find parameters $\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5$ so that the first five moments of the GLD match the corresponding first five moments of the original data. This will lead to a system of five nonlinear equations which can be solved by well known and dependable numerical methods. Often, more than one numerically acceptable solution is available. Thus, a goodness-of-fit test should be performed to establish the validity of the results. If this test fails, it is necessary to use numerical procedures to find suitable parameters. Several studies have been done about the method of moments (Karian and Dudewicz 2003; Headrick and Mugdadi 2006; Asquith 2007; Karvanen and Nuutinen 2008; Su 2007). To verify the goodness-of-fit, the sample data are first arranged in a frequency histogram having k class intervals. Let O_i be the observed frequency and E_i be the expected frequency in the *i-th* class interval. For finding E_i, the y_i values are obtained by solving the percentile function $X(y)$ at each class interval endpoint. Multiplying those y_i-values with the number of observations will give the expected cumulative frequencies from which E_i-values can easily be obtained. The test statistic has approximately a chi-square distribution with k-L-1 degrees of freedom, where L (=5, in this case) represents the number of parameters of the hypothesized GLD distribution.

$$\chi_0^2=\sum_{i=1}^{k}\frac{(O_i-E_i)^2}{E_i} \tag{4}$$

The null hypothesis, which states that the sample is drawn from GLD, is rejected if the calculated value of the test statistic $\chi_0^2>\chi^2_{\alpha,k-6}$ where α is the significance level.

3.2 Introduction of GLD in SPC methodology

In process control methodology, accurate modeling of the process probability density function (pdf) must be used. Forbes et al. (2004), who used a Gram–Charlier pdf to obtain a standard state-space model of the process. In the framework of dynamic processes, Wang (2000) proposed a general time-evolving model of the process pdf based on B-splines. Several other attempts have been made to find a family of statistical distributions that could accurately model a broad class of distribution shapes. For example, Albers et al. (2004) used the normal power family and Lin and Chou (2005) used the Burr distribution.

GLD's flexibility with regard to the shape of many distributions enables to consider the GLD in a more global SPC strategy without making any preconceived and risky choice about the unknown underlying distribution. Spedding and Rawlings (1994) address the problem of non-normality by using the GLD Distribution which provide the good model for production's data distribution function. Fournier et al. introduced the use of the GLD in the SPC methodology (Fournier et al. 2003). Pal (2005) used the lambda distributions to calculate the capability indices (CP and CPK) widely used in statistical process control.

In this paper, a control chart is proposed that monitors the conformance of a sample using the quantile or inverse cumulative distribution function named five-parameter Generalized Lambda Distribution (GLD). This method also helps to detect changes in the distributional shape, which may be undetected in control charts that are based on summary statistics.

4 A control chart based on 5-parameter GLD

Control limits are limits that include a specific proportion of the population at a given confidence level. In the context of process control, they are used to make sure that production will not be outside specifications. The proposed method can be applied to almost all SPC methods, such as EWMA, CUSUM. Here we will only concentrate the Shewhart type control charts as an example to demonstrate the method.

In this section, we develop a Shewhart control chart based on finding process's data distribution by 5 parameters GLD. Because the population distribution is unknown in practice, as Teyarachakul et al. (2007) did, the idea of this paper is to treat the given dataset as the population. In other words, the empirical distribution function is used to estimate the distribution function of the underlying population. Therefore, a GLD based control chart can be constructed in the following steps:

Step 1: set up production line and make sure it is in control.

Step 2: determine probability values of type-1 error (α), and type-2 error (β), in control variable value under ideal condition (μ), and accepted deviation value from ideal value by the use of real limitation (such as expense, time, technical limitations and etc).

Step 3: find OC curve which corresponds to above data, and extract sample size (n).

Step 4: Take a sample with size n.

Step 5: find λ values in Eq. (1) by using the method of moments (or any method mentioned in 3.1), as the best GLD distribution for sample distribution is resulted. Call $\hat{\lambda}$for any value of finding λ.

Step 6: Calculate the Parameters of GLD based control Chart (CL, LCL and UCL). There are two approaches to express these parameters, namely probability limits approach and 3-sigma limits approach. In case of asymmetric distributional behavior of relevant estimator, the probability limits approach is preferred. If the distributional behavior of relevant estimator is symmetric or nearly symmetric, then 3-sigma limits approach is a good alternative. The parameters of proposed chart using both the approaches are expressed in the following two subsections:

4.1 Probability limits approach

Assuming the probability of making a Type-I error to be less than a specified value say α, control limits (which are actually the true probability limits) for proposed chart are defined as:

LCL $= x_L$ with $P(X = x_L) \leq \alpha_L$, UCL $= x_u$ with $P(X = x_u) \geq 1 - \alpha_u$, where $\alpha = \alpha_L + \alpha_u$, and CL $= \mu$

According to Eqs. (1) and (3), this means that:

$$LCL = \hat{\lambda}_1 + \hat{\lambda}_2 \alpha_L^{\hat{\lambda}_3} - \hat{\lambda}_4 (1 - \alpha_L)^{\hat{\lambda}_5} \quad \text{and} \quad UCL = \hat{\lambda}_1 + \hat{\lambda}_2 (1 - \alpha_U)^{\hat{\lambda}_3} - \hat{\lambda}_4 \alpha_U^{\hat{\lambda}_5}$$

4.2 3-Sigma limits approach

Parameters of Q chart with the usual 3-sigma control limits are given as:

UCL $= \mu + 3\sigma$, CL $= \mu$, LCL $= \mu - 3\sigma$. This means:

$$LCL = \mu - 3\sqrt{E(x^2) - (E(x))^2}$$

Using Eq. (3) gives the following result:

$$LCL = \mu - 3 \times \sqrt{\int_0^1 (\hat{\lambda}_1 + \hat{\lambda}_2 y^{\hat{\lambda}_3} - \hat{\lambda}_4 (1-y)^{\hat{\lambda}_5})^2 dy - \int_0^1 (\hat{\lambda}_1 + \hat{\lambda}_2 y^{\hat{\lambda}_3} - \hat{\lambda}_4 (1-y)^{\hat{\lambda}_5}) dy)^2}$$

and

$$UCL = \mu + 3 \times \sqrt{\int_0^1 (\hat{\lambda}_1 + \hat{\lambda}_2 y^{\hat{\lambda}_3} - \hat{\lambda}_4 (1-y)^{\hat{\lambda}_5})^2 dy - \int_0^1 (\hat{\lambda}_1 + \hat{\lambda}_2 y^{\hat{\lambda}_3} - \hat{\lambda}_4 (1-y)^{\hat{\lambda}_5}) dy)^2}$$

A problem of LCL calculated by this approach is that, sometimes LCL results into a negative value. A negative value for dispersion measure has no realistic meanings. Therefore, in such situations it is assigned the value of zero.

Step 7: pick up the next sample and using control limits determine whether the production process is in control or not.

Step 8: if the production process is in control, then go back to step 5, else stop the production, find the defect, remove it and go back to step 4.

Using this method, in each time of sampling you will get a distribution which states the real distribution of production line outputs and by assessing them you can understand the rate of changes in production line. Furthermore, as at the time of installing production line, people and machinery have not been perfectly adjusted or they do not coincide with ideal state, in the first repetitions, wider control limits and in later ones, narrower and stricter control limits

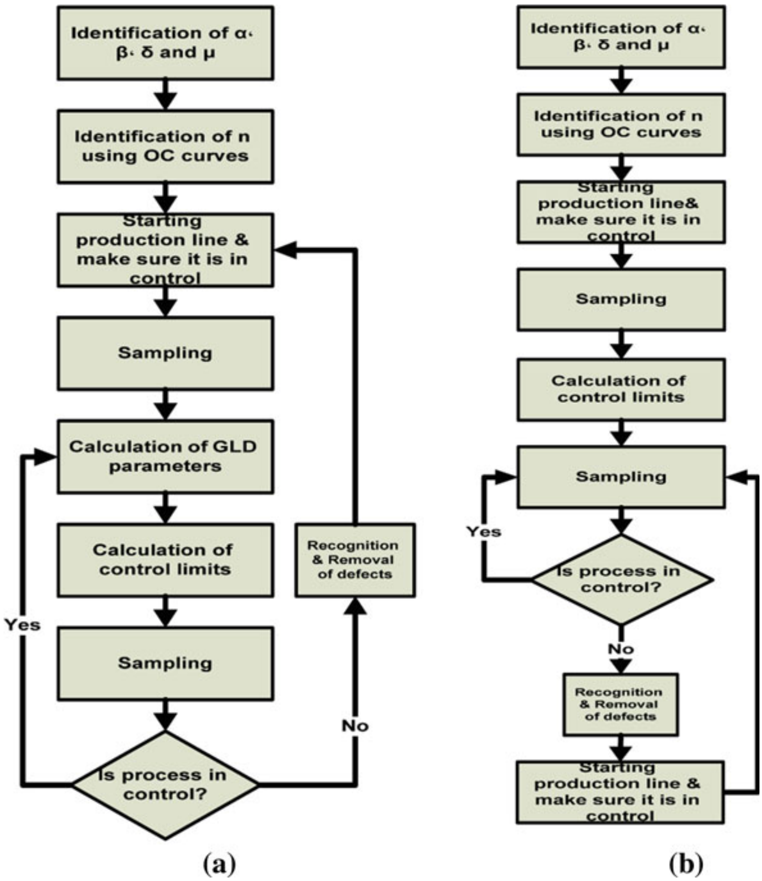

Fig. 3 The proposed (**a**) and usual (**b**) algorithms of control chart structuring

are expected to be calculated. In this way, applying this method can show learning in the system and monitor it. Figure 3 illustrates the usual algorithm of control chart Structuring, proposed algorithm, and improvement process using the new proposed method.

5 Numerical example

In this section, we use empirical data to illustrate the calculations needed to construct the proposed control chart. To do this, we illustrate the method for fitting a GLD curve to sample data which is far from being normal. The estimated curve is then used to compute the control chart parameters.

After implementing the steps 1 and 2, the values of α, β and σ were chosen while the value of n from OC chart is equal to 200 and ideal value for the under control value $\mu = 0.5$.

In the next step a sample data of size 200 are collected over a short period of time. The randomness of the data is verified to ensure the stability of the process. The sample data are given in Appendix Table 3. The frequency distribution and the moments of the sample data are given in Table 1.

Table 1 Frequency distribution and the moments

Class intervals	Frequency	Moments
0.05–0.14	10	
0.14–0.23	15	
0.23–0.32	23	
0.32–0.41	24	$M_1 = 0.4968$
0.41–0.50	24	$M_2 = 0.2926$
0.50–0.59	30	$M_3 = 0.1903$
0.59–0.68	28	$M_4 = 0.1324$
0.68–0.77	25	$M_5 = 0.0966$
0.77–0.86	14	
0.86–0.95	7	

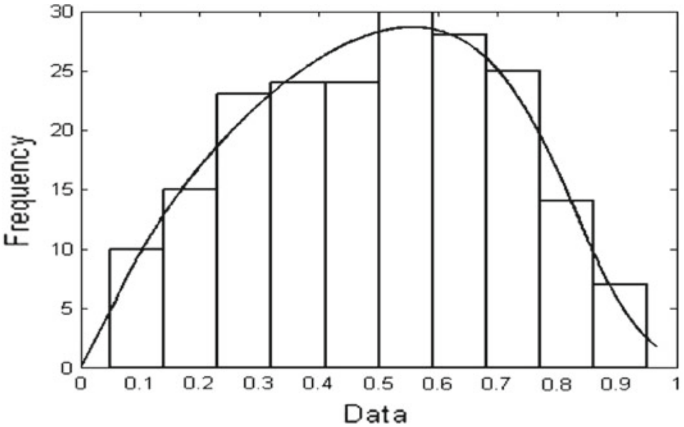

Fig. 4 Histogram of original data and fitted GLD density curve

Above moments calculated using sample data (Appendix Table 3) and below equation:

$$M_m = \frac{\sum_{i=1}^{N} x_i^m}{N}$$

In the next step, using moment-matching method and according to (3), the m-th moment of GLD calculated as a function of five parameters λ_1 to λ_5. The empirical moments of data distribution derived from Table 1. This led to a system of five equations:

$$\int_0^1 (\lambda_1 + \lambda_2 y^{\lambda_3} - \lambda_4 (1-y)^{\lambda_5})^i dy = M_i$$

$$1 \le i \le 5$$

After solving the system of equations, we obtain the lambda values as $\lambda_1 = 0.6682$, $\lambda_2 = 0.6451$, $\lambda_3 = 0.5802$, $\lambda_4 = 0.6451$ and $\lambda_5 = 0.1128$. The estimated m-th moments (m=1 to 5) of GLD, are respectively 0.4967, 0.2924, 0.1902, 0.1322 and 0.0964 which are very close to empirical moments. Figure 4 shows the relative frequency histogram for the sample data and the fitted probability density curve corresponding to the above lambda values.

For verifying the goodness-of-fit of the GLD distribution, X^2-test is carried out. The expected frequencies and computed X^2 (using (4)) are given in Table 2.

Table 2 Test for goodness-of-fit

Class intervals	Observed frequency	Expected frequency	Computed χ^2
Below 0.05	0	0.8	0.80
0.05–0.14	10	9	0.11
0.14–0.23	15	16	0.06
0.23–0.32	23	21	0.19
0.32–0.41	24	24.6	0.01
0.41–0.50	24	27.6	0.47
0.50–0.59	30	28.6	0.07
0.59–0.68	28	27.4	0.01
0.68–0.77	25	23.2	0.14
0.77–0.86	14	15	0.07
0.86–0.95	7	5.6	0.35
Above 0.95	0	1.2	1.20
Total	200	200	3.49

The computed value of $X^2 = 3.49$ is much less compared to the tabulated value of $X^2_{0.05,6} = 12.06$ which indicates that the above GLD curve fits the sample data quite well.

In the next step, from the fitted GLD curve, we calculate the LCL and UCL values by using Probability Limits Approach at the p-values as 0.001 and 0.999 respectively, as $\mathrm{LCL}_{\mathrm{GLD}} = 0.0349$, $UCL_{\mathrm{GLD}} = 1.017$ and CL = 0.5.

In the next section, in order to show efficiency of suggested method, parameters of control diagram are calculated with usual and proposed methods. Then, values of the type-2 error and by using that values of ARL_1 for different values of deviation (δ), in both diagrams will be calculated and compared with each other.

6 Efficiency comparison

In this section, assuming normal distribution, firstly, control limits will be calculated. To do this, suppose we are given X_{ij} for *i* = *1, …,m* and *j* = *1, …,n,* where n is the number of samples and m is the sample size. All these observations are assumed to come from the process when it was in-control. Suppose we need to construct a chart with a false alarm rate of 2a. Under the normality assumption, the control chart parameters, CL, LCL and UCL, are

$$\mathrm{CL} = \bar{x}_{nm}, \mathrm{LCL} = \bar{x}_{nm} - \frac{z_{\frac{\alpha}{2}} s}{\sqrt{n}} \quad \text{and} \quad \mathrm{UCL} = \bar{x}_{nm} + \frac{z_{(1-\frac{\alpha}{2})} s}{\sqrt{n}} \tag{5}$$

where $\bar{x}_{nm}$ is the average of n sample averages, s is an estimate of the σ of the process and z is the normal deviate $\mathrm{N}(\mu, \sigma)$. Note that $\mathrm{Z}_{(1-\alpha)}\mathrm{S}/\sqrt{n}$ is an estimate of the $100(1-\alpha)\%$ point of the distribution of $\bar{x}_{nm}$ when X is normal.

Here, we are given n = 1 and m = 200 (all of existing data). Therefore, the average and standard deviation of samples respectively are 0.4968 and 0.2138. Then, using (5), the CL, LCL and UCL values at $\alpha = 0.002(\mathrm{Z}_{0.001} = -3.09)$ and n = 1, respectively are 0.4968, –0.166 and 1.1596.

The type I error, α, is the probability that a point on the X control chart falls outside the control limits and the corresponding in-control average run length (ARL_0), is $1/\alpha$. The type II error, is the probability that a given sample will not signal an out-of-control when the process mean has indeed shifted to μ_1 and is given by $\beta(\mu - \mu_1) = P(LCL < X < UCL|\mu = \mu_1)$

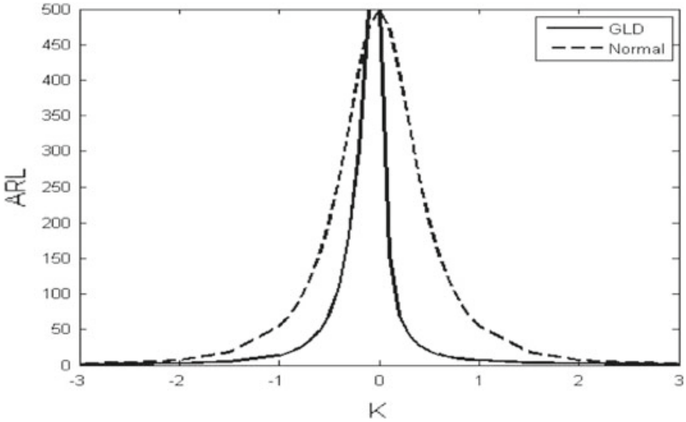

Fig. 5 Average run length calculated by GLD and traditional methods at $\alpha = 0.002, n = 1$

and the out-of-control average run length (ARL_1) is

$$ARL_1 = \frac{1}{1-\beta} \tag{6}$$

Under the assumption of normal distribution for x, simplifying this equation gives

$$\beta(\delta) = P(Z \leq Z_{\alpha/2} - \delta\sqrt{n}) - P(Z \leq -Z_{\alpha/2} - \delta\sqrt{n}) = \int_{-z_{\alpha/2}-\delta\sqrt{n}}^{z_{\alpha/2}-\delta\sqrt{n}} \varphi(z)dz \tag{7}$$

where $\delta = \frac{\mu_1 - \mu_0}{\sigma}$ (Alwan 2000).

The below algorithm shows the calculation method of $B(\delta)$ for a control chart which is calculated by GLD distribution:

Step 1: calculate GLD distribution parameters and control limits, by using suggested algorithm.

Step 2: Consider the value of δ as Kσ where $\sigma = S/\sqrt{n}$, s is the deviation of samples, $-3 \leq$ k $\leq +3$ and consider n=1 in this example.

Step 3: add the value of δ to all data, i.e., $\acute{x}_i = \acute{x}_i + \delta$ then by using of new data, compute new parameters of GLD distribution, name them $\hat{\acute{\lambda}}$, Therefore, fitted GLD distribution to new data set ($\acute{x}_l$) is:

$$X'(y) = \widehat{\lambda}'_1 + \widehat{\lambda}'_2 y^{\widehat{\lambda}'_3} - \widehat{\lambda}'_4 (1-y)^{\widehat{\lambda}'_5}$$

Step 4: find the value of y_L, y_U where $\acute{X}(y_L) = LCL$ and $\acute{X}(y_U) = UCL$. Then calculate $\beta(\delta)$ by $B(\delta) = y_u - y_L$. Using (6), ARL(δ) can be calculated.

Step 5: for all values of K, do steps 2 to 4 and correspond to each K, calculate the values of δ, then B(δ) and finally ARL(δ).

$\beta(\delta)$ for traditional control chart is calculated from (7) and based on it, ARL(δ) could be computed. By using these steps, the value of ARL was calculated by using traditional method and GLD (Fig. 5).

In-control ARL(ARL_0) should be large so that the false alarm rate can be kept small, while the out-of-control ARL($ARL_1(\delta)$) should be small so that a prompt and rapid response detection of the process change can be achieved (Torng and Lee 2008).

7 Conclusions

Using the five parameters GLD for finding control limitations has some advantages as below:

a. In non-normal distribution, using GLD, makes Type-I and II errors to be less. Therefore the cost of holding manufacturing line and/or defect products will be lower.
b. Huge sample of data, which cost a lot, is needed to assure of data distribution type. Even in this case, only the opposite hypothesis test can be rejected and data distribution can not be specified sufficiently accurately. Using method based on GLD, there is no requirement to huge sample of data.
c. By the use of GLD, defining data distribution type is not needed. So using of GLD, not only could lead lower cost and more accuracy in making control charts. But also the sample data number could be lower in some cases.
d. Number of observations needed for presented control chart setup is less than Shewhart's control chart.
e. As it is mentioned in (Fournier et al. 2006), even for normal distribution, the result of GLD method could be same with result of traditional method just with 100 number of sampling.
f. As we know, by some transformations (such as Box-Cox transformation) we can transform the data from a non-normal to a normal situation. However, these transformations are only suitable for non-negative and non-discrete distribution. However, the GLD method can be used for any set of data (Quesenberry 1995).
g. The most of continuous distributions can be defined very simply in terms of the generalized distribution function. This approach to define distributions enables the two tails of a distribution to be almost independently modeled. This is a very useful property for handling non-normal and even unknown distributions.

Appendix

See Table 3.

Table 3 Sample data

0.4468	0.0502	0.539	0.6309	0.7348	0.2361	0.4051	0.6104	0.669	0.1349
0.2107	0.8104	0.3923	0.5198	0.6193	0.5184	0.9003	0.6954	0.3241	0.8226
0.6072	0.2867	0.6944	0.5097	0.1694	0.2753	0.7476	0.6191	0.8243	0.2478
0.3924	0.485	0.8136	0.2169	0.5604	0.4122	0.1374	0.7373	0.2519	0.1811
0.3875	0.7275	0.8534	0.6502	0.5679	0.4705	0.2554	0.9499	0.2583	0.4637
0.7009	0.9296	0.4221	0.9037	0.391	0.3781	0.6033	0.2946	0.8414	0.5974
0.7588	0.6317	0.1959	0.789	0.6436	0.5254	0.3984	0.752	0.1551	0.4678
0.6648	0.5692	0.7337	0.8153	0.341	0.6373	0.6957	0.7278	0.7098	0.44
0.0828	0.6832	0.789	0.1743	0.6902	0.9177	0.2513	0.2899	0.7548	0.2824
0.6091	0.3695	0.3864	0.3706	0.7912	0.5598	0.7316	0.6036	0.3647	0.5319
0.9206	0.5424	0.6126	0.6069	0.4877	0.2413	0.2123	0.7052	0.2501	0.4382
0.3026	0.4129	0.0901	0.3246	0.4366	0.1012	0.634	0.2124	0.5098	0.6169
0.8265	0.5172	0.4326	0.4425	0.5399	0.5337	0.6105	0.3856	0.2007	0.2911

Table 3 continued

0.874	0.7554	0.1986	0.734	0.4779	0.5238	0.4458	0.2619	0.5497	0.8368
0.3006	0.4694	0.8562	0.1365	0.5422	0.522	0.6537	0.3785	0.2736	0.6827
0.3328	0.6202	0.0721	0.578	0.2534	0.2604	0.4003	0.2492	0.7766	0.5893
0.6942	0.3131	0.5601	0.4143	0.5285	0.4527	0.5094	0.3616	0.3678	0.5575
0.3645	0.6569	0.3604	0.5366	0.641	0.4497	0.6582	0.6897	0.7029	0.4749
0.3668	0.179	0.6746	0.1966	0.3446	0.3159	0.5616	0.5326	0.2053	0.702
0.1856	0.4424	0.0679	0.4229	0.0674	0.6198	0.4702	0.53	0.5767	0.5937

References

Albers, W., Kallenberg, W.C.M.: Estimation in Shewhart control charts: effects and corrections. Metrika **59**, 207–234 (2004)

Albers, W., Kallenberg, W.C.M., Nurdiati, S.: Parametric control charts. J. Stat. Plann. Inf. **124**, 159–184 (2004)

Alwan, L.C.: Statistical process analysis. Mcgraw-Hill, New York (2000)

Asquith, W.H.: l-moments and tl-moments of the generalized lambda distribution. Comput. Stat. Data Anal. **51**, 4484–4496 (2007)

Bakir, S.T.: A distribution-free Shewhart quality control chart based on signed-ranks. Qual. Eng. **16**(4), 613–623 (2004)

Bigerelle, M., Najjar, S., Fournier, B., Rupin, N., Iost, A.: Application of lambda distributions and bootstrap analysis to the prediction of fatigue lifetime and confidence intervals. Int. J. Fatigue **28**, 223–236 (2005)

Chakraborti, S.: Run length, average run length and false alarm rate of Shewhart x chart: exact derivations by conditioning. Commun. Stat. Simul. Comput. **29**, 61–81 (2000)

Chang, Y.S., Bai, K.C.: Control charts for positively skewed populations with weighted standard deviations. Qual. Reliab. Eng. Int. **17**, 397–406 (2001)

Dengiz, B.: The generalized lambda distribution in simulation of m/m/1 queue systems. J. Fac. Eng. Arch. Gazi Univ. **3**, 161–171 (1988)

Elaney, H.D., Vargha, A.: The effect on non-normality on student's two-sample t-test. In: The Annual Meeting of the American Educational Research Association, New Orleans (2000)

Farnum, N.R.: Statistical quality control and improvement. Duxbury, Belmont (1994)

Forbes, M.G., Guay, M., Forbes, J.F.: Control design for first-order processes: shaping the probability density of the process state. J. Process Control **14**, 399–410 (2004)

Fournier, B., Minguet, L., Niamien, A., Rupin, N.: Les lambda distributions applique ′es au spc, Unpublished report (2003)

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

Ganeshan, R.: Are more suppliers better?: generalizing the gau and ganeshan procedure. J. Oper. Res. Soc. **52**, 122–123 (2001)

Gunter, B.: The use and abuse of c, part l–4. Quality progress, part 1: vol. 22, no. 1, pp. 72–73, part 2: vol. 22, no. 3, pp. 108–109, part 3: vol. 22, no. 5, pp. 79–80, and part 4: vol. 22, no. 7, pp. 86–87 (1989)

Headrick, T.C., Mugdadi, A.: On simulating multivariate non-normal distributions from the generalized lambda distribution. Comput. Stat. Data Anal. **50**, 3343–3353 (2006)

Joiner, B.L., Rosenblatt, J.R.: Some properties of the range in samples from Tukey's symmetric lambda distributions. J. Am. Stat. Assoc. **66**, 394 (1971)

Karian, Z.A., Dudewicz, E.J.: Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Method. CRC Press, London (2000)

Karian, Z.A., Dudewicz, E.J.: Comparison of gld fitting methods: superiority of percentile fits to moments in l2 norm. J. Iran. Stat. Soc. **2**, 171–187 (2003)

Karvanen, J.: Generation of correlated non-gaussian random variables from independent components. In: Fourth International Symposium on Independent Component Analysis and Blind Signal Separation, Nara, Japan (2003)

Karvanen, J., Nuutinen, A.: Characterizing the generalized lambda distribution by l-moments. Comput. Stat. Data Anal. **52**, 1971–1983 (2008)

Katz, S., Johnson, N.L.: Process Capability Indices. Chapman and Hall, New York (1995)

Kittlitz, R.G.: Transforming the exponential for spc applications. J. Qual. Technol. **31**, 301–308 (1999)

Lee, S.J., Amin, R.W.: Process tolerance limits. Total Qual. Manag. **11**(3), 267–280 (2000)
Lin, Y., Chou, C.: On the design of variable sample size and sampling intervals x charts under non-normality. Int. J. Prod. Econ. **96**(2), 249–261 (2005)
Maravelakis, P.E., Panaretos, J., Psarakis, S.: Effect of estimation of the process parameters on the control limits of the univariate control charts for process dispersion. Commun. Stat.: A. **31**(3), 443–461 (2002)
Michael, B., Khoo, C.: Performance measures for the Shewhart x control chart. Qual. Eng. **16**(4), 585–590 (2004)
Montgomery, D.C.: Introduction to statistical quality control, 4th edn. John Wiley & Sons, New York (2001)
Najjar, D., Bigerelle, M., Lefebvre, C., Iost, A.: A new approach to predict the pit depth extreme value of a localized corrosion process. Isij **43**, 720–725 (2003)
Nedumaran, G., Pignatiello Jr, J.J.: On estimating x control chart limits. J. Qual. Technol. **33**, 206–212 (2001)
Öztürk, A., Dale, R.F.: A study of fitting the generalized lambda distribution to solar radiation data. J. Appl. Meteorol. **21**, 995–1004 (1982)
Pal, S.: Evaluation of non-normal process capability indices using generalized lambda distributions. Qual. Eng. 77–85 (2005)
Peam, W.L., Kotz, S., Johnson, N.L.: Distributional and inferential properties of process capability indices. J. Qual. Technol. **24**(4), 216–231 (1992)
Pearson, R.K.: Exploring process data. J. Process Control **11**, 179–194 (2001)
Pyzdek, T.: Processes control for short and small runs. Qual. Prog. **12**, 51–60 (1993)
Pyzdek, T.: Why normal distributions aren't (all that normal). Qual. Eng. **7**, 769–777 (1995)
Quesenberry, C.P.: On properties of q-charts for variables. J. Qual. Technol. **27**, 184–203 (1995)
Ramberg, J., Schmeiser, B.: An approximate method for generating asymmetric random variables. Commun. ACM **17**(2), 78–82 (1974)
Riaz, M.: A dispersion control chart. Commun. Stat.: Simul. Computat. **37**, 1239–1261 (2008)
Shewhart, W.: Economic control of quality of manufactured products. D. Van Nostrand, New York (1931)
Somerville, S.E., Montgomery, D.C.: Process capability indices and non-normal distributions. Qual. Eng. **9**(2), 305–316 (1996)
Spedding, T.A., Rawlings, P.L.: Non-normality in statistical process control measurements. Int. J. Qual. Reliab. Manag. **11**(6), 27–37 (1994)
Su, S.: Numerical maximum log likelihood estimation for generalized lambda distributions. Comput. Stat. Data Anal. **51**, 3983–3998 (2007)
Tarsitano, A.: Estimation of the generalized lambda distribution parameters for grouped data. Commun. Stat. Theory Methods **34**, 1689–1709 (2005)
Teyarachakul, S., Chand, S., Tang, J.: Estimating the limits for statistical process control charts: a direct method improving upon the bootstrap. Eur. J. Oper. Res. **178**, 472–481 (2007)
Torng, C.-C., Lee, P.-H.: ARL performance of the tukey's control chart. Commun. Stat. Simul. Comput. **37**, 1904–1913 (2008)
Tukey, J.W.: The future of data analysis. Ann. Math. Stat. **33**(1), 1–67 (1962)
Upadhyay, R.R., Ezekoye, O.A.: Treatment of design fire uncertainty using quadrature method of moments. Fire Saf. J. **43**, 127–139 (2008)
Wang, H.: Bounded dynamic stochastic systems—modelling and control. Springer, London (2000)
Woodall, W.H.: Controversies and contradictions in statistical process control. J. Qual. Technol. **32**(4), 341–350 (2000)
Yang, S.-F.: An approach to controlling process variability for short production runs. Total Qual. Manag. **10**(8), 1123–1129 (1999)
Yourstone, S., Zimmer, W.: Non-normality and the design of control charts for averages. Decis. Sci. **23**, 1099–1113 (1992)