Ballard (1987) who suggested to split the skewed distribution into two parts at its mean, instead of its mode, and then compute the standard deviations of the two new distributions using the semivariance approximation of Choobineh and Branting (1986). Finally, we can cite the recent works of Seppala (1995) who suggests to use the "Bootstrap" in the computation of control limits, Willemain and Runger (1996) who proposes to use the notion of "Statistically Equivalent Blocks" to design nonparametric control charts, and Castagliola (1997) who proposes an extension of the Weighted Variance method called the "Scaled Weighted Variance" method.

- Transform the data in order to make them quasi-normal. This approach was chosen by Pyzdek (1992), Farnum (1997) who used the Johnson system of distributions as a general tool for transforming the data to normality.

The method proposed in this paper, which follows the last approach, is devoted only to the designing of "classical" control charts (mean, median, standard deviation, range, EWMA, CUSUM, etc.) for data having a symmetrical distribution with a positive kurtosis (leptokurtic distribution). This method is based on the properties of the symmetrical Johnson S_U distributions which will be examined in the following section.

2. The Symmetrical Johnson S_U Distributions

Let us focus on transformations of form $Z = a + bg(Y)$ of the random variable Y, where a and $b > 0$ are two parameters, where g is a monotone increasing function, and where Z is a $(0, 1)$ normal random variable. It is very easy to show that the random variable Y has the following characteristics:

- cumulative distribution:
$$F_Y(y) = \Phi[a + bg(y)]$$
- inverse cumulative distribution:
$$F_Y^{-1}(\alpha) = g^{-1}\left[\frac{\Phi^{-1}(\alpha) - a}{b}\right]$$
- density function:
$$f_Y(y) = bg'(y)\phi[a + bg(y)]$$
- noncentral moments of order s:
$$m_s(Y) = \int_{-\infty}^{+\infty} \left[g^{-1}\left(\frac{z-a}{b}\right)\right]^s \phi(z)dz \tag{1}$$

If c and $d > 0$ are two additional parameters such that $Y = (X - c)/d$, then we can straightforwardly deduce the characteristics of the random variable X, i.e., $F_X(x) = F_Y[(x - c)/d]$ and $F_X^{-1}(\alpha) = c + dF_Y^{-1}(\alpha)$. There are a large number of possibilities for choosing an adequate function g. Johnson (1949) has proposed a very popular system of distributions based on a set of three different functions:

- $g_L(Y) = \ln(Y)$ and $d = 1$. The distributions defined by this function, called Johnson S_L distributions, are defined on $[c, +\infty]$.
- $g_B(Y) = \ln[Y/(1-Y)]$. The distributions defined by this function, called Johnson S_B distributions, are defined on $[c, c+d]$.
- $g_U(Y) = \ln(Y^2 + \sqrt{Y^2+1}) = \sinh^{-1}(Y)$. The distributions defined by this function, called Johnson S_U distributions, are defined on $]-\infty, +\infty[$.

Johnson has proved in his paper that (a) for every skewness coefficient $\gamma_1 = \mu_3/\mu_2^{3/2}$ and every kurtosis coefficient $\gamma_2 = \mu_4/\mu_2^2 - 3$ such that $\gamma_2 \geq \gamma_1^2 - 2$ there is one and only one Johnson distribution, (b) the S_B and S_U distributions occupy nonoverlapping regions covering the whole of the skewness-kurtosis plane, and the S_L distributions are the transitional distributions separating them (see Fig. 1).

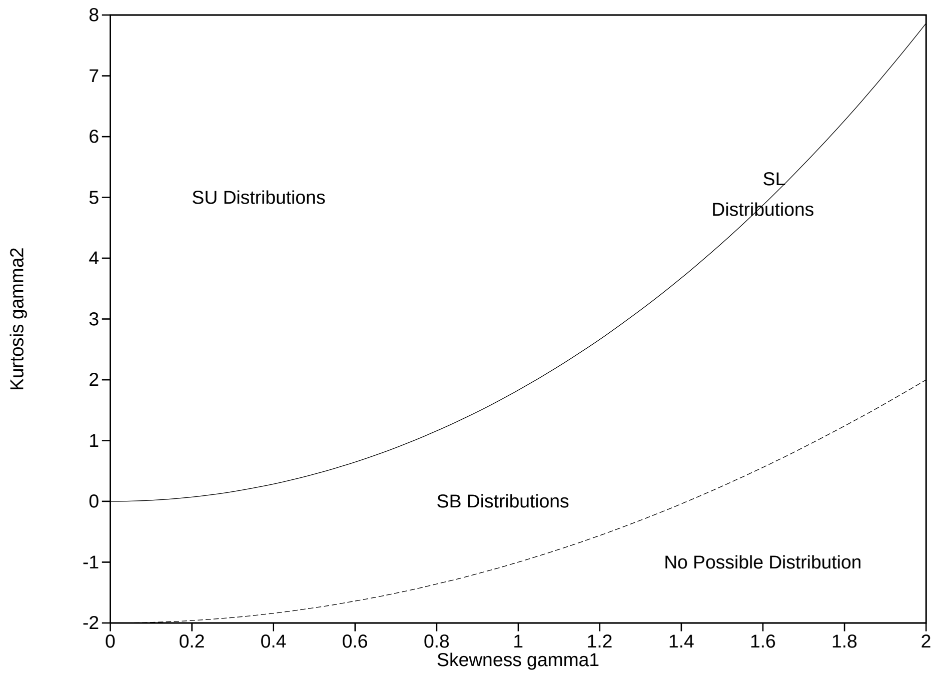

Fig. 1. The (γ_1, γ_2) plane for the Johnson distributions.

If we look at this figure, we can see that among the symmetrical Johnson distributions ($\gamma_1 = 0$) the S_U distributions are leptokurtic while the S_B ones are platykurtic. For this reason, we will now focus more precisely on Johnson S_U distributions which are symmetrical (about the mean $m_1(X) = m_1$). It is clear that a necessary and sufficient condition for a Johnson S_U to be symmetrical is that $a = 0$ and $c = m_1$. Consequently, the characteristics of a symmetrical Johnson S_U random variable X are:

- cumulative distribution:

$$F_X(x) = \Phi\left[b\sinh^{-1}\left(\frac{x-m_1}{d}\right)\right]$$

- inverse cumulative distribution:

$$F_X^{-1}(\alpha) = m_1 + d\sinh\left[\frac{\Phi^{-1}(\alpha)}{b}\right] \quad (2)$$

- density function:

$$f_X(x) = \frac{b}{\sqrt{x^2+d^2}}\phi[b\sinh^{-1}(x/d)]$$

Let $\mu_2(X) = \mu_2$ and $\gamma_2(X) = \gamma_2$ be, respectively, the variance and kurtosis coefficients of the random variable X. If X is a symmetrical Johnson S_U distribution, then we proved in Castagliola (1998) that parameters b and d are related to μ_2 and γ_2 using the following equations (see the appendix for the proof):

$$b = \sqrt{\frac{2}{\ln(\sqrt{2(\gamma_2+2)}-1)}} \quad (3)$$

$$d = \sqrt{\frac{2\mu_2}{\sqrt{2(\gamma_2+2)}-2}} \quad (4)$$

3. Application to Control Charts

Let $X_1, \ldots, X_n$ be a sample of n independent random variables corresponding to training data taken when the process is considered to be "in control", from which we have to compute control limits. Let $\hat{m}_1$, $\hat{\mu}_2$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$ be, respectively, the (moment) estimators of the mean, variance, skewness, and kurtosis. We will assume now (see the appendix) that a first statistical test leads to the conclusion $\gamma_1 = 0$ (the data distribution seems symmetrical) and a second one leads to the conclusion $\gamma_2 \geq 0$ (the data distribution seems leptokurtic). If these conditions are verified, we suggest to compute control limits as presented below:

- Compute $\hat{b}$ and $\hat{d}$ using Eqs. (3) and (4) in which μ_2 and γ_2 have been replaced by their estimators.
- Transform each new observation X to a quasi-normal $N(0, 1)$ random variable Z using the following equation:

$$Z = b\sinh^{-1}\left(\frac{X-m_1}{d}\right) \quad (5)$$

- Use "classical" control limits (mean, median, standard deviation, range, EWMA, CUSUM, etc.) corresponding to a normal $N(0, 1)$ distribution.