

# Process Capability Indices for Non-Normal Data

MARTIN KOVÁŘÍK, LIBOR SARGA  
 Department of Statistics and Quantitative Methods  
 Tomas Bata University in Zlín  
 Mostní 5139, 760 01 Zlín  
 CZECH REPUBLIC  
[m1kovarik@fame.utb.cz](mailto:m1kovarik@fame.utb.cz), [sarga@fame.utb.cz](mailto:sarga@fame.utb.cz)

*Abstract:* - When probability distribution of a process characteristic is non-normal,  $C_p$  and  $C_{pk}$  indices calculated using conventional methods often lead to erroneous interpretation of process capability. Various methods have been proposed for surrogate process capability indices (PCIs) under non-normality but few literature sources offer their comprehensive evaluation and comparison, in particular whether they adequately capture process capability under mild and severe departures from normality, and what is the best method to compute true capability under each of these circumstances. We overview 9 methods as to their performance when handling PCI non-normality. The comparison is carried out by simulating log-normal and data and the results presented using box plots. We show performance to be dependent on the ability to capture tail behavior of the underlying distribution.

*Key-Words:* - process, capability, index, distribution, method, transformation, asymmetric

## 1 Introduction

Process capability indices (PCIs) are widely used to determine whether a given process is capable to output products in a given tolerance band.  $C_p$  and  $C_{pk}$ , ranking among the most popular, are defined as:

$$C_p = \frac{\text{tolerance width}}{\text{process width}}, \quad (1)$$

$$C_{pk} = \min\{C_{pu}, C_{pl}\} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\}, \quad (2)$$

where  $USL$  and  $LSL$  denote upper and lower specification limits, respectively, and  $C_{pk}$  the minimum of  $(C_{pu}, C_{pl})$ . Due to unknown mean process value  $\mu$  and variance  $\sigma^2$ , they are usually estimated from sample counterparts,  $\bar{x}$  and  $s^2$ . English and Taylor analyzed non-normality in PCIs and posited  $C_{pk}$  is more sensitive to deviations than  $C_p$ . Kotz and Johnson tested properties of the indices, e.g., Clements' method, Johnson-Kotz-Pearn method and "distribution-independent" PCIs under non-normal distributions as well. Index by Wright [19] integrates skewness correction factor in denominator of  $C_{pmk}$  proposed by Johnson. Choi and Bai [5] suggested a weighted variance heuristic method to correct PCIs for skewness by considering standard deviation above and beyond mean individually. Some authors proposed new-generation PCIs based on assumptions about the population distribution. Pearn and Kotz [16] tied

their capability indices to Pearson distribution. Johnson et al. [11] substituted  $6\sigma$  in the numerator of equation (2) by  $6\theta$  where  $\theta$  was selected so that „capability“ was not substantially influenced by the shape of the distribution. Castagliola presented computational methods for non-normal PCIs by estimating ratio of unsatisfactory products using Burr distribution. Vännman introduced  $C_p(u, v)$  index family with  $(u, v)$  comprising many other indices as special cases. Deleryd investigated suitable  $u$ 's and  $v$ 's for skewed process distributions.  $C_p(1, 1)$ , equivalent to  $C_{pmk}$ , is recommended as the best fit for treating non-normality in PCIs. Even though it is accepted  $C_p$  and  $C_{pk}$  are not suitable for process capability with non-normal distributions and the previously-mentioned methods constitute viable substitutes, a complex treatment is missing. Experts are interested in which methods are mutually comparable, which are sensitive to the normality assumption, and which are suitable for use with non-normal distributions. We will investigate 9 methods with respect to process capability by comparing  $C_{pu}$  obtained from simulations with a target  $C_{pu}$ . In case the data are non-normally distributed, it will be necessary to obtain estimates as  $\hat{L}_{0.00135}$ ,  $\hat{U}_{0.99865}$ , and  $\hat{M}e_{0.5}$  quantiles; we will evaluate some methods.

Sometimes, the shape of the distribution is known a priori or based on knowledge of physical process properties. It is then necessary to use goodness of fit test to determine whether the model correctly represents the data, estimates distribution parameters, and computes the quantiles for log-normal, Weibull, exponential distributions, etc. Some literature sources list equations for parameter and quantile estimations which can be also inputted to software. Fig. 1 depicts how a PCI is obtained for a non-normally distributed quality attribute [16,17,18].

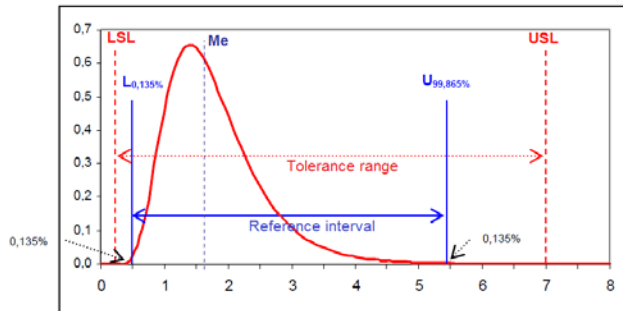


Fig. 1 PCI for a non-normally distributed quality attribute [8]

## 2 Surrogate PCIs for Non-Normal Distributions

Here we will present methods to compute PCIs for non-normal data distributions.

### 2.1 Probability Graph

A popular approach to compute PCIs is to use normal distribution graph which allows to investigate the normality assumption. As with normal distribution where process width spans the 0.135 and 99.865 percentile range, a surrogate PCI can be obtained from suitable probability distribution:

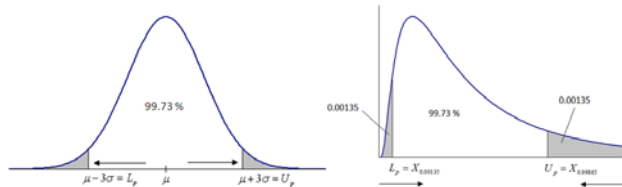


Fig. 2 Normal distribution: interval spanning 99.73% of values, and non-normal distribution: interval spanning 99.73% of values [Own work]

$$C_p = \frac{USL - LSL}{Upper\ 0.135\% - lower\ 0.135\% \ centiles} = \frac{USL - LSL}{U_p - L_p}, \quad (3)$$

where  $L_p$  and  $U_p$  are 0.135 and 99.865 percentiles, respectively, trivially obtained in statistical software such as Minitab, NCSS, XLStatistics, etc. Due to median being preferred as asymmetric distributions' location parameter,  $C_{pu}$  and  $C_{pl}$  are defined as

$$C_{pu} = \frac{USL - median}{x_{0.99865} - median}, \quad (4)$$

$$C_{pl} = \frac{median - LSL}{median - x_{0.00135}}, \quad (5)$$

Because the method is graphical, it can produce errors if statistical software is eschewed in favor of a probability paper where  $x$  axis is linear and  $y$  probabilistic, conforming to a selected probability distribution.  $N$  observations are diagramed according to  $(x(i))$  with the respective counterparts of the empirical distribution function  $((i) / N)$ . Interpolated line is the best fit for a function of quality attribute's expected probability distribution. Quantiles corresponding to (0,00135; 0,50; 0,99865) can be inferred and instated to the performance indicators. However, we repeat the method can be erroneous if statistical software is not involved [19,20].

### 2.2 Distribution-Independent Tolerance Intervals

Chan et al. [4] devised a method utilizing distribution-independent tolerance intervals to compute  $C_p$  a  $C_{pk}$  for non-normally distributed processes:

$$C_p = \frac{USL - LSL}{6\sigma} = \frac{USL - LSL}{\frac{3}{2}(4\sigma)} = \frac{USL - LSL}{3(2\sigma)}. \quad (6)$$

This implies

$$C_p = \frac{USL - LSL}{\omega} = \frac{USL - LSL}{\frac{3}{2}\omega_2} = \frac{USL - LSL}{3\omega_3}, \quad (7)$$

$$C_{pk} = \frac{\min[(USL - \mu, \mu - LSL)]}{\omega/2}, \quad (8)$$

where  $\omega$  is tolerance interval's width. If 99% reliability is required, the interval ensures 75% coverage of population statistics, the method is thus the only one giving different result even if the underlying distribution is normal. In that case, 99.73% of the quality attribute's values lie within  $\pm 3s$  tolerance bounds,  $C_p = 1$  means 99.72% of values lie within the bound. In equation (6), 6 in denominator is generally dependent on sample size

out of which  $s$  is estimated, and on quality attribute's distribution. The non-normality limitation can be mitigated by selecting suitable percentage of items within the tolerance bounds, e.g., for 99%, 5.15 may be used, a constant valid for many probability distributions with skewness from 0 to 3.111 and kurtosis from 1 to 5.997 [5,10].

**2.3 Weighted Variance Method**

Choi and Bai [5] proposed a heuristic weighted variance method for PCI correction according to population skewness. Let  $P_x$  be probability a variable  $X$  is lower or equal to the population mean

$$P_x = \frac{1}{n} \sum_{i=1}^n I(\bar{X} - X_i), \tag{9}$$

where  $I(x) = 1$  when  $x > 0$ , and  $I(x) = 0$  when  $x < 0$ . PCI is then defined as [5]

$$C_p = \frac{USL - LSL}{6\sigma W_x}, \tag{10}$$

where  $W_x = \sqrt{1 + |1 - 2P_x|}$ . Therefore

$$C_{pu} = \frac{USL - \mu}{2\sqrt{2P_x}\sigma}, \tag{11}$$

$$C_{pl} = \frac{\mu - LSL}{3\sqrt{2(1 - P_x)}\sigma}. \tag{12}$$

**2.4 Clements' Method**

Clements [6] supplanted  $6\sigma$  in equation (6) with  $U_p - L_p$ :

$$C_p = \frac{USL - LSL}{U_p - L_p}, \tag{13}$$

where  $U_p$  is the 99.865 percentile. For  $C_{pk}$ ,  $\mu$  is estimated by median  $M$  and  $3\sigma$  by  $U_p - M$  and  $M - U_p$ . We get

$$C_{pk} = \min \left[ \frac{USL - M}{U_p - M}, \frac{M - LSL}{M - L_p} \right]. \tag{14}$$

Clements' method employs traditional estimates for skewness and kurtosis based on the third and the fourth central moments both of which are unreliable for small sample sizes. The PCIs were devised for a class of *Pearson distributions*, i.e., *beta*, *gamma*, and *Student* [6,16].

Procedure for calculating PCIs:

a) tolerance bounds  $USL, LSL$ ,

- b)  $\bar{x}, s, Sk$  (skewness),  $Ku$  (kurtosis),
- c) standardized quantiles  $L'_p, U'_p$  (for selected  $p$  and known  $Sk, Ku$  can be obtained from tables [6] or in software),
- d) standardized median  $M'$  from tables [6],
- e)  $L_p, U_p : L_p = \bar{x} - s \cdot L'_p, U_p = \bar{x} + s \cdot U'_p$ ,
- f)  $M: M = \bar{x} + s \cdot M'$
- g)  $C_p, C_{pk}$  from equations (13) and (14).

Pearn and Kotz [16] expanded the method with  $C_{pm}, C_{pm*}, C_{pmk}$  defined as:

$$C'_{pm} = \frac{USL - LSL}{6\sqrt{\left(\frac{U_p - L_p}{6}\right)^2 + (M - T)^2}}, \tag{15}$$

$$C'_{pm*} = \frac{\min(USL - T, T - LSL)}{3\sqrt{\left(\frac{U_p + L_p}{6}\right)^2 + (M - T)^2}}, \tag{16}$$

$$C'_{pmk} = \min \left[ \frac{\frac{USL - M}{3\sqrt{\left(\frac{U_p - M}{3}\right)^2 + (M - T)^2}}}{\frac{M - LSL}{3\sqrt{\left(\frac{M - L_p}{3}\right)^2 + (M - T)^2}}} \right]. \tag{17}$$

**2.5 Burr Percentile Method**

Burr [2,3] proposed Burr XII distribution to compute the random variable  $X$ 's percentiles. Its probability density function,  $y$ , is of the form

$$f(y|c,k) = \begin{cases} \frac{cky^{c-1}}{(1+y^c)^{k+1}} & \text{when } y \geq 0, c \leq 1, k \geq 1, \\ 0 & \text{when } y < 0, \end{cases} \tag{18}$$

where  $c$  and  $k$  represent skewness and kurtosis, respectively.

Liu and Chen introduced modified Clements' method with percentiles from Burr XII distribution instead of from the Pearson's curve.

1. Estimate sample mean  $\bar{x}$  and standard deviation  $s$ . Skewness  $s_3$  and kurtosis  $s_4$  from empirical sample size are calculated as:

$$s_3 = \frac{n}{(n-1)(n-2)} \sum \left( \frac{x_j - \bar{x}}{s} \right)^3, \tag{19}$$

and

$$s_4 = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left( \frac{x_j - \bar{x}}{s} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}. \quad (20)$$

2. Compute standardized skewness and kurtosis moments  $\alpha_3$  and  $\alpha_4$ , respectively:

$$\alpha_3 = \frac{(n-2)}{\sqrt{n(n-1)}} s_3, \quad (21)$$

and

$$\alpha_4 = \frac{(n-2)(n-3)}{(n^2-1)} s_4 + 3 \frac{(n-1)}{(n+1)}. \quad (22)$$

Standardized kurtosis is known as excess; if it equals 3, the distribution is normal (i.e., mesokurtic). Negative values indicate platykurtic (broader) distribution with higher standard deviation, positive values leptokurtic (slender) distribution with lower standard deviation [1,2,3].

3. Use  $\alpha_3$  and  $\alpha_4$  to select suitable  $c$  and  $k$ , then calculate Burr XII distribution standardized value

$$Z = \frac{Y - \mu}{\sigma}, \quad (23)$$

where  $Y$  is the selected Burr value,  $\mu$  and  $\sigma$  mean and standard deviation, respectively. All central moment characteristics are in tables [2,3] which also supply standardized 0.00135, 0.5, and 0.99865 percentiles corresponding to  $Z_{0.00135}$ ,  $Z_{0.5}$ , and  $Z_{0.99865}$ . We get them by comparing

$$\frac{X - \bar{x}}{s} = \frac{Y - \mu}{\sigma}.$$

4. We now have estimates for lower, medium, and upper percentiles:

$$L_p = \bar{x} + s Z_{0.00135},$$

$$M = \bar{x} + s Z_{0.5},$$

$$U_p = \bar{x} + s Z_{0.99865}.$$

5. Compute the PCI from equations for the Clements' method.

Other procedures can be used to estimate Burr distribution's parameters apart from the third and the fourth central moments, e.g., maximum likelihood method, probability weighted moments, etc. Moment characteristics are the most usual due to their simplicity and speed of computation [2,3].

## 2.6 Transformation Techniques

**1. Variance stabilization** requires a transformation  $y = g(x)$  with constant variance  $\sigma^2(y)$ . In case variance of  $x$  is a function of the type  $\sigma^2(x) = f_1(x)$ ,  $\sigma^2(y)$  is computed as

$$\sigma^2(y) \approx \left[ \frac{dg(x)}{dx} \right]^2, f_1(x) = C, \quad (24)$$

where  $C$  is a constant;  $g(x)$  is then obtained by solving a differential equation

$$g(x) \approx C \int \frac{dx}{\sqrt{f_1(x)}}. \quad (25)$$

Multiple methods and tools were devised for constant relative error  $\delta(x)$ , e.g.,  $\sigma^2(x)$  is given by  $f_1(x) = \delta^2(x)$ ,  $x^2 = \text{constant}$ . After instating, we get  $g(x) = \ln x$ . Here, logarithmic transformation is optimal together with geometric mean. When  $\sigma^2(x) = f_1(x)$  exhibits a power relation,  $g(x)$  will be of power type as well. For normal distribution, mean is independent on variance, i.e., variance-stabilizing transformation will ensure convergence to normality.

**2. Sample distribution symmetrization** is possible using simple power transformation

$$y = g(x) = \begin{cases} x^\lambda & \text{for } \lambda > 0 \\ \ln x & \text{for } \lambda = 0 \\ -x^{-\lambda} & \text{for } \lambda < 0. \end{cases} \quad (26)$$

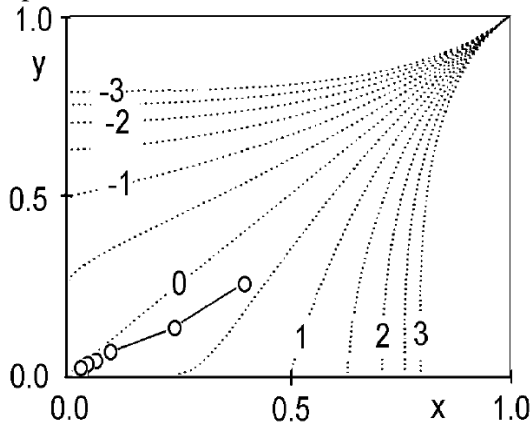
Scale is not preserved, is discontinuous due to  $\lambda$ , and suitable only for positive values. Optimal  $\hat{\lambda}$  estimate is found by minimizing the suitable asymmetry characteristics. Skewness defined in the following equation is robust and can be used in place of  $\hat{g}_1(y)$ :

$$\hat{g}_{1R}(y) = \frac{(\bar{y}_{0.75} - \bar{y}_{0.50}) - (\bar{y}_{0.50} - \bar{y}_{0.25})}{\bar{y}_{0.75} - \bar{y}_{0.25}}. \quad (27)$$

For a symmetric distribution,  $\hat{g}_1(y) = 0$ , as well as  $\hat{g}_1(y)$  and  $\hat{g}_{1R}(y)$ . Optimal  $\hat{\lambda}$  is found from a rankit plot; its quantiles will lie approximately on a line.

**3. Hine-Hines selection graph** ( $x$  axis:  $\tilde{x}_{0.5} / x_{1-P_i}$ ,  
 $y$  axis:  $\tilde{x}_{P_i} / \tilde{x}_{0.5}$ )

A selection graph as a diagnostic tool for finding optimal  $\lambda$  value.



**Fig. 3** Selection graph for a sample from log-normal distribution [14]

The graph originates from asymmetry requirements of quantiles around median

$$\left(\frac{\tilde{x}_{P_i}}{\tilde{x}_{0.5}}\right) + \left(\frac{\tilde{x}_{0.5}}{\tilde{x}_{1-P_i}}\right)^{-\lambda} = 2, \tag{28}$$

where for each ordinal probability, values of  $P_i = 2^{-i}$ ,  $i = 2, 3$  are usually assigned. To compare experimental and theoretical observation for selected  $\lambda$ ,  $y^\lambda + x^{-\lambda} = 2$  for  $0 \leq x \leq 1$  and  $0 \leq y \leq 1$  is plotted into the graph:

- a) for  $\lambda = 0$ , the solution is a line  $y = x$ ,
- b) for  $\lambda < 0$ , the solution is a ratio  $y = (2 - x^{-\lambda})^{1/\lambda}$ ,
- c) for  $\lambda > 0$ , the solution is a ratio  $x = (2 - y^\lambda)^{-1/\lambda}$ .

Judging from the positions of experimental points on theoretical curves in the selection graph,  $\lambda$  can be inferred and quality of transformation assessed in varying distances from median. To converge the sample distribution to normal, Box-Cox transformation is a viable method due to skewness and kurtosis [14].

**2.6.1 Box-Cox Transformation**

Disadvantages of a simple power transformation (discontinuity around zero and non-aligned transformed scales) are mitigated by Box-Cox transformation  $X^{(\lambda)}$ , a linear variation of a simple power transformation  $X_p^{(\lambda)}$ . Box-Cox class of polynomial transformations is of the form

$$X^{(\lambda)} = \begin{cases} \frac{X^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \ln X & \text{for } \lambda = 0. \end{cases} \tag{29}$$

To determine its quality and optimal  $\lambda$ , quantile dispersion graphs (QDG) and quantile graphs (Q-Q) can be plotted, with normality tests after power transformation more suitable. A well-known Shapiro-Wilk test is equivalent to testing tangent's significance in a Q-Q plot, i.e., linearity can be analyzed as well. The Box-Cox family of transformations depends on  $\lambda$  estimated by Maximum Likelihood Estimate (MLE) or Least-Squares Estimate (LSE). For  $\lambda = 1$ , additive measurement model is suitable while for  $\lambda = 0$ , a multiplicative one should be preferred. First,  $\lambda$  is chosen from a selected range, and the following statistic computed

$$L_{\max} = -\frac{1}{1} \ln \hat{\sigma}^2 + \ln J(\lambda, X) = \tag{30}$$

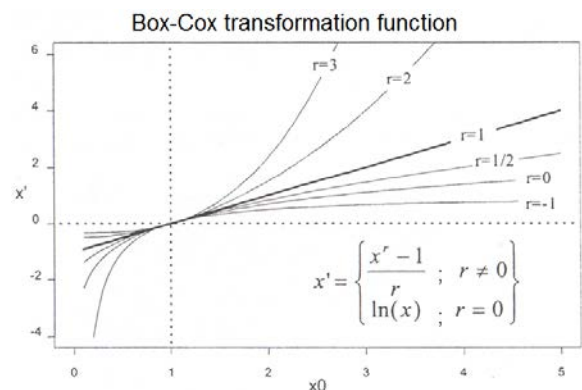
$$= \frac{1}{2} \ln \hat{\sigma}^2 + (\lambda - 1) \sum_{i=1}^n \ln X_i,$$

where

$$J(\lambda, X) = \prod_{i=1}^n \frac{\partial W_i}{\partial X_i} = \prod_{i=1}^n X_i^{\lambda-1} \forall \lambda,$$

so that  $\ln J(\lambda, X) = (\lambda - 1) \sum_{i=1}^n \ln X_i$ . The  $\hat{\sigma}^2$  estimate for fixed  $\lambda$  is  $\hat{\sigma}^2 = S(\lambda)/n$  where  $S(\lambda)$  is a residual sum of squares in analysis of variance  $X(\lambda)$ . After computing  $L_{\max}(\lambda)$  for several  $\lambda$  in the range,  $L_{\max}(\lambda)$  is plotted against  $\lambda$ . MLE for  $\lambda$  is computed from the  $\lambda$  maximizing  $L_{\max}(\lambda)$ .

With optimal  $\lambda^*$ , every  $X$  in the specification bounds is transformed to a normal value with the help of equation (29) [4,18].



**Fig. 4** Shape of Box-Cox transformation for  $r = -1, 0, 0.5, 1, 2, 3$  [13]

Box-Cox transformation estimates  $\lambda$  by minimizing standard deviation of the normalized transformed variable; software is capable to evaluate many of them.  $y$  is a transformed value of the  $x$  quality attribute [18].

$\lambda$  denotes the function's shape and allows to find  $F(x)$  satisfying the maximum normality or symmetry condition. Different shapes of the Box-Cox transformation function (specified as  $r$ ) for various  $\lambda$  are depicted in Fig. 4. The aim is to find  $r$  ensuring maximum symmetry "measured" by skewness, or (better) maximum normality "measured" by likelihood, after transformation. Maximum symmetry (zero skewness condition) or maximum likelihood condition is computed iteratively [15].

**2.6.2 Backward Transformation After Box-Cox and Power Transformation**

If a suitable transformation leading to approximate normality is found,  $\bar{y}, s^2(y)$ , and confidence intervals  $\bar{y} \pm t_{1-\alpha/2}(n-1)s(y)/\sqrt{n}$  can be found and statistical tests performed. However, we first need to transform the statistics back to their original forms.

1. **Incorrect approach** simply backward transforms  $\bar{x}_R = g^{-1}(\bar{y})$ . For a simple power case, backward transformation leads to mean of the form

$$\bar{x}_R = \bar{x}_\lambda = \left[ \frac{\sum_{i=1}^n x_i^\lambda}{n} \right]^{1/\lambda} \tag{31}$$

For  $\lambda = 0$ ,  $\ln x$  is used instead of  $x^\lambda$  and  $e^x$  instead of  $x^{1/\lambda}$ .  $\bar{x}_R = \bar{x}_{-1}$  represents harmonic,  $\bar{x}_R = \bar{x}_0$  geometric,  $\bar{x}_R = \bar{x}_1$  arithmetic, and  $\bar{x}_R = \bar{x}_2$  quadratic mean, respectively. The approach leads to biased results.

2. **Correct (more exact) approach** is based on Taylor expansion  $y = g(x)$  around  $\bar{y}$ . For untransformed  $\bar{x}_R$ , an approximate relation is

$$\bar{x}_R \approx g^{-1} \left[ \bar{y} - \frac{1}{2} \frac{d^2 g(x)}{dx^2} \left( \frac{dg(x)}{dx} \right)^{-2} \cdot s^2(y) \right] \tag{32}$$

for mean and

$$s^2(x_R) \approx \left( \frac{dg(x)}{dx} \right)^{-2} \cdot s^2(y) \tag{33}$$

for variance. Each derivation is computed at  $x = \bar{x}_R$ .

For  $100(1-\alpha)\%$  confidence interval for mean of original data, it holds that  $x_R - I_L \leq \mu \leq x_R + I_U$  where

$$I_L = g^{-1} \left( \bar{y} + G - t_{1-\alpha/2}(n-1) \frac{s(y)}{\sqrt{n}} \right), \tag{34}$$

$$I_U = g^{-1} \left( \bar{y} + G + t_{1-\alpha/2}(n-1) \frac{s(y)}{\sqrt{n}} \right), \tag{35}$$

$$G = -0.5 \frac{d^2 g(x)}{dx^2} \left( \frac{dg(x)}{dx} \right)^{-2} \cdot s^2(y). \tag{36}$$

$t_{1-\alpha/2}(n-1)$  denotes  $100(1-\alpha/2)\%$  Student quantile with  $(n-1)$  degrees of freedom. With  $y = g(x)$  and  $\bar{y}, s^2(y)$ , it is trivial to compute  $\bar{x}_R$  and  $s^2(x_R)$ :

a) For a special case of  $\lambda = 0$ , i.e., logarithmic transformation of the type  $g(x) = \ln x$ :

$$\bar{x}_R \approx \exp[\bar{y} + 0.5s_2(y)], \quad s^2(x_R) \approx \bar{x}_R^2 s^2(y).$$

b) For  $\lambda \neq 0$ ,  $\bar{x}_R$  is the following quadratic equation's root:

$$\bar{x}_{R,1,2} = \left[ \frac{0.5(1 + \lambda\bar{y}) \pm \sqrt{0.25(1 + \lambda\bar{y})^2 + 2\lambda(\bar{y} + s^2(y)) + (\lambda^2 - 2s^2(y))}}{\lambda} \right]^{1/\lambda} \tag{37}$$

$\bar{x}_{R,i}$  closer to  $\bar{x}_{0.5} = g^{-1}(\bar{y}_{0.5})$  is taken as an estimate of  $\bar{x}_R$ . With known untransformed mean, variance can be computed as  $s^2(x) = \bar{x}_R^{-2\lambda+2} \cdot s^2(y)$  [14].

**2.6.3 Johnson Transformation**

When observations are non-normally distributed, Johnson transformation can be used; the data will then be  $N(0,1)$ -normally distributed. It's of the form

$$z = \gamma + \eta\tau(x; \varepsilon, \lambda), \tag{38}$$

$$\eta > 0, -\infty < \gamma < \infty, \lambda > 0, -\infty < \varepsilon < \infty,$$

where  $z$  is a standardized normal random variable and  $x$  a variable fitted using Johnson transformation. Four parameters to be estimated are  $\gamma, \eta, \varepsilon$ , and  $\lambda$  with  $\tau$  an arbitrary function taking one of the following three forms. Johnson transformation therefore selects one type depending on whether the random variable is "log-normal", "bounded", or "unbounded" [9,11,12].











- [6] J. A. Clements, Process capability indices for non-normal calculations, *Qual. Prog.* 22 (1989) 49-55.
- [7] M. Dlouhý, J. Fábry, M. Kuncová, T. Hladík, *Simulace podnikových procesů*, Brno, Computer Press, 2007.
- [8] F. Fabian, V. Horálek, J. Křepela, J. Michálek, V. Chmelík, J. Chodounský, J. Král, *Statistické metody řízení jakosti*, Praha, ČSJ, 2007.
- [9] I. D. Hill, R. Hill, R. L. Holder, Algorithm AS99: Fitting Johnson curves by moments, *J. R. Stat. Soc.* 25:2 (1976) 180-189.
- [10] K. Ishikawa, *What Is Total Quality Control? The Japanese Way*, New York, Prentice Hall, 1988.
- [11] N. L. Johnson, S. Kotz, W. L. Pearn, Flexible process capability indices, *Pak. J. Stat.* 10 (1992) 23-31.
- [12] V. E. Kane, Process capability indices, *J. Qual. Technol.* 18 (1986) 41-52.
- [13] K. Kupka, *Statistické řízení jakosti*, Pardubice, TriloByte, 2001.
- [14] M. Meloun, J. Militký, *Statistická analýza experimentálních dat*, second ed., Praha, Academia, 2004.
- [15] M. Meloun, J. Militký, *Kompéndium statistického zpracování dat*, second ed., Praha, Academia, 2006.
- [16] W. Pearn, S. Kotz, Application of Clements' method for calculating second- and third-generation process capability indices for non-normal Pearsonian populations, *Qual. Eng.* 7:1 (1994) 139-145.
- [17] L. A. R. Rivera, N. F. Hubele, F. D. Lawrence, Cpk index estimation using data transformation, *Comput. Ind. Eng.* 29:1-4 (1995) 55-58.
- [18] L. C. Tang, S. E. Than, B. W. Ang, A graphical approach to obtaining confidence limits of Cpk, *Qual. Reliab. Eng. Int.* 13:6 (1997) 337-346.
- [19] P. A. Wright, A process capability index sensitive to skewness, *J. Stat. Comput. Sim.* 52:3 (1995) 195-203.
- [20] K. Yang, B. El-Haik, *Design For Six Sigma: A Roadmap for Product Development*, New York, McGraw-Hill, 2003.