Research Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.1967

Published online 19 February 2016 in Wiley Online Library

# Process Capability Analysis via Continuous Ranked Probability Score

**Liangxing Shi,[a*†] Hongye Ma[a] and Dennis K. J. Lin[b]**

**Process capability analysis is an important aspect of quality control. Various process capability indices are proposed when the distribution is normal. For non-normal cases, percentile and yield-based indices have been introduced. These two methods use partial features of a process distribution, such as key percentiles and the proportion of non-conforming (PNC) to estimate the process capability. However, these local features may not reflect the uniformity of a process appropriately when the distribution is non-normal. In this paper, continuous ranked probability score (CRPS) is introduced to process capability analysis and a CRPS-based approach is proposed. This method can assess the dispersion of process variation across the overall distribution and is applicable to any continuous distribution. An example and simulations show that CRPS-based indices are more stable and accurate indicators of process capability than the existing indices in reflecting the degree of process fluctuation.** 

**Keywords:** process capability index; non-normal distribution; quality control; continuous improvement

## 1. Introduction

Process capability analysis plays an important role in quality control and continuous improvement. To manufacture high quality products with 1ow cost, it is essential to ensure the process capability. Process capability analysis is widely used in factories to guarantee the uniformity of the process. It is also an important factor considered by manufacturers to choose suppliers.

Process capability indices (PCIs) are used to reflect process capability. Indices such as $C_p$ and $C_{pk}$ [1] are widely applied. $C_p$ measures the potential process capability, and $C_{pk}$ also takes into account the location parameter of the process distribution. However, the indices $C_p$ and $C_{pk}$ are based on the assumption of normality. These indices may not be appropriate when the process distribution is non-normal.

Various PCIs for non-normal distribution have been proposed. However, the methods developed to date have a very limited focus on the extremes of the distribution and are not sensitive to major variations in the bulk of the distribution. A percentile-based method proposed by Clements[2] calculates an index of process variation based on the 0.135, 50 and 99.865 percentiles. This method is widely applied in process capability analysis when the distribution is non-normal. Pearn *et al.*[3] studied the second and third generation PCIs. Chen *et al.*,[4] Tong *et al.*[5] and Peng[6] gave a unified form of PCIs named $C_p(u,v)$. Yang *et al.*[7] modified the natural tolerance and proposed indices based on the highest density interval. These methods are proposed based on the percentile-based method and also use percentiles to reflect the process capability, so we mainly study the percentile-based method. While percentile-based indices are not sensitive to the changes of the points on the process distribution except the 0.135, 50 and 99.865 percentiles, and therefore give an identical process capability evaluation to production lines with very different probability of producing ideal or extremely deviant part sizes. Therefore this method does not accurately evaluate the potential capability of non-normal distributions. Yield-based indices have also been studied by many researchers (see Castagliola[8], Chen[9], Chao *et al.*,[10] Abbasi *et al.*[11]). But this approach focuses exclusively on the proportion of products that do not conform to the defined specification limits, and is insensitive to variations in the distribution shape within the specification interval.

Methods such as data transformation are also suggested (see Box *et al.*,[12] Somerville *et al.*[13]). However, it is often difficult or impossible to find a proper transformation function,[14] and the relationship between the transformed data and the original specification can be unclear.[15,16]

In this paper, PCIs based on continuous ranked probability score (CRPS) are proposed. CRPS is widely used in probabilistic forecasting field (see Hersbach,[17] Gneiting *et al.*[18]). CRPS is an effective scoring rule as Nau[19] proved that the outcomes given by CRPS are meaningfully ordered. When using CRPS to evaluate the capability, only a cumulative distribution function (CDF) needs to be

[a]*Department of Industrial Engineering, Tianjin University, Tianjin, China, 300072*
[b]*Department of Statistics, The Pennsylvania State University, State College, PA 16801, USA*
**Correspondence to: Liangxing Shi, Department of Industrial Engineering, Tianjin University, Tianjin, China, 300072.*
[†]*E-mail: shi@tju.edu.cn*



 

known, and CRPS can be applied to any data distribution. This scoring rule considers all the points on the distribution, so it gives a more complete indication of process quality. CRPS can measure the distance from the distribution to its center, so it can be used to reflect process variation. Based on this, CRPS-based indices can effectively evaluate the process capability.

The paper is organized as follows: commonly used PCIs under non-normal distributions are illustrated and discussed in Section 2. Then CRPS-based PCIs are proposed in Section 3. In Sections 4 and 5, simulations and an example are given to illustrate the effectiveness of the CRPS-based method. The discussion and conclusions are made in Section 6 and 7.

## 2. Methods for assessing process capability

### 2.1. Case of normal distribution

Process capability is used to reflect whether the process fluctuation meets customers' requirements. It can be obtained through comparing the specification spread with the actual process spread. If the process spread is within the specification interval, the process capability is sufficient. Two process capability evaluation indices that are commonly used are $C_p$ and $C_{pk}$. The $C_p$ is defined as the ratio of the product tolerance and the process range, as

$$C_p = \frac{USL - LSL}{6\sigma_p} \tag{1}$$

*where* $USL$ and $LSL$ are upper and lower specification limits, respectively. $\sigma_p$ is standard deviation of the process distribution. $C_{pk}$ is used to judge whether the process mean value coincides with the specification center. It can be simply defined as

$$C_{pk} = \min\{C_{pu}, C_{pl}\} \tag{2}$$

*where* $C_{pu} = \frac{USL-\mu}{3\sigma_p}$ *and* $C_{pl} = \frac{\mu-LSL}{3\sigma_p}$. *If* $C_{pk} < C_p$, it indicates that the process is biased toward either the $USL$ or $LSL$.

This method can effectively measure the process capability when the process is normally distributed. However, when the process distribution is non-normal, these indices may be misleading.

### 2.2. Case of non-normal distribution

With respect to the process capability of non-normal distribution, various indices have been proposed. Clements[2] proposed percentile-based indices to evaluate the process capability under non-normal distribution, defined as

$$C_{p(p)} = \frac{USL - LSL}{P_{99.865} - P_{0.135}} \tag{3}$$

$$C_{pk(p)} = \min\{C_{pu(p)}, C_{pl(p)}\} = \min\left\{\frac{USL - m}{P_{99.865} - m}, \frac{m - LSL}{m - P_{0.135}}\right\} \tag{4}$$

*where* $P_\alpha$ stands for the α percentile of the fitting distribution and $m$ is the median value of the process. Wu *et al.*[20] indicated that this method may be inaccurate when the distribution is extremely skewed.

Castagliola[8] and Chen,[9] for example, studied the relationship between PCIs and proportion of non-conforming (PNC). Yield-based indices are also proposed to evaluate the capability of non-normal process, defined as

$$C_{p(y)} = \frac{\Phi^{-1}(0.5 + 0.5[F(USL)\text{-}F(LSL)])}{3} \tag{5}$$

$$C_{pk(y)} = \min\{C_{pu(y)}, C_{pl(y)}\} = \min\left\{\frac{\Phi^{-1}[F(USL)]}{3}, \frac{\Phi^{-1}[1\text{-}F(LSL)]}{3}\right\} \tag{6}$$

*where* $\Phi$ is the CDF of standard normal distribution and $F$ represents the CDF of the distribution.

Percentile-based indices are commonly applied. However, these indices may be misleading because they only consider the percentiles, ignoring other features of the distribution. When the processes that need to be evaluated are distributed differently, but have the same 0.135 percentile, median value and 99.865 percentile, the method gives the same evaluation result despite the differences in process capability that these differing distributions represent. Furthermore, accurately identifying the 0.135 percentile of the distribution requires an extremely large amount of data. Thus, this method may not give exact evaluation results.

Yield-based indices can effectively reflect the PNC, while they fail to measure the dispersion degree of the process within the specification limits. Processes distributed differently may also have the same PNC. Figure 1 shows probability density functions (PDFs) and CDFs of normal and uniform distributions which have the same PNC. Yield-based indices give both the normally and uniformly distributed data the same evaluation result, even though the two distributions show processes with very different quality (the products from a normally distributed manufacturing process are far more likely to function perfectly, while products from the process

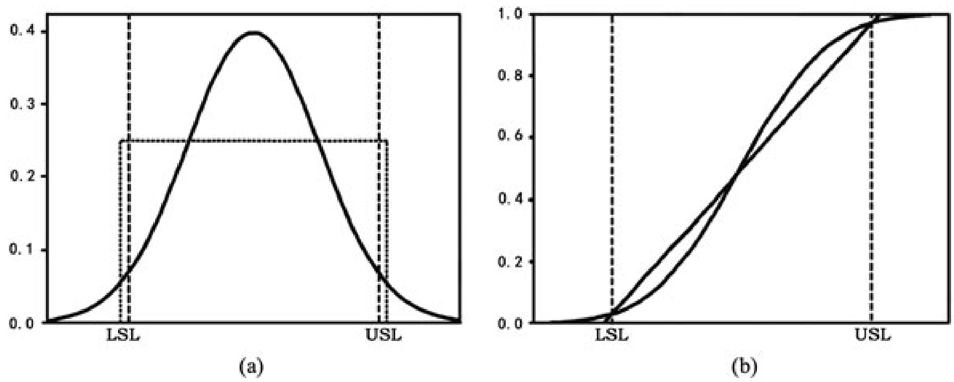


**Figure 1.** Distributions with the same PNC: (a) PDF and (b) CDF

with the uniform distribution are far more likely to be borderline-defective). Therefore, the yield-based method may not be appropriate when the process is non-normally distributed.

These percentile-based and yield-based methods consider specific features of the process distribution, but ignore the impact of the large amount of distribution points on the process capability. In the estimation of the process dispersion, all the points, but not some certain points on the process distribution should be concerned. To address this issue, a CRPS evaluation method will be introduced. Nau[19] proved CRPS is meaningful in ranking the distributions, so CRPS is used to measure the dispersion degree of the process distribution and CRPS-based process evaluation indices will be proposed in the next section.

## 3. Process capability analysis via CRPS

### 3.1. Continuous ranked probability score

CRPS, short for continuous ranked probability score, is an ensemble-specific assessment tool used to evaluate the forecasting performance of the ensemble. CRPS has drawn much attention in recent research and is well applied in probabilistic forecasting field (see Hersbach,[17] Pinson *et al.*,[21] Thorarinsdottir *et al.*[22]). CRPS can access sharpness as well as accuracy of a forecasting process simultaneously.[23] It can be obtained by computing the integral of the square difference between two CDFs of deterministic and probabilistic forecasts.[24,25] CRPS is defined as:

$$S(F,y)=\int_{-\infty}^{+\infty}(F(t)-H(t-y))^2\,dt \tag{7}$$

*where* $H(t\text{-}y)$ is the indicative function. If $t<y$, the function value is 0; otherwise, the function value is 1. The schematic diagram of CRPS can be seen in Figure 2. The CRPS can be presented as the shaded area between CDF and $y$.

CRPS generalizes the absolute error,[26,27] and it can be used as a metric to compare the deterministic and probabilistic forecasts directly. CRPS is also negative oriented: the smaller the value of CRPS is, the more accurate the prediction will be. When the forecast is deterministic, that is $F(t)=H(t-y)$, CRPS reaches the smallest value 0.

According to Equation (7), CRPS can also be expressed as

$$S(F,y)=\int_{-\infty}^{y}(F(t))^2dt+\int_{y}^{+\infty}(F(t)-1)^2\,dt=S_l(F,y)+S_u(F,y) \tag{8}$$

*where* $S_l(F,y)$ is the CRPS value below $y$ and $S_u(F,y)$ is the CRPS value beyond $y$.

Gneiting *et al.*[18] pointed out that when $F_N$ represents CDF of normal distribution with mean $\mu$ and variance $\sigma^2$, the formula of CRPS is

$$S\left(F_N\left(\mu,\sigma^2\right),y\right)=\sigma\left(\left(\frac{y-\mu}{\sigma}\right)\left(2\Phi\left(\frac{y-\mu}{\sigma}\right)-1\right)+2\phi\left(\frac{y-\mu}{\sigma}\right)-\frac{1}{\sqrt{\pi}}\right) \tag{9}$$

*where* $\phi$ represents PDF of standard normal distribution.

Nau[19] studied the relationship between CRPS and distance measures on classes of probability distributions. With respect to any distance $d$ which is symmetric and satisfies the triangle inequality, it is shown that:

$$S(F_1,y)\geq S(F_2,y)\Leftrightarrow d(f_1,y)\geq d(f_2,y) \tag{10}$$

*where* $f_1$ and $f_2$ represent two PDFs while $F_1$ and $F_2$ are corresponding CDFs. And when $f_1=f_2$, the equality holds. CRPS can effectively measure the distance between a distribution function and a certain point.[17] When $F_1$ and $F_2$ represent different CDFs of process distributions, $y$ is the distribution center, (10) still holds. Therefore, CRPS is able to reflect the dispersion degree of the process distribution.

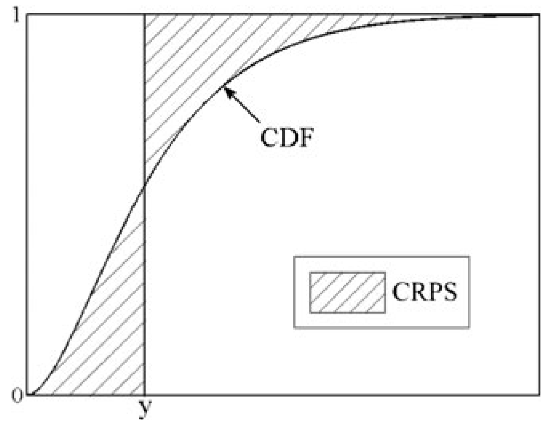


**Figure 2.** Schematic diagram of CRPS

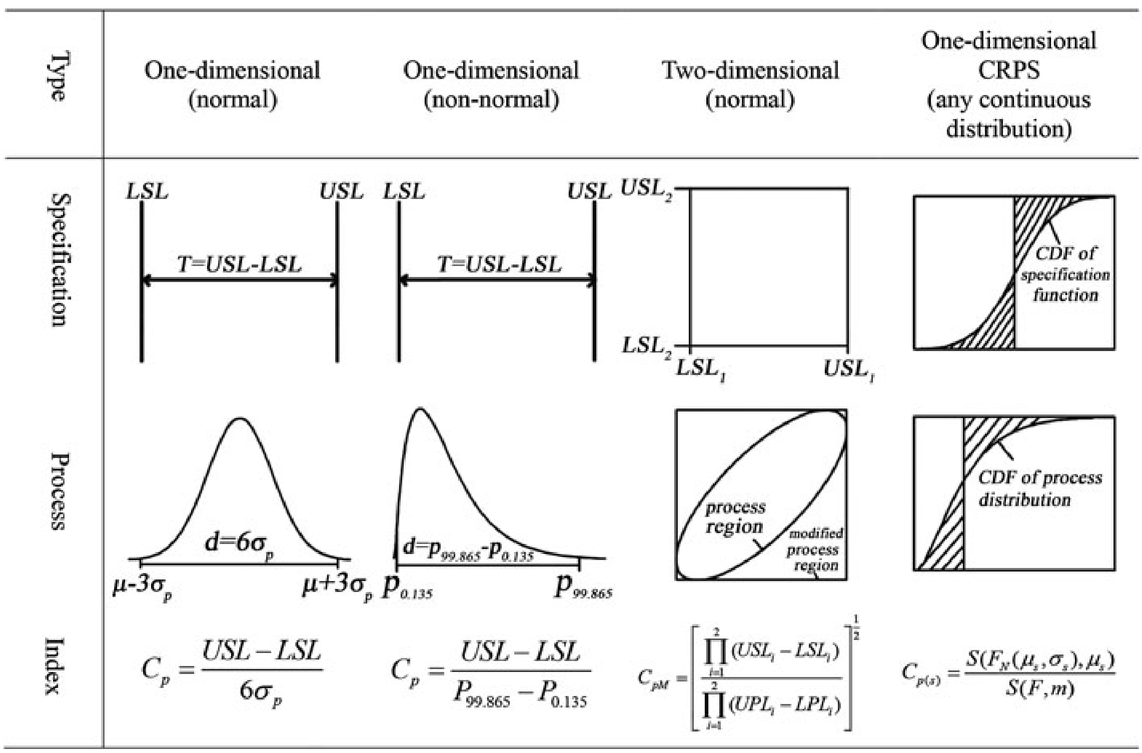


**Figure 3.** The expression of specification limits and process spread in capability indices.
Note: $UPL_i$ and $LPL_i$ (i = 1, 2) represent the upper and lower process limits; $T$ is the width of the tolerance

A smaller CRPS value indicates a more capable process, while a larger CRPS value indicates a worse performance. As CRPS can be applied to any continuous distribution, it can be used in process capability analysis to reflect the process spread of normal or non-normal distributions.

### *3.2. The proposed CRPS-based indices*

PCI is defined as the ratio of the specification spread to the process spread. If $F$ is set as the CDF of the process distribution, process spread can be represented by $S(F, m)$ using the CRPS method, whether the process distribution is normal or not. Median value $m$ is always regarded as the center of a non-normal distribution.[28] Moreover, for any type of continuous distribution, the distance computed by CRPS from the distribution to the median value is the smallest, compared with distances from the distribution to the mean or other indicators of central tendency (see Appendix A). So $S(F, m)$ is able to reflect the dispersion of the process distribution.

To evaluate the process capability, the process spread is usually compared with the specification. The expression of the specification and the process range should keep consistent. Figure 3 presents the expressions of specification limits and process range in different indices. It shows that when the process is one dimensional, the PCIs are obtained by comparing the widths of the specification and the process spread; and in multivariate cases, such as two variables, the PCI is the ratio of the areas between the process spread and the specification.[29,30] If we want to construct CRPS-based indices, both of the process spread and the specification should be expressed by CRPS.

To obtain the CRPS form of the specification, we build a specification function which follows the form of normal distribution with mean $\mu_s = \frac{USL+LSL}{2}$ and standard deviation $\sigma_s = \frac{USL-LSL}{6}$. Then $S(F_N(\mu_s, \sigma_s), \mu_s)$ is used to reflect the specification range. Then CRPS-based $C_{p(s)}$ is proposed to measure the potential capability of the process, defined as

$$C_{p(s)} = \frac{S(F_N(\mu_s, \sigma_s), \mu_s)}{S(F, m)}. \tag{11}$$

CRPS-based $C_{pk(s)}$ is also proposed. The process distribution is divided by $m$ value into two deviations, the upper-sided distribution and the lower-sided one. When the process capability for upper-sided specification is considered, the dispersion degree of the upper-sided distribution is expressed as $S_u(F, m)$ by CRPS. The corresponding specification interval is $[m, USL]$. We also define a function for the upper-sided specification which follows normal distribution with mean $\mu_u = m$ and standard deviation $\sigma_u = (USL - m)/3$. Then upper-sided specification range can be expressed by CRPS as $S(F_N(m, \sigma_u), m)/2$. So PCI for the upper-sided specification is

$$C_{pu(s)} = \frac{S(F_N(m, \sigma_u), m)}{2S_u(F, m)}. \tag{12}$$

Similarly, PCI for the lower-sided specification is

$$C_{pl(s)} = \frac{S(F_N(m, \sigma_l), m)}{2S_l(F, m)} \tag{13}$$

*where* $S_l(F, m)$ reflects the range of the lower distribution and $\sigma_l = (m - LSL)/3$. So PCI for the two-sided specification is

$$C_{pk(s)} = \min\{C_{pu(s)}, C_{pl(s)}\}. \tag{14}$$

When the process distribution is normal with mean $\mu$ and standard deviation $\sigma_p$, we have

$$C_{p(s)} = C_p \tag{15}$$

*and*

$$C_{pk(s)} = C_{pk}. \tag{16}$$

Namely, $C_p$ and $C_{pk}$ are special cases of CRPS-based indices (The proof can be seen in Appendix B).

CRPS-based indices work equally well for any continuous distribution, whether it is normal or not. In the following sections, we apply CRPS to non-normal distributions. As Wang *et al.*[31] proved, $S(F_e, y)$ is an asymptotic unbiased estimator of $S(F, y)$, where $F_e$ represents empirical CDF. Therefore, we use empirical CDF to curve the unknown distributions. The detailed calculations of $S(F_e, m)$ can be seen in Appendix C. The calculation process is complex when done manually, but with the help of computers, it is very easy.

## 4. Simulations

To illustrate the CRPS-based method outperforms percentile-based method and yield-based method, respectively, we simulate three scenarios:

Scenario 1: different process distributions with the same mean and variance;
Scenario 2: different process distributions with the same $P_{0.135}$, $P_{50}$ and $P_{99.865}$;
Scenario 3: different process distributions with the same PNC.

Scenario 1 is presented to illustrate how the CRPS-based method responds to changes of skewness and kurtosis of the distribution. Scenario 2 and Scenario 3 are special cases that hold percentiles or PNC equal and show that CRPS reflects quality variations that others cannot detect.

The proposed indices are applied for each scenario to estimate the process capability. The results are compared with the existing methods which include percentile and yield-based methods. In each scenario, every simulation runs for 10 000 times and the evaluation result is averaged. For the sake of data fitting, any evaluation result greater than 4 is set as 4.

1 Scenario 1: distributions with the same mean and variance

The Pearson distribution family is used in the simulation part, because it consists of the most commonly used distributions, such as normal, Gamma, t, F, Beta and log-normal distributions. Pearson distributions have four parameters, mean ($\mu$), standard deviation ($\sigma$), skewness ($\alpha_3$) and kurtosis ($\alpha_4$). In Table I and Table II, 10 kinds of Pearson distributions with the same mean and standard deviation ($\mu = 3$, $\sigma = 1$) are generated. The specification limits are [0, 6], and the target value is 3. For all these cases, both $C_p$ and $C_{pk}$ equal to 1. Evaluation results given by CRPS, percentile and yield-based indices are shown in Table I and Table II as sample size $n$, skewness and

**Table I.** $C_{p(s)}$, $C_{p(p)}$ and $C_{p(y)}$ of Pearson distributions

| N | $\alpha_3$ | $\alpha_4$ | $n=10$ | | | $n=50$ | | | $n=100$ | | | $n=1\,000\,000$(known parameter case) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ |
| (a) | 0 | 3 | 1.13 | 2.08 | 3.91 | 1.02 | 1.36 | 3.59 | 1.01 | 1.21 | 3.28 | 1.00 | 1.00 | 1.00 |
| (b) | 0.5 | 3 | 1.11 | 2.12 | 3.87 | 1.00 | 1.43 | 3.43 | 0.99 | 1.30 | 2.99 | 0.98 | 1.10 | 0.96 |
| (c) | 1 | 3 | 1.17 | 2.29 | 4.00 | 0.99 | 1.70 | 4.00 | 0.97 | 1.62 | 4.00 | 0.95 | 1.53 | 4.00 |
| (d) | 0 | 4 | 1.18 | 2.11 | 3.76 | 1.09 | 1.29 | 3.02 | 1.08 | 1.11 | 2.35 | 1.07 | 0.85 | 0.89 |
| (e) | 0.5 | 4 | 1.17 | 2.12 | 3.77 | 1.08 | 1.31 | 3.01 | 1.07 | 1.14 | 2.36 | 1.05 | 0.89 | 0.90 |
| (f) | 1 | 4 | 1.16 | 2.20 | 3.66 | 1.04 | 1.45 | 2.74 | 1.03 | 1.30 | 2.02 | 1.02 | 1.08 | 0.86 |
| (g) | 0 | 6 | 1.25 | 2.14 | 3.64 | 1.15 | 1.24 | 2.68 | 1.14 | 1.04 | 1.95 | 1.13 | 0.75 | 0.85 |
| (h) | 0.5 | 6 | 1.24 | 2.16 | 3.67 | 1.15 | 1.25 | 2.70 | 1.14 | 1.06 | 1.97 | 1.13 | 0.77 | 0.86 |
| (i) | 1 | 6 | 1.23 | 2.18 | 3.65 | 1.13 | 1.30 | 2.66 | 1.12 | 1.11 | 1.90 | 1.10 | 0.83 | 0.85 |
| (j) | 1.5 | 6 | 1.28 | 2.31 | 3.51 | 1.14 | 1.44 | 2.24 | 1.12 | 1.26 | 1.50 | 1.10 | 0.99 | 0.81 |

**Table II.** $C_{pk(s)}$, $C_{pk(p)}$ and $C_{pk(y)}$ of Pearson distributions

| N | $\alpha_3$ | $\alpha_4$ | $n=10$ | | | $n=50$ | | | $n=100$ | | | $n=1\,000\,000$ (known parameter case) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ | $C_{p(s)}$ | $C_{p(p)}$ | $C_{p(y)}$ |
| (a) | 0 | 3 | 1.19 | 2.24 | 3.95 | 1.04 | 1.38 | 3.77 | 1.02 | 1.23 | 3.58 | 1.00 | 1.00 | 1.00 |
| (b) | 0.5 | 3 | 1.15 | 2.07 | 3.85 | 0.98 | 1.23 | 3.42 | 0.96 | 1.09 | 2.93 | 0.94 | 0.89 | 0.89 |
| (c) | 1 | 3 | 1.05 | 1.78 | 4.00 | 0.85 | 1.17 | 4.00 | 0.82 | 1.10 | 4.00 | 0.80 | 1.03 | 4.00 |
| (d) | 0 | 4 | 1.29 | 2.33 | 3.87 | 1.11 | 1.33 | 3.44 | 1.09 | 1.13 | 3.01 | 1.07 | 0.85 | 0.89 |
| (e) | 0.5 | 4 | 1.24 | 2.20 | 3.77 | 1.07 | 1.20 | 3.04 | 1.04 | 1.02 | 2.41 | 1.02 | 0.75 | 0.83 |
| (f) | 1 | 4 | 1.15 | 1.96 | 3.65 | 0.98 | 1.10 | 2.69 | 0.95 | 0.95 | 1.95 | 0.93 | 0.76 | 0.78 |
| (g) | 0 | 6 | 1.35 | 2.41 | 3.83 | 1.18 | 1.29 | 3.21 | 1.16 | 1.08 | 2.68 | 1.13 | 0.75 | 0.85 |
| (h) | 0.5 | 6 | 1.34 | 2.31 | 3.74 | 1.15 | 1.20 | 2.94 | 1.12 | 0.99 | 2.29 | 1.10 | 0.68 | 0.81 |
| (i) | 1 | 6 | 1.28 | 2.18 | 3.65 | 1.10 | 1.11 | 2.60 | 1.07 | 0.92 | 1.87 | 1.05 | 0.64 | 0.77 |
| (j) | 1.5 | 6 | 1.22 | 1.97 | 3.48 | 1.02 | 1.01 | 2.21 | 1.00 | 0.86 | 1.43 | 0.97 | 0.64 | 0.72 |

kurtosis increase. A total of 1 000 000 data sets are also simulated for each distribution as a known parameter case and the results given by different methods are estimates of their "actual values".

Table I and Table II show that:

- i All of the indices give different evaluation of process quality to these distributions. Even though they have the same mean and variance, their process capability can be different, and the choice of indicator matters.
- ii When the process parameters are unknown, the results given by CRPS-based method are closer to its actual value than that given by percentile and yield-based method. CRPS-based method is more stable.
- iii For the known parameters case, when only skewness increases, the $C_{p(s)}$ given by CRPS-based method tends to decrease. While the $C_{p(p)}$ given by the percentile-based method increases and $C_{p(y)}$ given by yield-based method has no clear trend; when only kurtosis increases, the $C_{p(s)}$ increases, while $C_{p(p)}$ and $C_{p(y)}$ decrease. In measuring the process departure, these methods all have trends of decreasing with the increase of skewness, while only $C_{pk(s)}$ has the same trend with the change of kurtosis.

2 Scenario 2: distributions with the same $P_{0.135}$, $P_{50}$ and $P_{99.865}$

In this group of simulations, the 0.135, 50 and 99.865 percentiles of the process distribution remain unchanged and the distribution shape varies. Normal, Pearson, uniform and bimodal distributions are simulated, as shown in Figure 4. The specification limits is defined as [14, 36], and the target value is 25. A total of 10 000 data points are generated for each distribution. The evaluation results are shown in Table III.

In this case, for the four distributions are all symmetric, $C_{p(s)}$ equals to $C_{pk(s)}$. So it is with the other two methods. Although the shapes of distributions (a–d) in Table III are quite different, the percentile-based method gives all of these distributions the same evaluation. As is evident in Figure 4, (b) is the most concentrated distribution while (c) and (d) are not desired ones; this ranking is identical with the evaluation results given by CRPS-based method. The CRPS-based method is able to identify the fluctuations in different distribution shapes which have the same 0.135, 50 and 99.865 percentiles, while the percentile-based method gives out the same result, despite obvious differences in quality.

3 Scenario 3: different distributions with the same PNC

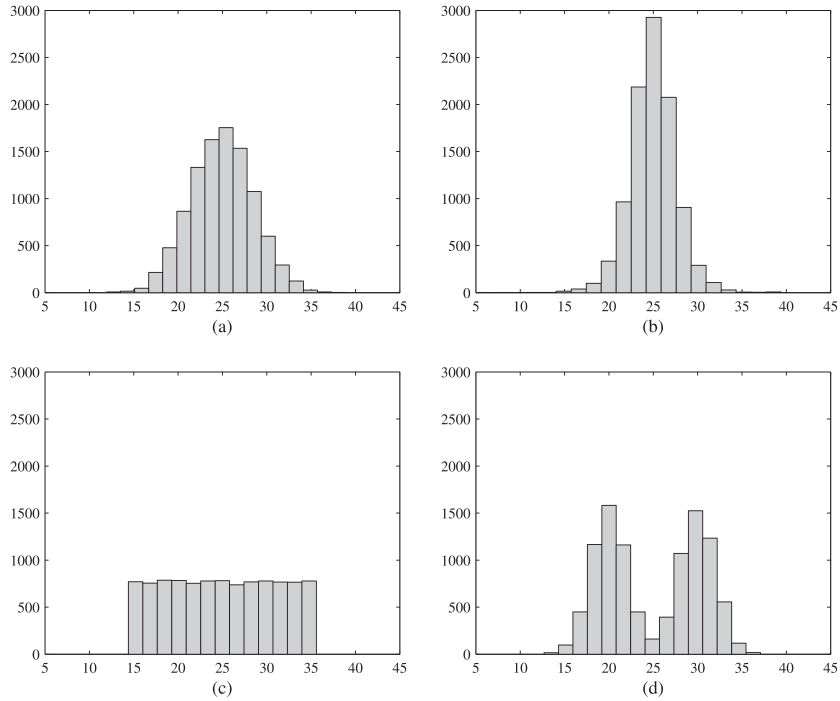


**Figure 4.** Histograms of distributions with the same $P_{0.135}$, $P_{50}$ and $P_{99.865}$

**Table III.** Evaluation results of distributions with the same $P_{0.135}$, $P_{50}$ and $P_{99.865}$

| N | Distribution | $C_{p(s)}$ | $C_{pk(s)}$ | $C_{p(p)}$ | $C_{pk(p)}$ | $C_{p(y)}$ | $C_{pk(y)}$ |
|---|---|---|---|---|---|---|---|
| (a) | $N(25, 3.52^2)$ | 1.04 | 1.04 | 1.04 | 1.04 | 1.04 | 1.04 |
| (b) | Pearson(25,2.64,0,6) | 1.57 | 1.57 | 1.04 | 1.04 | 1.02 | 1.02 |
| (c) | $U$ (14.4,35.6) | 0.49 | 0.49 | 1.04 | 1.04 | 4.00 | 4.00 |
| (d)* | Bimodal | 0.44 | 0.44 | 1.04 | 1.04 | 1.07 | 1.07 |

Note: *This is a symmetrical bimodal distribution generated by two normal distributions, namely, $N(20, 2^2)$ and $N(30, 2^2)$.

Under this scenario, the PNC of the process distribution is fixed. Three kinds of distributions are simulated, which are normal, uniform and gamma distributions, as is shown in Figure 5. The specification limits are defined as [0.1, 7.5] and the target value is 3.8. A total of 10 000 data points are generated for each distribution. The evaluation results for each method are shown in Table IV.

In Table IV, yield-based indices give both distribution (a) and (b) the same evaluation. While in Figure 5, distribution (b) has more marginal products and would be worse than (a) if tighter process capability requirements were needed in the future. CRPS-based indices can correctly identify the capability of normal and uniform distribution. Two skewed distributions (c) and (d) are symmetrical with respect to the target value. (c) is right skewed and (d) is left skewed. CRPS-based indices correctly identify that (c) and (d) are skewed in different directions, while yield-based indices fail.

From the evaluation results given in Tables I-IV, it can be concluded that the CRPS-based method is more stable and effective in reflecting the process fluctuation than others.

## 5. An example

In this section, we select an example to illustrate the application of CRPS-based indices.

Mahmoud *et al.*[32] provided data on the manufacture of the rotor shaft with a specified diameter between 7.986 and 7.995 mm. The histogram of the distribution is shown in Figure 6. It is clear that the distribution is non-normal. Evaluation results given by different methods are shown in Table V.

Table V shows the process capability given by CRPS-based indices is smaller than the result given by percentile-based method. In Figure 6, many data points fall on the left edge of the distribution, but not the distribution center. For that reason, this process distribution is worse than condensed distributions that have more observations near the median than near the 0.135 percentile of

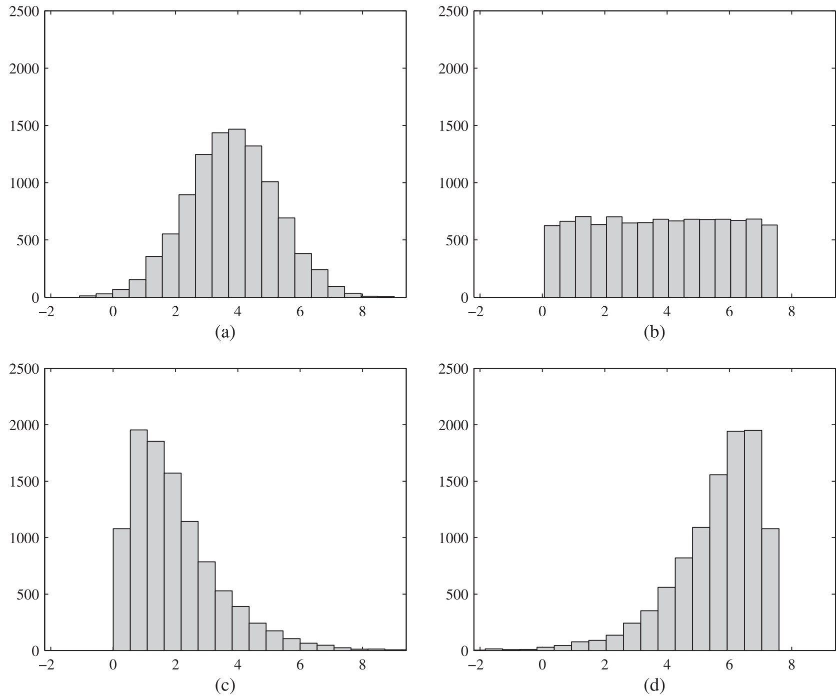


**Figure 5.** Histograms of distributions with the same PNC

**Table IV.** The evaluation results of distributions with the same PNC

| N | Distribution | CRPS | | | Percentile | | | Yield | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $C_{p(s)}$ | $C_{pu(s)}$ | $C_{pl(s)}$ | $C_{p(p)}$ | $C_{pu(p)}$ | $C_{pl(p)}$ | $C_{p(y)}$ | $C_{pu(y)}$ | $C_{pl(y)}$ |
| (a) | *N*(3.8,1.425) | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 |
| (b) | *U*(0.065,7.535) | 0.46 | 0.46 | 0.46 | 0.99 | 0.99 | 0.99 | 0.87 | 0.87 | 0.87 |
| (c) | Gamma(2,1) | 0.96 | 1.26 | 0.51 | 0.84 | 0.81 | 0.97 | 0.87 | 0.87 | 0.87 |
| (d) | 7.6-Gamma(2,1) | 0.96 | 0.51 | 1.26 | 0.84 | 0.97 | 0.81 | 0.87 | 0.87 | 0.87 |

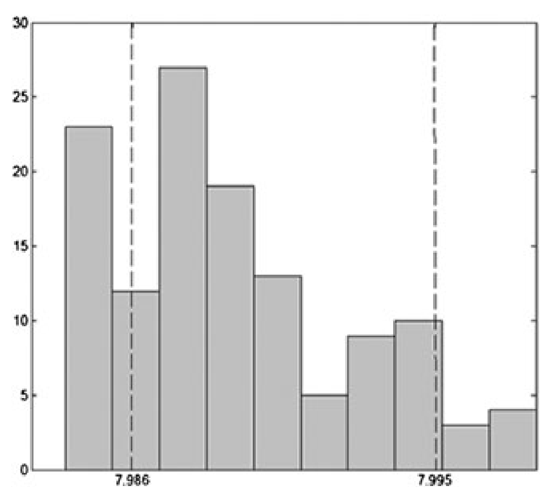


**Figure 6.** Histogram of the shaft diameter

**Table V.** Evaluation results using different methods

| | | | |
|---|---|---|---|
| CRPS-based method | $C_{p(s)}$ | $C_{pl(s)}$ | $C_{pu(s)}$ |
| | 0.39 | 0.22 | 0.66 |
| Percentile-based method | $C_{p(p)}$ | $C_{pl(p)}$ | $C_{pu(p)}$ |
| | 0.64 | 0.60 | 0.67 |
| Yield-based method | $C_{p(y)}$ | $C_{pl(y)}$ | $C_{pu(y)}$ |
| | 0.39 | 0.30 | 0.53 |

the distribution. So the result given by CRPS-based method is a better reflection of quality than that given by percentile-based method. Moreover, yield-based method performs almost as well as CRPS-based method in this case.

## 6. Discussion

While in the application of CRPS-based method, the evaluation result can be affected by the sample size. When the process distribution is normal, the requirement for the sample size is consistent with that of the traditional method; for non-normal process distribution, empirical CDF is recommended to fit the sample in this paper. If the sample size is big enough, the empirical CDF can be smooth, and the result given by CRPS-based method is much accurate; however, if the sample size is limited, the result given by CRPS-based method can be unreliable.

Furthermore, CRPS can also be used in constructing indices when the process is off target. Additionally, Burdick *et al.*[33] illustrate the existence of non-normal system measurement system error in factories. CRPS can also be used in measurement system analysis to measure the variance of measurement system error whose distribution is normal or non-normal.

## 7. Conclusions

PCIs $C_p$ and $C_{pk}$ are usually used to evaluate the process capability when the distribution is normal. When the process distribution is non-normal, $C_p$ and $C_{pk}$ are often misleading. PCIs under non-normal distribution are proposed using different methods, such as percentile and yield-based indices. However, these indices only consider certain features of the process distribution and may not be appropriate to reflect the process capability when the distribution shape changes.

In this paper, a CRPS-based process capability method is proposed. This method is suitable for any continuous distribution. The proposed CRPS-based method considers the distance between the process distribution and its median value, so it can effectively reflect the dispersion degree of the process distribution. Particularly, $C_p$ and $C_{pk}$ are special cases of CRPS-based indices under normal conditions. A set of simulation analysis and an example are given and comparisons between the CRPS, percentile and yield-based method are listed. The results show that the CRPS-based method is better than the others in the identification of process fluctuations and the true degree of consistency in manufacturing or process control.

## Acknowledgements

The authors are grateful for the help and advice on this paper from Joe Gyekis (Pennsylvania State University). This research is supported by the National Nature Science Foundation of China (71102140).

## References

1. Juran JM. *Quality Control Handbook* (3rd edn). McGraw-Hill: New York, 1974.
2. Clements JA. Process capability calculations for non-normal distributions. *Quality Progress* 1989; **22**:95–97.
3. Pearn WL, Kotz S. Application of Clements' method for calculating second- and third-generation process capability indices for non-normal Pearson populations. *Quality Engineering* 1994; **7**:139–145.
4. Chen KS, Pearn WL. An application of non-normal process capability indices. *Quality and Reliability Engineering International* 1997; **13**:355–360.
5. Tong LI, Chen JP. Lower confidence limits of process capability indices for nonnormal process distributions. *International Journal of Quality & Reliability Management* 1998; **15**:907–919.
6. Peng C. Estimating and testing quantile-based process capability indices for processes with skewed distributions. *Journal of Data Science* 2010; **8**:253–268.
7. Yang J, Gang T, Cheng Y, Xie M. Process capability indices based on the highest density interval. *Quality and Reliability Engineering International* 2014. doi:10.1002/qre.1665.

8. Castagliola P. Evaluation of non-normal process capability indices using Burr's distributions. *Quality Engineering* 1996; **8**:587–593.
9. Chen JP. Re-evaluating the process capability indices for non-normal distributions. *International Journal of Production Research* 2000; **38**:1311–1324.
10. Chao MT, Lin DKJ. Another look at the process capability index. *Quality and Reliability Engineering International* 2006; **22**:153–163.
11. Abbasi B, Niaki STA. Estimating process capability Indices of multivariate nonnormal processes. *The International Journal of Advanced Manufacturing Technology* 2010; **50**:823–830.
12. Box GE, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society: Series B Methodological* 1964; **26**:211–252.
13. Somerville SE, Montgomery DC. Process capability indices and non-normal distributions. *Quality Engineering* 1996; **9**:305–316.
14. Pal S. Evaluation of nonnormal process capability indices using generalized lambda distribution. *Quality Engineering* 2004; **17**:77–85.
15. Abbasi B. A neural network applied to estimate process capability of non-normal processes. *Expert Systems with Applications* 2009; **36**:3093–3110.
16. Chang YS, Choi IS, Bai DS. Process capability indices for skewed populations. *Quality and Reliability Engineering International* 2002; **18**:383–393.
17. Hersbach H. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting* 2000; **15**:559–570.
18. Gneiting T, Raftery AE, Westveld AH III, Goldman T. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review* 2005; **133**:1098–1118.
19. Nau RF. Should scoring rules be effective? *Management Science* 1985; **31**:527–535.
20. Wu HH, Wang JS, Liu TL. Discussions of the Clements-based process capability indices. *In Proceedings of the 1998 CIIE National Conference* 1998: 561–566.
21. Pinson P, Reikard G, Bidlot JR. Probabilistic forecasting of the wave energy flux. *Applied Energy* 2012; **93**:364–370.
22. Thorarinsdottir TL, Gneiting T. Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2010; **173**:371–388.
23. Fraley C, Raftery AE, Gneiting T. Calibrating multimodel forecast ensembles with exchangeable and missing members using Bayesian model averaging. *Monthly Weather Review* 2010; **138**:190–202.
24. Boucher MA, Laliberté JP, Anctil F. An experiment on the evolution of an ensemble of neural networks for streamflow forecasting. *Hydrology and Earth System Sciences* 2010; **14**:603–612.
25. Voisin N, Schaake JC, Lettenmaier DP. Calibration and downscaling methods for quantitative ensemble precipitation forecasts. *Weather and Forecasting* 2010; **25**:1603–1627.
26. Grimit EP, Gneiting T, Berrocal VJ, Johnson NA. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. *Quarterly Journal of the Royal Meteorological Society* 2006; **132**:2925–2942.
27. Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 2007; **69**:243–268.
28. Tang LC, Than SE. Computing process capability indices for non-normal data: a review and comparative study. *Quality and Reliability Engineering International* 1999; **15**:339–353.
29. Pan JN, Lee CY. New capability indices for evaluating the performance of multivariate manufacturing processes. *Quality and Reliability Engineering International* 2010; **26**:3–15.
30. Shahriari H, Abdollahzadeh M. A new multivariate process capability vector. *Quality Engineering* 2009; **21**:290–299.
31. Wang W, Lin DKJ. Another look at run length distribution. 2013.
32. Mahmoud MA, Aufy SA. Process capability evaluation for a non-normally distributed one. *Engineering &Technology Journal* 2013; **31**:2345–2358.
33. Burdick RK, Borror CM, Montgomery DC. A review of methods for measurement systems capability analysis. *Journal of Quality Technology* 2003; **35**:342–354.

# Appendix A. Proof of the minimum CRPS value of a continuous distribution

According to Equation (8), $S(F,y)=\int_{-\infty}^{y}(F(t))^2dt+\int_{y}^{+\infty}(F(t)-1)^2dt$. If $y$ moves $\Delta y$ ($\Delta y>0$ and $\Delta y\to 0$), Then

$$\begin{aligned}S(F,y+\Delta y)&=\int_{-\infty}^{y+\Delta y}(F(t))^2dt+\int_{y+\Delta y}^{+\infty}(F(t)-1)^2dt\\&=\int_{-\infty}^{y}(F(t))^2dt+\int_{y}^{+\infty}(F(t)-1)^2dt+\int_{y}^{y+\Delta y}(F(t))^2dt-\int_{y}^{y+\Delta y}(1-F(t))^2dt\\&=S(F,y)+\int_{y}^{y+\Delta y}\left[(F(t))^2-(1-F(t))^2\right]dt\\&=S(F,y)+\int_{y}^{y+\Delta y}[2F(t)-1]dt.\end{aligned}$$

$F$ is an increasing function and $F(m)=0.5$. When $y+\Delta y<m$, $2F(t)-1<0$, so $\int_{y}^{y+\Delta y}[2F(t)-1]dt<0$ and $S(F,y+\Delta y)<S(F,y)$; when $y\geq m$, $2F(t)-1\geq 0$, So $\int_{y}^{y+\Delta y}[2F(t)-1]dt>0$ and $S(F,y+\Delta y)>S(F,y)$.

So it can be proved that $S(F,y)$ first decreases and then increases. When $y=m$, $S(F,y)$ reaches its minimum value.

# Appendix B. Proof of theorem

According to Equation (9), when the process distribution follows $N\left(\mu,\sigma_p^2\right)$ and $m$ coincides with $\mu$, then

$$C_{p(s)}=\frac{S(F_N(\mu_s,\sigma_s),\mu_s)}{S\left(F_N\left(\mu,\sigma_p\right),\mu\right)}=\frac{\sigma_s\left(2\phi(0)-\frac{1}{\sqrt{\pi}}\right)}{\sigma_p\left(2\phi(0)-\frac{1}{\sqrt{\pi}}\right)}=\frac{USL-LSL}{6\sigma_p}=C_p$$

$$C_{pu(s)}=\frac{S(F_N(m,\sigma_u),m)}{2S_u\left(F_N\left(\mu,\sigma_p\right),\mu\right)}=\frac{S(F_N(m,\sigma_u),m)}{S\left(F_N\left(\mu,\sigma_p\right),\mu\right)}=\frac{\sigma_u\left(2\phi(0)-\frac{1}{\sqrt{\pi}}\right)}{\sigma_p\left(2\phi(0)-\frac{1}{\sqrt{\pi}}\right)}=\frac{USL-\mu}{3\sigma_p}=C_{pu}.$$

In the same way, $C_{pl(s)}=C_{pl}$. Then $C_{pk(s)}=\min\{C_{pu},C_{pl}\}=C_{pk}$.

# Appendix C. Calculation of CRPS for empirical CDF

The formula of an empirical CDF is

$$F_e(x) = \begin{cases} 0 & x < x_1 \\ \frac{i}{n} & x_i \le x < x_{i+1} \\ 1 & x \ge x_n \end{cases}$$

*where* $n$ is the sample size and $x_1, x_2, \ldots, x_i, \ldots, x_n$ ($i = 1, 2, \ldots, n$) are ordered samples. $m$ stands for median value. Then the CRPS value can be calculated by

a. if $n$ is an odd number ($n > 1$),

$$S(F_e(x), m) = \sum_{i=1}^{\frac{n-1}{2}} \left(x_{(i+1)} - x_{(i)}\right)\left(\frac{i}{n}\right)^2 + \sum_{i=\frac{n+1}{2}}^{n-1} \left(x_{(i+1)} - x_{(i)}\right)\left(1 - \frac{i}{n}\right)^2$$

$$S_l(F_e(x), m) = \sum_{i=1}^{\frac{n-1}{2}} \left(x_{(i+1)} - x_{(i)}\right)\left(\frac{i}{n}\right)^2$$

$$S_u(F_e(x), m) = \sum_{i=\frac{n+1}{2}}^{n-1} \left(x_{(i+1)} - x_{(i)}\right)\left(1 - \frac{i}{n}\right)^2$$

b. if $n$ is an even number ($n > 2$),

$$S(F_e(x), m) = \sum_{i=1}^{\frac{n}{2}} \left(x_{(i+1)} \text{-} x_{(i)}\right)\left(\frac{i}{n}\right)^2 + \sum_{i=\frac{n}{2}+1}^{n\text{-}1} \left(x_{(i+1)} \text{-} x_{(i)}\right)\left(1 - \frac{i}{n}\right)^2$$

$$S_l(F_e(x), m) = \sum_{i=1}^{\frac{n}{2}\text{-}1} \left(x_{(i+1)} - x_{(i)}\right)\left(\frac{i}{n}\right)^2 + \left(m - x_{\left(\frac{n}{2}\right)}\right)\left(\frac{1}{2}\right)^2$$

$$S_u(F_e(x), m) = \left(x_{\left(\frac{n}{2}+1\right)} - m\right)\left(\frac{1}{2}\right)^2 + \sum_{i=\frac{n}{2}+1}^{n\text{-}1} \left(x_{(i+1)} \text{-} x_{(i)}\right)\left(1 - \frac{i}{n}\right)^2$$

*Authors' biographies*

**Liangxing Shi** is an Associate Professor in the Department of Industrial Engineering at Tianjin University, China. He received his PhD from Tianjin University in 2008. His research interests include quality engineering, six sigma, industrial engineering, and operations management.

**Hongye Ma** is a Master candidate in the Department of Industrial Engineering at Tianjin University, China. Her research interests focus on quality management and industrial engineering.

**Dennis K. J. Lin** is a Distinguished Professor of Statistics and Supply Chain Management at the Pennsylvania State University, USA. He received his PhD in Statistics from the University of Wisconsin–Madison, USA, in 1988. His research interests are quality assurance, industrial statistics (design of experiment, reliability, statistical process control, quality assurance), data mining and response surface.