Research Article

Nonparametric Confidence Limits of Quantile-Based Process Capability Indices

Cheng Peng¹ and Jiaqing Xu²

¹ Department of Mathematics and Statistics, University of Southern Maine, 96 Falmouth Street, Portland, ME 04104, USA

² College of Sciences, Ningbo University of Technology, Ningbo, Zhejiang 315211, China

Correspondence should be addressed to Cheng Peng, cpeng@usm.maine.edu

Received 1 August 2011; Accepted 23 December 2011

Academic Editor: Suk joo Bae

Copyright © 2012 C. Peng and J. Xu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose an asymptotic nonparametric confidence interval for quantile-based process capability indices (PCIs) based on the superstructure $C_{Np}(u, v)$ modified from $C_p(u, v)$ which contains the four basic PCIs, C_p , C_{pk} , C_{pm} , and C_{pmk} , as special cases. Since the asymptotic variance of the estimator for quantile-based PCIs involves the density function of the underlying process, the existing asymptotic results cannot be used directly to construct confidence limits for PCIs. To obtain a consistent estimator for the asymptotic variance of the estimated quantile-based PCIs, in this paper, we propose to use the kernel density estimator for the underlying process. Consequently, the confidence limits for PCIs are established based on the consistent estimates. A real-life example from manufacturing engineering is used to illustrate the implementation of the proposed methods. Simulation studies are also presented in this paper to compare the two quantile estimators that are used in the definition of PCIs.

1. Introduction

The process capability index (PCI) as a quality monitoring tool has been playing an increasingly important role in analyzing process quality and productivity. It has been widely applied in industrial engineering particularly in quality monitoring and improvement programs in manufacturing systems as well as reliability engineering and environmental engineering. Many PCIs have been proposed since Juran et al. [1] proposed the first PCI C_p . Let USL and LSL be the upper and lower specification limits, $d = (USL - LSL)/2$, $m = (LSL + USL)/2$ and T be the target values. The process mean and standard deviation are denoted by μ and σ . Vännman [2] proposed the following unified superstructure which includes the commonly used four basic PCIs, C_p , C_{pk} , C_{pm} , and C_{pmk} ,

$$C_p(u, v) = \frac{d - u|\mu - m|}{3\sqrt{\sigma^2 + v(\mu - T)^2}}, \quad (1)$$

where u and v are nonnegative constants. We can see from (1) that $C_p(0, 0) = C_p$, $C_p(1, 0) = C_{pk}$, $C_p(0, 1) = C_{pm}$, and $C_p(1, 1) = C_{pmk}$. Since the process mean and standard

deviation-based PCIs implicitly assume the normality of the underlying process, applying such PCIs to skewed process may cause invalid results. Let ξ_α be the α th quantile of the process, that is, $P(X < \xi_\alpha) = \alpha$. Define $\xi = (\xi_{\alpha_1}, \xi_{\alpha_2}, \dots, \xi_{\alpha_k})$ to be the vector of quantiles to be used in the definition of PCIs. Chen and Pearn [3] modified Vännman's [2] $C_p(u, v)$ and proposed a quantile-based PCI superstructure without assuming implicitly the normality of the underlying process as follows:

$$C_{Np}(u, v) = \frac{d - u|\xi_{p_2} - m|}{3\sqrt{\left((\xi_{p_3} - \xi_{p_1})/6\right)^2 + v(\xi_{p_2} - T)^2}}, \quad (2)$$

where $p_1 = 0.00135$, $p_2 = 0.50$ and $p_3 = 0.99865$ which guarantee that the proportion of nonconforming parts of the process is at most 0.27% if the corresponding four PCIs take value greater than 1 for on target normal processes. $(p_1, p_2, p_3) = (0.00135, 0.5, 0.9986)$ will be used throughout this paper. Since two different types of quantile functions will be used in this paper, we will use notation $C_{Np}(u, v) = C_{Np}(u, v; \xi)$ meaning that the PCI superstructure is defined based on the specified quantile ξ .

Chen and Hsu [4] studied the asymptotic properties of $C_{Np}(u, v)$ and proved that the point estimators of quantile-based superstructure using different quantile estimators are asymptotically normally distributed. Since the variance of the estimator involves the density function of the underlying process, the asymptotic results cannot be used directly to construct confidence limits of and perform testing hypotheses on PCIs.

In this paper, a consistent nonparametric estimator of the asymptotic variance of the quantile-based PCIs using the kernel density estimator of the underlying process. Based on the consistent variance estimator, we can construct confidence limits of desired PCIs within the frame work of superstructure.

The rest of this paper is organized as follows. Two point estimators of quantile-based PCIs will be proposed in Section 2. The distribution and inferential properties of the consistent estimators of $C_{Np}(u, v)$ based on different quantile estimators are addressed in Section 3. Nonparametric asymptotic confidence limits will be detailed in Section 4. In Section 5, we present an illustrative example showing the steps of constructing confidence intervals using proposed methods. Some comparative analysis based on simulations will also be presented in this section. The proof of Theorem 2 is given in the appendix.

2. Asymptotic Property of the Quantile-Based Estimators

Since the quantile-based PCIs are functions of process quantiles and some preset numerical characteristics regarding process capability such as process target and specification limits, different estimators of process quantiles give different point estimators of PCIs. In this section, we introduce two point estimators of process quantile functions, namely, (raw) sample quantiles and interpolated quantiles based on which the consistent PCI estimators are proposed.

2.1. Notations and Definitions. Let $\xi = (\xi_{p_1}, \xi_{p_2}, \xi_{p_3})$ be the vector of (raw) sample quantiles. For ease of developing asymptotic results of estimators of the quantile-based PCIs, we reexpress (2) as follows:

$$\begin{aligned} C_{Np}^1(u, v; \xi) &= \frac{d}{3\sqrt{\left((\xi_{p_3} - \xi_{p_1})/6\right)^2 + v(\xi_{p_2} - T)^2}}, \\ &\quad \text{if } \xi_{p_2} = m; \\ C_{Np}^2(u, v; \xi) &= \frac{d - u(\xi_{p_2} - m)}{3\sqrt{\left((\xi_{p_3} - \xi_{p_1})/6\right)^2 + v(\xi_{p_2} - T)^2}}, \\ &\quad \text{if } \xi_{p_2} > m; \\ C_{Np}^3(u, v; \xi) &= \frac{d + u(\xi_{p_2} - m)}{3\sqrt{\left((\xi_{p_3} - \xi_{p_1})/6\right)^2 + v(\xi_{p_2} - T)^2}}, \\ &\quad \text{if } \xi_{p_2} < m. \end{aligned} \quad (3)$$

Let

$$\Delta = \left(\frac{\xi_{p_3} - \xi_{p_1}}{6} \right)^2 + v(\xi_{p_2} - T)^2. \quad (4)$$

Furthermore, we define vectors, for $i = 0, 1, 2$,

$$V_i(u, v; \xi) = \left(\frac{\partial C_{Np}^i(u, v; \xi)}{\partial \xi_{p_1}}, \frac{\partial C_{Np}^i(u, v; \xi)}{\partial \xi_{p_2}}, \frac{\partial C_{Np}^i(u, v; \xi)}{\partial \xi_{p_3}} \right), \quad (5)$$

where

$$\begin{aligned} \frac{\partial C_{Np}^1(u, v; \xi)}{\partial \xi_{p_1}} &= \frac{d(\xi_{p_3} - \xi_{p_1})}{108\Delta^{3/2}}, \\ \frac{\partial C_{Np}^1(u, v; \xi)}{\partial \xi_{p_2}} &= -\frac{dv(\xi_{p_2} - T)}{3\Delta^{3/2}}, \\ \frac{\partial C_{Np}^2(u, v; \xi)}{\partial \xi_{p_1}} &= \frac{[d - u(\xi_{p_2} - m)](\xi_{p_3} - \xi_{p_1})}{108\Delta^{3/2}}, \\ \frac{\partial C_{Np}^2(u, v; \xi)}{\partial \xi_{p_2}} &= -\frac{u\Delta + v[d - u(\xi_{p_2} - m)](\xi_{p_2} - T)}{3\Delta^{3/2}}, \\ \frac{\partial C_{Np}^3(u, v; \xi)}{\partial \xi_{p_1}} &= \frac{[d + u(\xi_{p_2} - m)](\xi_{p_3} - \xi_{p_1})}{108\Delta^{3/2}}, \\ \frac{\partial C_{Np}^3(u, v; \xi)}{\partial \xi_{p_2}} &= \frac{u\Delta - v[d + u(\xi_{p_2} - m)](\xi_{p_2} - T)}{3\Delta^{3/2}}, \\ \frac{\partial C_{Np}^i(u, v; \xi)}{\partial \xi_{p_1}} &= -\frac{\partial C_{Np}^i(u, v; \xi)}{\partial \xi_{p_3}}, \quad \text{for } i = 1, 2, 3. \end{aligned} \quad (6)$$

Finally, denote

$$\begin{aligned} V(u, v; \xi) &= V_1(u, v; \xi)I_{[\xi_{p_2}=m]} \\ &\quad + V_2(u, v; \xi)I_{[\xi_{p_2}>m]}, \end{aligned} \quad (7)$$

where $I_{[\cdot]}$ is an indicator function taking value 1 if the argument is true and 0 otherwise. That is, $V(u, v; \xi)$ is a linear combination of $V_1(u, v; \xi)$, $V_2(u, v; \xi)$, and $V_3(u, v; \xi)$. Note also that $V(u, v; \xi)$ depends only on the form of $C_{Np}(u, v)$.

2.2. Two Estimators of Quantile-Based PCIs. Among all available quantile estimators in the literature, the simplest consistent estimator of a process quantile is the sample quantile. We consider point estimators of the quantile based PCIs using sample quantiles and the refined estimator of quantiles-based on the linear interpolation of sample quantiles.

Let $\hat{\xi} = (\hat{\xi}_{p_1}, \hat{\xi}_{p_2}, \hat{\xi}_{p_3})^\tau$ be the vector of sample quantiles and $f(x)$ be the density function of the underlying process.

It is well known that ξ is a strong consistent estimator of $\xi = (\xi_{p_1}, \xi_{p_2}, \xi_{p_3})^T$. Furthermore,

$$\Gamma_0(f; \xi) = \text{var}(\hat{\xi}) = (\delta_{ij}(f; \xi)),$$

$$\text{with } \delta_{ij}(f; \xi) = \frac{p_i(1-p_j)}{nf(\xi_{p_i})f(\xi_{p_j})}, \quad (8)$$

where $\delta_{ij}(f; \xi)$ is the covariance of $\hat{\xi}_i$ and $\hat{\xi}_j$ (see Serfling [5], 1980, Section 2 for details). The first point estimator of (2) is proposed as follows

$$\hat{C}_{Np}(u, v; \xi) = C_{Np}(u, v; \hat{\xi})$$

$$= \frac{d - u \left| \hat{\xi}_{p_2} - m \right|}{3\sqrt{\left((\hat{\xi}_{p_3} - \hat{\xi}_{p_1})/6 \right)^2 + v(\hat{\xi}_{p_2} - T)^2}}. \quad (9)$$

Pearn and Chen [6] generalized Chang and Lu's [7] interpolation-based quantile estimators and proposed the following estimators (referred to hereafter as *interpolated quantiles*), for $i = 1, 2, 3$,

$$\hat{\xi}_{p_i} = \hat{\xi}_{[R_i]/n} + (R_i - [R_i]) \left(\hat{\xi}_{([R_i]+1)/n} - \hat{\xi}_{[R_i]/n} \right), \quad (10)$$

where $R_i = np_i + 1 - p_i$ and n is the sample size and $[x]$ is the floor function which returns the largest integer less than or equal to x . Let $\hat{\zeta} = (\hat{\zeta}_1, \hat{\zeta}_2, \hat{\zeta}_3)$. Since the interpolation-based quantile, $\hat{\zeta}$, is a linear combination of some (raw) sample quantiles, $\hat{\zeta}$ is a consistent estimator of ζ . The second point estimator of PCI based on interpolated quantiles is proposed with the following form:

$$\hat{C}_{Np}(u, v; \zeta) = C_{Np}(u, v; \hat{\zeta})$$

$$= \frac{d - u \left| \hat{\zeta}_{p_2} - m \right|}{3\sqrt{\left((\hat{\zeta}_{p_3} - \hat{\zeta}_{p_1})/6 \right)^2 + v(\hat{\zeta}_{p_2} - T)^2}}. \quad (11)$$

For $i = 1, 2, 3$, define $q_{2i-1} = [R_i]/n$, $q_{2i} = ([R_i] + 1/n)$, and $w_i = R_i - [R_i]$. Denote

$$Q = \begin{pmatrix} 1 - w_1 & w_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - w_2 & w_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - w_3 & w_3 \end{pmatrix}, \quad (12)$$

and $\rho = (\xi_{q_1}, \xi_{q_2}, \xi_{q_3}, \xi_{q_4}, \xi_{q_5}, \xi_{q_6})^T$, where ξ_{q_i} is q_i th quantile. The sample version of ρ is given by $\hat{\rho} = (\hat{\xi}_{q_1}, \hat{\xi}_{q_2}, \hat{\xi}_{q_3}, \hat{\xi}_{q_4}, \hat{\xi}_{q_5}, \hat{\xi}_{q_6})^T$, where $\hat{\xi}_{q_i}$ is the q_i th sample quantile. With these notation, we have the following linear relationship between $\hat{\zeta}$ and $\hat{\rho}$:

$$\hat{\zeta} = Q\hat{\rho}. \quad (13)$$

Using a similar argument on (8) gives the 6×6 variance-covariance matrix of $\hat{\rho}$ as follows:

$$\Sigma(f; \rho) = \text{var}(\hat{\rho}) = (\sigma_{ij}(f; \rho)),$$

$$\text{with } \sigma_{ij}(f; \rho) = \frac{q_i(1-q_j)}{nf(\xi_{q_i})f(\xi_{q_j})}. \quad (14)$$

Therefore, the variance-covariance matrix of $\hat{\zeta} = (\hat{\zeta}_{p_1}, \hat{\zeta}_{p_2}, \hat{\zeta}_{p_3})^T$ expressed in sample quantiles and the density function of underlying process is given by

$$\Gamma_1(f; \zeta) = \text{var}(\hat{\rho}) = Q^T \Sigma(f; \rho) Q. \quad (15)$$

2.3. Asymptotic Normality of $\hat{C}_{Np}(u, v; \xi)$ and $\hat{C}_{Np}(u, v; \zeta)$. Chen and Hsu [4] have proposed several estimators of $C_{Np}(u, v)$ based on several quantile estimators available in the literature. They also proved the corresponding central limit theorems associated with the proposed estimators. Here we prove a version of central limit theorem regarding $\hat{C}_{Np}(u, v)$ in a similar way but with a simpler form of the asymptotic variance of $\hat{C}_{Np}(u, v)$. Similar to that of Chen and Hsu [4], our variance is dependent on the explicit expression of the density of the process.

Lemma 1. Let $f(y)$ be the density function of the underlying process. Assume further that $f(y)$ is positive and continuous in a neighborhood of p_i for $i = 1, 2, 3$ and q_j for $j = 1, 2, 3, 4, 5, 6$, respectively. Let $(\hat{\xi}_{p_1}, \hat{\xi}_{p_2}, \hat{\xi}_{p_3})^T$ and $(\hat{\xi}_{q_1}, \hat{\xi}_{q_2}, \hat{\xi}_{q_3}, \hat{\xi}_{q_4}, \hat{\xi}_{q_5}, \hat{\xi}_{q_6})^T$ be the vectors of sample quantiles of $(\xi_{p_1}, \xi_{p_2}, \xi_{p_3})^T$ and $(\xi_{q_1}, \xi_{q_2}, \xi_{q_3}, \xi_{q_4}, \xi_{q_5}, \xi_{q_6})^T$, respectively. The vector $(\hat{\zeta}_{p_1}, \hat{\zeta}_{p_2}, \hat{\zeta}_{p_3})^T$ is defined as in previous subsection. Then

$$\sqrt{n} \begin{pmatrix} \hat{\xi}_{p_1} - \xi_{p_1} \\ \hat{\xi}_{p_2} - \xi_{p_2} \\ \hat{\xi}_{p_3} - \xi_{p_3} \end{pmatrix} \rightarrow_d N(0, n\Gamma_0(f; \xi)),$$

$$\sqrt{n} \begin{pmatrix} \hat{\zeta}_{p_1} - \zeta_{p_1} \\ \hat{\zeta}_{p_2} - \zeta_{p_2} \\ \hat{\zeta}_{p_3} - \zeta_{p_3} \end{pmatrix} \rightarrow_d N(0, n\Gamma_1(f; \zeta)), \quad (16)$$

where $\Gamma_0(f; \xi)$ and $\Gamma_1(f; \zeta)$ are specified in (8) and (15).

The proof of this Lemma follows immediately from Theorem B in Serfling [5, page 80]. The following theorem gives the asymptotic distributions of the sample quantile-based estimators $\hat{C}_{Np}(u, v; \xi)$ and $\hat{C}_{Np}(u, v; \zeta)$. The proof will be provided in the appendix.

Theorem 2. Under the assumption in Lemma 1, one has

$$\sqrt{n}(\hat{C}_{Np}(u, v; \xi) - C_{Np}(u, v; \xi)) \rightarrow_d N(0, \Omega_0),$$

$$\sqrt{n}(\hat{C}_{Np}(u, v; \zeta) - C_{Np}(u, v; \zeta)) \rightarrow_d N(0, \Omega_1), \quad (17)$$

where $\Omega_0 = nV(u, v; \xi)\Gamma_0(f; \xi)V^T(u, v; \xi)$, $\Omega_1 = nV(u, v; \zeta)\Gamma_1(f; \zeta)V^T(u, v; \zeta)$, and $V(u, v; \cdot)$ is specified in (7).

Remark 3. The point estimators $\hat{C}_{Np}(u, v; \xi)$ and $\hat{C}_{Np}(u, v; \zeta)$ are asymptotically unbiased estimators of $C_{Np}(u, v)$.

The primary application of Theorem 2 is to construct asymptotic confidence limits of and test hypotheses on PCIs.

We will address how to construct confidence limits of PCIs in the following section.

3. Asymptotic Confidence Intervals

If the density function of the underlying process is known, we can use Theorem 2 to construct the asymptotic confidence limits for $C_{Np}(u, v)$. If the underlying distribution is unknown, we cannot use Theorem 2 directly to construct confidence limits for $C_{Np}(u, v)$. In this section, we first propose a consistent kernel estimator for the density function, $f(x)$, of the underlying process. Then we can estimate the asymptotic variances in Theorem 2 using the sample information.

3.1. A Consistent Variance Estimator for $\hat{C}_{Np}(u, v)$. Suppose that $\{X_1, X_2, \dots, X_n\}$ is a random sample collected from the underlying process with density $f(x)$. The kernel density estimator of $f(x)$, denoted by $\hat{f}_n(x)$, is given by

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (18)$$

where $K(\cdot)$ is the kernel function which is nonnegative, unimodal, and symmetric with respect to the vertical axis and integrates to unity. h is the bandwidth which controls the degree of smoothing of the estimation. Among several commonly used kernel functions, Gaussian kernels are the most often used. The quality of a kernel estimate depends less on the shape of the K than on the value of its bandwidth h . In this paper, we will use Gaussian kernel and the rule of thumb of Silverman [8] for choosing the bandwidth, which yields a mean integrated square error within 10% of the optimum for all the t distributions, of a Gaussian kernel density estimator.

The kernel density estimator has many desirable properties. It is asymptotically unbiased, consistent in a mean-square sense, and uniformly consistent in probability. For references Silverman [8] and Wand and Jones [9] are the excellent monographs. Bowman and Azzalini [10] is a nice nontechnical introduction to kernel density estimation.

Using the kernel estimator of the density as well as the quantile estimators quantile-based PCIs for the underlying process, we propose the following variance estimators:

$$\begin{aligned} \widehat{\text{var}}(\hat{\Omega}_0) &= \hat{V}(u, v; \hat{\xi}) \Gamma_0(\hat{f}; \hat{\xi}) \hat{V}^T(u, v; \hat{\xi}), \\ \widehat{\text{var}}(\hat{\Omega}_1) &= \hat{V}(u, v; \hat{\zeta}) \Gamma_1(\hat{f}; \hat{\zeta}) \hat{V}^T(u, v; \hat{\zeta}). \end{aligned} \quad (19)$$

Since both variances are continuous function of the estimated density function $f(x)$ and the estimated process

quantiles, the consistency of the the two kernel density estimator-based variance estimators $\widehat{\text{var}}(\hat{\Omega}_0)$ and $\widehat{\text{var}}(\hat{\Omega}_1)$ follows immediately from the fact that the $\hat{f}(x)$ is (strongly) consistent with $f(x)$.

3.2. Asymptotic Confidence Interval. With the kernel estimator of the density of the underlying process, we define the asymptotic confidence interval for the quantile-based PCIs. Let $\hat{\xi}_\alpha$ be the α th sample percentile and define

$$\begin{aligned} \widehat{\text{s.e.}}[\hat{C}_{Np}(u, v; \xi)] &= \sqrt{\widehat{\text{var}}[\hat{C}_{Np}(u, v; \xi)]}, \\ \widehat{\text{s.e.}}[\hat{C}_{Np}(u, v; \zeta)] &= \sqrt{\widehat{\text{var}}[\hat{C}_{Np}(u, v; \zeta)]}. \end{aligned} \quad (20)$$

The $100(1 - \alpha)\%$ lower confidence limits of $C_{Np}(u, v)$ using different quantile-based estimators are given by

$$\begin{aligned} &(\hat{C}_{Np}(u, v; \xi) - z_\alpha \widehat{\text{s.e.}}[\hat{C}_{Np}(u, v; \xi)], +\infty), \\ &(\hat{C}_{Np}(u, v; \zeta) - z_\alpha \widehat{\text{s.e.}}[\hat{C}_{Np}(u, v; \zeta)], +\infty). \end{aligned} \quad (21)$$

Since the above standard errors are explicitly expressed in data measurements and the estimated density function, the confidence limits can be calculated through simple programming.

4. Numerical Results

We now present some numerical results in this section to apply the proposed procedures. The first example is based on real-life data. We also provide a simulation study to compare the accuracy of the two estimators of quantile functions and the impact on the estimators of PCIs.

4.1. An Illustrative Example. The rubber-edge is one of the key components in audio speaker manufacturing that reflect sound quality of the driver units and clarity of the sound. As an illustrative example, we use the data given Chen and Pearn [3]. One characteristic of the rubber edge is the weight. The upper and lower specification limits, USL and LSL, of the weight for a particular model of rubber edge are set at 8.94 and 8.46 (in grams). The target value is the midpoint between LSL and USL, which is 8.70 grams.

In order to assess the capability of the process, Chen and Pearn [3] collected following random sample with 100 data measurements (see also Table IV of Chen and Pearn [3]).

8.61	8.64	8.53	8.76	8.68	8.55	8.66	8.67	8.76	8.71	8.81	8.68	8.74	8.71	8.55
8.71	8.72	8.71	8.64	8.81	8.72	8.98	8.59	8.71	8.73	8.74	8.81	8.63	8.54	8.60
8.69	8.70	8.69	8.67	8.67	8.70	8.63	8.74	8.71	8.64	8.65	8.74	8.70	8.67	8.71
8.62	8.78	8.67	8.69	8.71	8.64	8.75	9.03	8.68	8.73	8.61	8.64	8.69	8.80	8.75
8.68	8.66	8.83	8.69	8.67	8.79	8.66	8.69	8.70	8.67	8.74	9.00	8.87	8.74	8.68
8.69	8.63	8.68	8.59	8.73	8.68	8.64	8.79	8.80	8.69	8.68	8.71	8.70	8.53	8.61
8.67	8.70	8.68	8.59	8.74	8.77	8.99	8.81	8.74	8.84

FIGURE 1: (a) and (b) are histograms with overlaid nonparametric kernel density curves using Silverman's [8] rule-of-thumb bandwidth selection. (c) and (d) are control charts.

There are four measurements which are above the upper specification limit $USL = 8.94$ (see Figure 1) and the data also failed to pass the normality test. Chen and Pearn [3] used quantile-based PCI values and claimed that the process was incapable at the time of data collection.

After some adjustments were performed on the process, another 100 data measurements were sampled (see also Table V of Chen and Pearn [3]).

8.70	8.65	8.72	8.80	8.94	8.65	8.63	8.67	8.66	8.69	8.69	8.66	8.80	8.70	8.68
8.74	8.70	8.68	8.69	8.65	8.71	8.69	8.72	8.52	8.62	8.75	8.92	8.69	8.68	8.93
8.70	8.71	8.94	8.65	8.70	8.69	8.71	8.67	8.69	8.64	8.66	8.69	8.81	8.73	8.69
8.70	8.67	8.69	8.68	8.67	8.67	8.71	8.67	8.70	8.66	8.70	8.62	8.69	8.73	8.64
8.68	8.68	8.74	8.55	8.70	8.56	8.68	8.52	8.73	8.68	8.73	8.87	8.71	8.76	8.81
8.67	8.70	8.65	8.67	8.77	8.66	8.70	8.75	8.73	8.69	8.71	8.64	8.70	8.83	8.64
8.72	8.69	8.73	8.71	8.72	8.64	8.70	8.69	8.71	8.81					

There are still two measurements which are equal to the upper specification limit. Further more, since the underlying process failed passing normality tests, Chen and Pearn [3] reported the quantile-based PCIs again and found that PCIs are bigger than 1 which implies that the capability of the underlying process was improved.

In this analysis, we report the point estimates of quantile-based PCIs and the corresponding standard error (s.e.) and 95% lower confidence limit (LCL). The results are summarized in Table 1.

We can see from Table 1 that the standard errors of point estimates of PCIs based on sample quantiles are uniformly

TABLE 1: Point estimates and 95% asymptotic lower confidence limits of various $C_{Np}(u, v)$.

Process	PCI Type	Sample quantile ξ		Interpolation-based quantile ζ			
		LCL	$\hat{C}_{Np}(\xi)$	s.e.	LCL	$\hat{C}_{Np}(\zeta)$	s.e.
Before Adjust.	$C_{Np}(0, 0)$	0.911466	0.960000	0.02476212	0.918986	0.967761	0.02488492
	$C_{Np}(1, 0)$	0.846163	0.920000	0.03767177	0.849286	0.927437	0.03987269
	$C_{Np}(0, 1)$	0.902110	0.953162	0.02604660	0.907308	0.960756	0.02726949
	$C_{Np}(1, 1)$	0.824969	0.913447	0.04514164	0.827334	0.920725	0.04764796
After Adjust.	$C_{Np}(0, 0)$	1.104678	1.142857	0.01947940	1.056493	1.142857	0.04406315
	$C_{Np}(1, 0)$	1.038113	1.095238	0.02914537	0.997961	1.095238	0.04963102
	$C_{Np}(0, 1)$	1.088776	1.131371	0.02173196	1.043063	1.131371	0.04505517
	$C_{Np}(1, 1)$	1.011710	1.084230	0.03700042	0.977467	1.084230	0.05447085

smaller than those of PCIs-based on interpolated quantiles. The point estimates of PCIs based sample and interpolated quantiles happened to be the same after the process was adjusted.

4.2. A Monte Carlo Simulation Study. In this simulation, we assume the underlying process follows a Weibull distribution with density function

$$f(x) = \frac{\beta}{\theta} \left(\frac{x}{\theta}\right)^{\beta-1} \exp \left[-\left(\frac{x}{\theta}\right)^{\beta} \right], \quad x \geq 0, \quad (22)$$

where θ is the scale parameter and β is a shape parameter. The theoretical α th quantile of this Weibull process, denoted by ξ_{α} , is implicitly determined by

$$\alpha = 1 - \exp \left[-\left(\frac{\xi_{\alpha}}{\theta}\right)^{\beta} \right], \quad (23)$$

that is, the α th quantile of the Weibull process is explicitly expressed by

$$\xi_{\alpha} = \theta [-\ln(1 - \alpha)]^{1/\beta}. \quad (24)$$

The fundamental goal of this simulation is to compare the estimators of PCIs and the corresponding kernel density-based asymptotic variance defined based on sample quantiles with that defined based on interpolated quantiles using mean square error (MSE) and examine how sample size and shape of the process distribution affect the accuracy of PCI estimators and their associated variance estimators as well. To this end, we first choose fixed values of shape and scale parameter, say $\beta = 3.2$ and $\theta = 7$. The specification limits and the target value are chosen to be LSL = 0.30, USL = 12.7, $T = 6.5$. With these choices of quality characteristics, the true PCIs are $C_{Np}(0, 0) = 1.056$, $C_{Np}(1, 0) = 1.012$, $C_{Np}(0, 1) = 1.047$, and $C_{Np}(1, 1) = 1.004$. The sample sizes we used are $n = 50, 100, 150, 200, 400$, and 1000 . For each sample size, we simulated 1000 random samples from the Weibull population and evaluate the true PCIs and the corresponding asymptotic variance (in Theorem 2) according to the sample size. The simulated mean square errors (MSE) of the point estimate of PCIs and their corresponding asymptotic variances based on raw sample quantiles (ξ) and interpolated quantiles (ζ), respectively, in Table 2.

From Table 2, we can see that the (true value) asymptotic variance of PCIs decreases as the sample size increases; the MSEs of $\hat{C}_{Np}(\xi)$ and $\text{var}_n[\hat{C}_{Np}(\xi)]$ are uniformly greater than that of $\hat{C}_{Np}(\zeta)$ and $\text{var}_n[\hat{C}_{Np}(\zeta)]$, respectively; the MSEs of (raw) sample and interpolated quantile-based point estimators of PCIs and their corresponding asymptotic variances decreases as the sample size increases. Based on these observations it turns out that the raw sample quantile-based PCIs have better asymptotic performance than the interpolated quantile-based ones.

Figure 2 gives the simulated density curves of quantile-based PCIs under different sample sizes. It is seen that all curves with sample size as small as 50 are bell shaped and similar to the density curves of normal random variables which indicate the validity of asymptotic inference.

Next we examine how the change of shape parameter (related to skewness) in the Weibull process affects the biases of the point estimators of PCIs defined using sample quantile and interpolated quantiles. We still use the same LSL, USL and the process target as those used in the previous simulation. The values of shape parameter are chosen in the sequence from 2.0 to 5.0 with increment 0.2. Since Figure 2 indicates that the asymptotic normal approximation is reasonably well with sample size as small as 50, we choose the fixed sample size 50 to simulate the biases using different values shape parameter specified in the above sequence.

It turns out, from Figure 3, that biases of the estimators of PCIs based on interpolated quantiles are greater than those defined based on raw sample quantiles across the various values of the shape parameter. But the significance of the differences of bias needs to be rigorously tested.

The analysis conducted in this section is carried out using the free open source statistical package R. The code can also be run in the commercial package S-PLUS. The program is available on request.

5. Discussion and Summary

The formulation of C_p , C_{pk} , C_{pm} , and C_{pmk} using underlying process mean and standard deviation implicitly assumed the certain degree of normality which limits the application of PCIs in many skewed processes. The quantile-based basic PCIs of Chen and Pearn [3] eliminate the implicit normality

TABLE 2: Simulation results the performance of various 95% interval estimates of $C_{Np}(u, v)$ based sample quantile and interpolated quantile using Weibull process.

Size (n)	PCI Type	Asymptotic	Sample quantile		Interpolated quantile	
		PCI Var.	MSE [$\hat{C}_{Np}(\xi)$]	var _n [$\hat{C}_{Np}(\xi)$]	MSE [$\hat{C}_{Np}(\zeta)$]	var _n [$\hat{C}_{Np}(\zeta)$]
50	$C_{Np}(0, 0)$	$2.78004e - 07$	0.1224934	$6.833435e - 06$	0.1281274	0.0003281428
	$C_{Np}(1, 0)$	$4.04514e - 06$	0.09835884	0.0001175441	0.0991082	0.0005820965
	$C_{Np}(0, 1)$	$9.06451e - 07$	0.07808902	0.001147389	0.0874771	0.001795696
	$C_{Np}(1, 1)$	$7.43255e - 06$	0.08175114	0.0001861944	0.0976206	0.0005729443
100	$C_{Np}(0, 0)$	$1.39002e - 07$	0.04463229	$6.763005e - 06$	0.0518551	$4.923725e - 05$
	$C_{Np}(1, 0)$	$2.02257e - 06$	0.03884461	$2.997771e - 05$	0.0470966	0.0001031592
	$C_{Np}(0, 1)$	$4.53225e - 07$	0.03483475	0.0001387058	0.0438483	0.0002384231
	$C_{Np}(1, 1)$	$3.71628e - 06$	0.03572435	$1.961133e - 05$	0.0426112	$7.817404e - 05$
200	$C_{Np}(0, 0)$	$6.95009e - 08$	0.01623415	$8.483628e - 06$	0.0205617	$8.252656e - 06$
	$C_{Np}(1, 0)$	$1.01129e - 06$	0.01489944	$1.519576e - 05$	0.0197595	$1.852167e - 05$
	$C_{Np}(0, 1)$	$2.26613e - 07$	0.01422966	$3.028482e - 05$	0.0177817	$3.590338e - 05$
	$C_{Np}(1, 1)$	$1.85814e - 06$	0.01422652	$9.677983e - 06$	0.0192691	$1.137201e - 05$
400	$C_{Np}(0, 0)$	$9.26679e - 08$	0.00598806	$1.319295e - 05$	0.0091714	$2.284677e - 06$
	$C_{Np}(1, 0)$	$1.34838e - 06$	0.00586153	$1.430459e - 05$	0.0077462	$4.224512e - 06$
	$C_{Np}(0, 1)$	$3.02150e - 07$	0.00552098	$1.616761e - 05$	0.0076277	$7.120676e - 06$
	$C_{Np}(1, 1)$	$2.47752e - 06$	0.00560132	$1.155253e - 05$	0.0076023	$2.451394e - 06$
1000	$C_{Np}(0, 0)$	$6.95009e - 08$	0.00264029	$3.57955e - 06$	0.0029580	$1.819142e - 06$
	$C_{Np}(1, 0)$	$1.01129e - 06$	0.00256571	$4.026816e - 06$	0.0029113	$2.016832e - 06$
	$C_{Np}(0, 1)$	$2.26613e - 07$	0.00258448	$4.102949e - 06$	0.0028118	$2.470318e - 06$
	$C_{Np}(1, 1)$	$1.85814e - 06$	0.00244308	$3.640072e - 06$	0.0027096	$1.711628e - 06$

assumption and therefore can be used essentially in any processes. The asymptotic properties of quantile-based PCIs with superstructure have been studied by Chen and Hsu [4]. Since the asymptotic variance of quantile-based superstructure contains the underlying process density distribution function, the lower confidence limits of quantile-based PCIs cannot be constructed without knowing the density function of the underlying process.

In this paper, we use nonparametric kernel density estimation approach to estimate the underlying density function which allows to propose a consistent estimator of the asymptotic variance of the quantile-based PCIs. We also use two different quantile estimators, namely-raw sample quantile and interpolated quantile estimator originally proposed by Chen and Pearn [3]. Simulation study based on Weibull process indicates that, in terms of mean square error, the performance raw sample quantile-based PCIs are uniformly better than interpolated quantile-based PCIs.

There are at least 9 different quantile estimators available in different software packages. The detailed definitions of these quantile estimators can be found in Hyndman and Fan [11]. It is of interest to examine performance of quantile-based PCIs defined based on these different quantile estimators.

Another interesting issue is to compare the performance of process mean standard deviation-based superstructure of Vännman [2] and the quantile-based superstructure of Chen and Pearn [3] via simulation using both normal and nonnormal processes.

Appendix

Proof of Theorem 2. For the asymptotic normality of $\hat{C}_{Np}(u, v; \xi)$, we have the following approximation using Taylor expansion

$$\begin{aligned}
 \hat{C}_{Np}(u, v; \xi) &= C_{Np}(u, v; \xi) \\
 &+ \sum_{i=1}^3 \frac{\partial C_{Np}(u, v; \xi)}{\partial \xi_{p_i}} (\hat{\xi}_{p_i} - \xi_{p_i}) + o_p(\psi_n^0) \\
 &= C_{Np}(u, v; \xi) \\
 &+ V(u, v; \xi) \begin{pmatrix} \hat{\xi}_{p_1} - \xi_{p_1} \\ \hat{\xi}_{p_2} - \xi_{p_2} \\ \hat{\xi}_{p_3} - \xi_{p_3} \end{pmatrix} + o_p(\psi_n^0),
 \end{aligned} \tag{A.1}$$

where $\psi_n^0 = \sum_{i=1}^3 \|\hat{\xi}_{p_i} - \xi_{p_i}\|$ is the Euclidean distance. The $\sqrt{n}$ -consistency of $(\hat{\xi}_{p_1}, \hat{\xi}_{p_2}, \hat{\xi}_{p_3})$ from Lemma 1 implies that $\sqrt{n}o_p(\psi_n) = \sqrt{n}o_p(n^{-1/2}) = o_p(1)$. Therefore,

$$\begin{aligned}
 &\sqrt{n}(\hat{C}_{Np}(u, v; \xi) - C_{Np}(u, v; \xi)) \\
 &= V(u, v; \xi) \sqrt{n} \begin{pmatrix} \hat{\xi}_{p_1} - \xi_{p_1} \\ \hat{\xi}_{p_2} - \xi_{p_2} \\ \hat{\xi}_{p_3} - \xi_{p_3} \end{pmatrix} + o_p(1).
 \end{aligned} \tag{A.2}$$

FIGURE 2: Simulated kernel density curves of various PCIs using different sample sizes based on Weibull process with shape parameter $\beta = 3.2$ and scale parameter $\theta = 7$.

A direct application of Slutsky's theorem gives

$$\begin{aligned} & \sqrt{n}(\hat{C}_{Np}(u, v; \xi) - C_{Np}(u, v; \xi)) \\ & \rightarrow_d N(0, nV(u, v; \xi)\Gamma_0(\hat{f}; \hat{\xi})V^T(u, v; \xi)), \end{aligned} \quad (A.3)$$

where $\Gamma_0(\hat{f}, \hat{\xi})$ is specified in (8). For the asymptotic normality of $\hat{C}_{Np}(u, v; \xi)$, we have

$$\hat{C}_{Np}(u, v; \xi) = C_{Np}(u, v)$$

$$+ \sum_{i=1}^3 \frac{\partial C_{Np}(u, v; \xi)}{\partial \zeta_{p_i}} (\hat{\zeta}_{p_i} - \zeta_{p_i}) + o_p(\psi_n^0)$$

$$= C_{Np}(u, v; \xi)$$

$$+ V(u, v; \xi) \begin{pmatrix} \hat{\zeta}_{p_1} - \zeta_{p_1} \\ \hat{\zeta}_{p_2} - \zeta_{p_2} \\ \hat{\zeta}_{p_3} - \zeta_{p_3} \end{pmatrix} + o_p(\psi_n^0)$$

$$= C_{Np}(u, v; \xi) + V(u, v; \xi)Q \cdot$$

$$(\hat{\rho} - \rho) + o_p(\psi_n^0).$$

(A.4)

Note that

$$\sqrt{n}(\hat{\rho} - \rho) \rightarrow_d N(0, \text{var}(\hat{\rho})), \quad (A.5)$$

FIGURE 3: Comparing the bias of estimators of PCIs defined based on sample quantiles and interpolated quantiles respectively. The solid line represents sample quantile and the dashed line represents the interpolated quantile.

where $\text{var}(\hat{\rho})$ is specified in (14). Similar to the arguments used in proving the normality of $\hat{C}_{Np}(u, v; \xi)$, we have

$$\begin{aligned} & \sqrt{n}(\hat{C}_{Np}(u, v; \xi) - C_{Np}(u, v; \xi)) \\ & \rightarrow_d N(0, nV(u, v; \xi)\Gamma_1(f; \xi)V^T(u, v; \xi)), \end{aligned} \quad (\text{A.6})$$

where $\Gamma_1(f, \xi)$ is specified in (15). The proof of Theorem 2 is complete. $\square$

Acknowledgments

The authors are grateful to the valuable comments from an anonymous reviewer and the Editor that improved the presentation of the this paper. The research of J. Xu was supported by the Natural Science Foundation of the City of Ningbo, China, under Contract 2011A610166.

References

- [1] J. M. Juran, F. M. Gryna, and R. S. Bingham, *Quality Control Handbook*, McGraw-Hill, New York, NY, USA, 1974.
- [2] K. Vännman, "A unified approach to process capability indices," *Statistica Sinica*, vol. 5, pp. 805–820, 1995.
- [3] K. S. Chen and W. L. Pearn, "An application of non-normal process capability indices," *Quality and Reliability Engineering International*, vol. 13, no. 6, pp. 355–360, 1997.
- [4] S.-M. Chen and Y.-S. Hsu, "Asymptotic analysis of estimators for C_{Np} based on quantile estimators," *Journal of Nonparametric Statistics*, vol. 15, no. 2, pp. 137–150, 2003.
- [5] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York, NY, USA, 1980.
- [6] W. L. Pearn and K. S. Chen, "Capability indices for non-normal distributions with an application in electrolytic capacitor manufacturing," *Microelectronics Reliability*, vol. 37, no. 12, pp. 1853–1858, 1997.
- [7] P. L. Chang and K. H. Lu, "PCI calculations for any shape of distribution with percentile," *Quality World, Technical Section*, pp. 110–114, 1994.
- [8] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, UK, 1986.
- [9] M. Wand and M. Jones, *Kernel Smoothing*, Chapman and Hall, London, UK, 1995.
- [10] A. Bowman and A. Azzalini, *Applied Smoothing Techniques*, Oxford University Press, Oxford, UK, 1997.

- [11] R. J. Hyndman and Y. Fan, "Sample quantiles in statistical packages," *American Statistician*, vol. 50, no. 4, pp. 361–365, 1996.