$D_{p,q}$-DISTANCE AND ITS APPLICATION FOR CONSTRUCTING CONFIDENCE INTERVAL FOR C_{pm} BASED ON FUZZY DATA

B. SADEGHPOUR GILDEH , S. ASGHARI

ABSTRACT. A measurement control system ensures that measuring equipment and measurement processes are fit for their intended use and its importance in achieving product quality objectives. In most real life applications, the observations are fuzzy. In this paper, first we use the $D_{p,q}$ -distance for the point estimation of process capability index C_{pm} based on fuzzy data and use this estimation for constructing confidence interval for C_{pm} and present a membership function for $\tilde{C}_{pm}$. A numerical example has presented for showing the sensitivity of this new approach with respect to the ambiguity of target value.

1. Introduction

Statistical techniques can be helpful throughout the product cycle, including activities prior to manufacturing, in quantifying process variability, in analyzing this variability relative to product requirements or specifications, and in assisting development and manufacturing in eliminating or greatly reducing this variability. This general activity is called process capability analysis. Process capability refers to the uniformity of process. Obviously, the variability in the process is a measure of the uniformity of output. There may not exist a definition of the ''process capability'' but in high probability the (real valued) quality characteristic X of the produced items lies between some lower and upper specification limits LSL and USL (or tolerance interval limits). Therefore the idea of process capability implies that the fraction p of produced nonconforming items should be small if the process is said to be capable.
In the traditional quality management, the most commonly used capability assumption is that output process measurements are distributed as normal random variables. Experience shows that the normality assumption is often not met in real world. Application and observations usually contain fuzziness owing imprecise measurements or described by linguistic variables, such 'about 7', 'somewhere between 6 and 8' and so on. Chen, Lai and Nien [1] use the fuzzy analytic method concerning process capability index C_{pm} and calculate $\tilde{C}_{pm}$ for fuzzy observation. Perakis and Xekalaki [3] construct confidence interval for the index C_{pm} with crisp data. In this paper, by using fuzzy data, estimate index $\tilde{C}_{pm}$ and construct confidence interval based on the method presented by [3] . Parchami et al [2] obtained fuzzy confidence interval for a fuzzy process capability index . The organization of this paper is as follows. In section 2 we recall some notions of fuzzy number used in this paper. Section 3 contain the traditional definitions of process capability indices. Sections 4 and 5 assign to the presentation of point and interval estimations for C_{pm} based on fuzzy data. At last present a numerical example.

2. Preliminary Notes

2.1.Definition. The $D_{p,q}$-distance, indexed by parameter $1< p <\infty$, $0 \le q \le 1$, between two fuzzy numbers $\tilde{A}$ and $\tilde{B}$ is a nonnegative function on F(R) × F(R) gives as follows:

$$D_{p,q}(\tilde{A},\tilde{B})=\begin{cases}[(1-q)\int_0^1\left|A_\alpha^- - B_\alpha^-\right|^p d\alpha + q\int_0^1\left|A_\alpha^+ - B_\alpha^+\right|^p d\alpha]^{\frac{1}{p}} & if\ p<\infty\\ (1-q)\sup\limits_{0<\alpha\le 1}\left(\left|A_\alpha^- - B_\alpha^-\right|\right)+q\inf\limits_{0<\alpha\le 1}\left(\left|A_\alpha^+ - B_\alpha^+\right|\right) & if\ p=\infty.\end{cases}$$

The analytical properties of $D_{p,q}$ depend on the first parameter p, while the second parameter q of $D_{p,q}$ characterizes the subjective weight attributed to the sides of the fuzzy numbers. If there are no reason to distinguish any side of fuzzy numbers, $D_{p,\frac{1}{2}}$ is recommended.

2.2.Definition. A mapping $\tilde{X}:\Omega\to F$ (r) is said to be a fuzzy random variable associated with (Ω, A) if and

Key words and phrases: Dp,q-distance; Process capability index; Triangular fuzzy number; Fuzzy random variable.

only if

$$(\omega, x) : x \in X_\alpha(\omega) \in A \times B,$$

where B denote the σ -field of Borel set in R.

2.3.Definition. The central $D_{2,q}$ -mean square dispersion of $\widetilde{X}$ about $\widetilde{E}(\widetilde{X})$ (or $\widetilde{\mu}_{\widetilde{X}}$) is called $DVAR(\widetilde{X})$ given by the value (if it exists)

$$D\operatorname{var}(\widetilde{X}) = E([D_{2,q}(\widetilde{X}, \widetilde{\mu}_{\widetilde{X}})]^2) = \int_\Omega [(1-q)\int_0^1 (X_\alpha^-(w) - (\mu_{\widetilde{X}})_\alpha^-)^2 d\alpha \ + q\int_0^1 (X_\alpha^+(w) - (\mu_{\widetilde{X}})_\alpha^+)^2 d\alpha] dp(w).$$

Assume that $\widetilde{A}$ and $\widetilde{B}$ are triangular fuzzy numbers: $\widetilde{A} = tri(a_1, a_2, a_3)$ and $\widetilde{B} = tri(b_1, b_2, b_3)$, the α -cuts of $\widetilde{A}$ and $\widetilde{B}$ are as follows,

$$A_\alpha = [(1-\alpha)a_1 + a_2\alpha, a_3\alpha + (1-\alpha)a_4],$$
$$B_\alpha = [(1-\alpha)b_1 + b_2\alpha, b_3\alpha + (1-\alpha)b_4].$$

It can establish that

$$[D_{2,\frac{1}{2}}(\widetilde{A}, \widetilde{B})]^2 = \frac{1}{6}[(b_1 - a_1)^2 + 2(b_2 - a_2)^2 + (b_3 - a_3)^2 + (b_1 - a_1)(b_2 - a_2) + (b_3 - a_3)(b_2 - a_2)]. \tag{2.1}$$

3. Traditional Process Capability Indices

One of the proposed definitions on process capability index consider that as the ratio of the real performance of process to requested performance, that is,

$$C_p = \frac{USL - LSL}{6\sigma}.$$

In order to reflect departures from the target value $(M = \dfrac{USL + LSL}{2})$ as well as changes in the process variation several order indices have been proposed such as C_{pk} and C_{pm} given as

$$C_{pk} = C_p(1-k), k = \frac{|\mu - M|}{\frac{USL - LSL}{2}},$$

and

$$C_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - M)^2}},$$

where μ is the distribution center of characteristic X .

C_{pk} measures the distance between the process mean and the closest specification limit relation to the one-side actual process spread 3σ . Departures from the target value carry more weight with the other well-known capability index C_{pm} . In principal, C_{pm} behaved like C_{pk} but C_{pm} is bounded above as $\sigma \to 0$ and $\mu \neq M$. For $\mu = M$ it holds $C_p = C_{pk} = C_{pm}$.

4. The Point Estimation Of Cpm Based On Fuzzy Data

When the observations are crisp, a natural estimator of process capability index C_{pm} can be calculated as the following

$$\hat{C}_{pm}=\frac{USL-LSL}{6\sqrt{S^2+(\overline{X}-T)^2}}, \tag{4.1}$$

where $\overline{X}=\sum_{i=1}^{n}\frac{X_i}{n}$ and $S=\sqrt{\sum_{i=1}^{n}\frac{(X_i-\overline{X})^2}{n-1}}$ are the estimators of μ and σ respectively. For simplicity, when the observations are imprecise, assume that all the observations are triangular fuzzy numbers defined as

$$\tilde{X}_i = tri(X_{il}, X_{im}, X_{iu}),\ for\ \ i=1,2,...,n. \tag{4.2}$$

Suppose that the target T be a triangular fuzzy number defined as $\tilde{T}_i = tri(T_{il}, T_{im}, T_{iu})$. The sample mean $\overline{X}$ and sample variance S^2 are substituted by statistics $\tilde{\tilde{X}}$ and $D\,\text{var}(\tilde{X})$ respectively.

We estimate $D\,\text{var}$ by $\hat{D}\,\text{var}(\tilde{X})=\frac{\sum_{i=1}^{n}[D_{2,1/2}(\tilde{X}_i,\tilde{\tilde{X}})]^2}{n}$ or $DS^2=\frac{\sum_{i=1}^{n}[D_{2,1/2}(\tilde{X}_i,\tilde{\tilde{X}})]^2}{n-1}$, and we have

$$\hat{C}_{\tilde{p}m}=\frac{USL-LSL}{6\sqrt{DS^2+[D_{2,1/2}(\tilde{\tilde{X}},\tilde{T})]^2}}, \tag{4.3}$$

When all observations and target are degenerated to crisp values, equation (4.3) become identical to equation (4.1) of the classical model.

5. The Interval Estimation Of Cpm Based On Fuzzy Data

The statistic $\sum_{i=1}^{n}\frac{[D_{2,1/2}(\tilde{X}_i,\tilde{T})]^2}{D\,\text{var}}$ is distributed as the non-central chi-square with n degrees of freedom and Non-centrality parameter $n\delta$ where $\delta=\frac{[D_{2,1/2}(\tilde{\mu},\tilde{T})]^2}{D\text{var}}$ [5]. Therefore, it follows that

$$\text{p}\left(\chi'^2_{n,\alpha/2}(n\delta)<\sum_{i=1}^{n}\frac{[D_{2,\frac{1}{2}}(\tilde{X}_i,\tilde{T})]^2}{D\,\text{var}}<\chi'^2_{n,1-\alpha/2}(n\delta)\right)=1-\alpha\ , \tag{5.1}$$

where $\chi'^2_{n,\alpha/2}(n\delta)$ denotes the $100\alpha\%$ percentile of the non-central chi-square distribution with n degrees of freedom and non-centrality parameter $n\delta$. In particular, Pearson [4] considered an improvement of Patnaik's approximation by which the non-central chi-square distribution with v degrees of freedom and non-centrality parameter λ is approximated by a distribution of the form $c\chi^2_f+b$, where χ^2_f denotes the chi-square distribution with f degrees of freedom and c, f and b are some constants. The values of the constants c, f and b are obtained by equating the first three moments of the non-central chi-square distribution and $c\chi^2_f+b$. Then, the relation (5.1) can be approximated by

$$\text{p}\left(c\chi^2_{f,\alpha/2}+b<\sum_{i=1}^{n}\frac{[D_{2,\frac{1}{2}}(\tilde{X}_i,\tilde{T})]^2}{D\,\text{var}}<c\chi^2_{f,1-\alpha/2}+b\right)=1-\alpha\ , \tag{5.2}$$

it can be found that the appropriate values of c, f and b are given by $\frac{v+3\lambda}{v+2\lambda}$, $\frac{(v+2\lambda)^3}{(v+3\lambda)^2}$ and $\frac{-\lambda}{(v+3\lambda)}$, respectively. Thus the values of c, f and b can be simplified to $c=\frac{1+3\delta}{1+2\delta}$, $f=\frac{n(1+2\delta)}{C^2}$, $b=-\frac{n\delta^2}{1+3\delta}$ respectively. After some algebra, a $100(1-\alpha)\%$ approximate confidence interval for $C_{\tilde{p}m}$ given by

$$\left(\hat{c}_{\tilde{p}m}\sqrt{\frac{\hat{c}\chi^2_{\hat{f},\alpha/2}+\hat{b}}{n(1+\hat{\delta})}},\hat{c}_{\tilde{p}m}\sqrt{\frac{\hat{c}\chi^2_{\hat{f},1-\alpha/2}+\hat{b}}{n(1+\hat{\delta})}}\right).$$

Where $\hat{C}, \hat{f}$ and $\hat{b}$ are obtained substituting $\hat{\delta}$ for δ. Here, $\hat{\delta}$ Can be either $\hat{\delta}_1=\frac{\left[D_{2,\frac{1}{2}}(\tilde{\bar{X}},\tilde{T})\right]^2}{\hat{D}\,\text{var}}$ or

$$\hat{\delta}_2=\frac{\left[D_{2,\frac{1}{2}}(\tilde{\bar{X}},\tilde{T})\right]^2}{DS^2}.$$

To gain a better understanding of C_{pm} based on fuzzy data, we propose to put this confidence interval as a α -cut for $C_{\tilde{p}m}$ and plot the upper and lower limits of the interval estimation for different values of α. This method gives a membership function for $\tilde{C}_{pm}$.

6. A Numerical Example

Table I shows the Fuzzy data given in [1]. In this example, $USL=6.4$, $LSL=5.5$, $M=5.95$ and $T=6$. Assume that target is a triangular fuzzy number as follows: $tri(6\ 6\ 6)$. The value of $\hat{C}_{\tilde{p}m}$ using the above method is 0.6869018. Table II present $100(1-\alpha)\%$ approximate confidence interval limits for target $T=tri(6\ 6\ 6)$. Figure 1 shows the membership function for $\tilde{C}_{pm}$ that its membership grade in 0.7239642 is one (by using δ_1). The obtained membership function for $\tilde{C}_{pm}$ based on the method presented by [1] has membership grade one in 0.7657. Table III shows the sensitivity of index $\hat{C}_{\tilde{p}m}$ with respect to increasing of ambiguity on target. The mean of observations is triangular fuzzy number (5.843667 5.992333 6.129333). When the target tends to observations mean the value of $\hat{C}_{\tilde{p}m}$ will be large. If target is equal to observations mean, $\hat{C}_{\tilde{p}m}$ = 0.7428751.

TABLE I. 30 triangular fuzzy observation

$\tilde{X}_1$=[5.85 6.15 6.35]	$\tilde{X}_{11}$= [5.86 6.04 6.25]	$\tilde{X}_{21}$=[5.5 5.81 5.99]
$\tilde{X}_2$=[5.79 5.9 5.98]	$\tilde{X}_{12}$=[6.13 6.23 6.33]	$\tilde{X}_{22}$=[5.6 5.92 6.05]
$\tilde{X}_3$= [5.71 5.83 5.99]	$\tilde{X}_{13}$=[5.95 6.05 6.19]	$\tilde{X}_{23}$=[5.5 5.75 5.95]
$\tilde{X}_4$= [6.05 6.18 6.32]	$\tilde{X}_{14}$=[5.06 5.65 5.70]	$\tilde{X}_{24}$=[5.84 6.03 6.15]
$\tilde{X}_5$= [5.89 6.06 6.23]	$\tilde{X}_{15}$=[5.65 5.74 5.84]	$\tilde{X}_{25}$=[6.05 6.30 6.50]
$\tilde{X}_6$= [6.01 6.10 6.25]	$\tilde{X}_{16}$=[5.70 5.77 5.83]	$\tilde{X}_{26}$=[6.25 6.35 6.45]
$\tilde{X}_7$=[6.15 6.20 6.30]	$\tilde{X}_{17}$=[6.23 6.32 6.40]	$\tilde{X}_{27}$=[5.65 5.86 6.05]
$\tilde{X}_8$= [5.64 5.81 6.05]	$\tilde{X}_{18}$=[5.60 5.70 5.08]	$\tilde{X}_{28}$=[5.70 5.87 5.95]
$\tilde{X}_9$=[5.8 5.9 5.98]	$\tilde{X}_{19}$=[5.85 5.95 6.05]	$\tilde{X}_{29}$=[5.75 5.95 5.15]
$\tilde{X}_{10}$=[6.01 6.12 6.24]	$\tilde{X}_{20}$=[5.90 6.00 6.10]	$\tilde{X}_{30}$=[6.10 6.23 6.46]

TABLE II . Confidence interval $100(1-\alpha)\%$ for C_{pm} with fuzzy data

α	Twosided	
	LCL δ_1 δ_2	UCL δ_1 δ_2
0.00001	0.3336 0.3336	1.1743 1.1748
0.0001	0.3700 0.3699	1.0431 1.0434
0.001	0.4132 0.4131	0.9844 0.9847
0.025	0.4930 0.4929	0.8839 0.8840
0.05	0.5153 0.5152	0.8574 0.8575
0.1	0.5407 0.5406	0.8280 0.8281
0.3	0.5908 0.5907	0.7721 0.7722
0.6	0.6341 0.6341	0.7259 0.7260
1	0.6796 0.6795	0.6796 0.6795

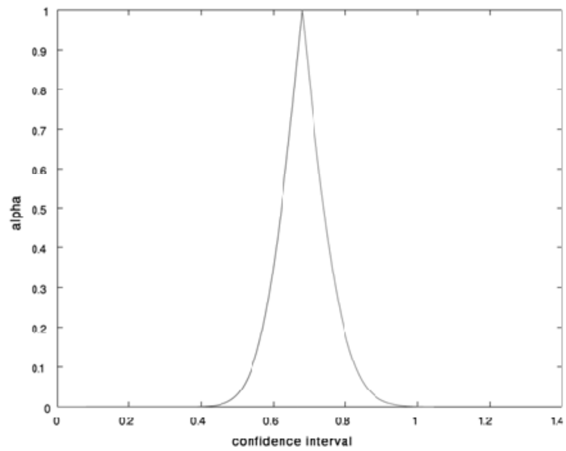

Figure 1. Membership function of $\widetilde{C}_{pm}$

TABLE III. Effect of fuzzy target on $C_{\widetilde{p}m}$

target	C_{pm}	$\left[D_{2,1/2}(\widetilde{\widetilde{X}},\widetilde{\widetilde{T}})\right]^2$
[6.00 6 6.00]	0.6869018	0.0069152963
[5.95 6 6.05]	0.7170677	0.0029875185
[5.90 6 6.10]	0.7363444	0.0007264074
[5.85 6 6.15]	0.7416758	0.0001319630
[5.80 6 6.20]	0.7321417	0.0012041852
[5.75 6 6.25]	0.7093643	0.0039430741
[5.70 6 6.30]	0.6768055	0.0083486296
[5.65 6 6.35]	0.6384898	0.0144208519
[5.60 6 6.40]	0.5979433	0.0221597407
[5.55 6 6.45]	0.5577162	0.0315652963
[5.50 6 6.50]	0.5193811	0.0426375185
⋮	⋮	⋮
[5.00 6 7.00]	0.2805835	0.2450264074

REFERENCES

[1] C. C. Chen, C. M. Lai, and H. Y. Nien, *Measuring process capability index Cpm with fuzzy data, Springer Science*, Business Media B.V., 2008.

[2] A. Parchami, M. Mashinchi, and H.R. Maleki, *Fuzzy confidence interval for fuzzy process capability index*, Journal of Intelligent & Fuzzy Systems, 17(2006), 287-295.

[3] M. Perakis, and E. Xekalaki, *A new method for constructing confidence intervals for the index Cpm*, Qual. Reliab. Engng.Int, 20 (2004), 651-665.

[4] E. Pearson, *Note on approximation to the dietribution of noncentral* χ^2 , Biometrika, 46 , 364,1959.

[5] N.N. Vakhania, *Probability distribution on linear space*, Elsevier science publishes B.v. North holland (1981).

B. SADEGHPOUR-GILDEH, Department of Statistics, Faculty of Basic Science, University of Mazandaran, Babolsar, Iran
E-mail address: **sadeghpour@umz.ac.ir**

S. ASGHARI*, Department of Statistics, Faculty of Basic Science, University of Mazandaran, Babolsar, Iran
E-mail address: **sasghari@umz.ac.ir**

* SPEAKER