Study on Process Capability Index of Fuzzy Set Distance Based on Beta Distribution

Ying Bai[1,a]
[1]Beijing information Science Technology University, China
[a]baiying401@163.com

Keywords: statistical process control, fuzzy set, process capability index, Beta.

Abstract: This paper focused on the question of measuring process capability index under the Beta distribution. First of all, analyze process capability index under Beta distribution, get the equation of process capability index. Then, in order to know the effects of process capability index on process capability, established the formula of C'_{P}、C'_{PU} and C'_{PK}. At last, based on bootstrap method and Monte Carlo method, do statistical simulation on estimated value of process capability index.

Introduction

As an important part of control chart, Process Capability Index （CP, CP_K） is a main measurement on process control. When production process is under statistical control, Process Capability Index is measurement on measuring how well the capability of production process can satisfy product quality standard.

Fuzzy set distance is an index on measuring how close two fuzzy sets approach. When fuzzy set distance goes to 1, it means these two fuzzy sets are similar. And when fuzzy set distance goes to 0, it means these two fuzzy sets are different. Fuzzy set distance is a quantized index of describing the degree of approximation between two fuzzy sets. Fuzzy set distance is usually shown as $N(A,B)$.[1]

Analysis of Process Capability Index under Beta Distribution

Assume fuzzy set distance as the random variable which follows Beta distribution $\mathrm{Beta}(a,b)$. According to abnormal distribution, analyzing process capability indexes. Set $\alpha = 0.135\%$, Then get ξ_{α} and $\xi_{1-\alpha}$ correspond to 0.135% and 99.865% quantile of fuzzy set distance, respectively. Then Process Capability Index under Beta Distribution C'_{P} is:[2]

$$C'_{P} = \frac{T}{\xi_{1-\alpha} - \xi_{\alpha}} \tag{1}$$

In Equation (1), T is tolerance, the range of production. Based on control chart of fuzzy set distance, Tolerance range is usually $0^{+T_U}_{-0}$ ($T_U \le 1$), and $T = T_U$. T_U is the upper limit of common difference. Only if unilateral process capability index C'_{PU} is below the upper limit of common difference, it has significant practical value in the control chart of fuzzy set distance. C'_{PU} is shown below:

$$C'_{PU} = \frac{T_U - \xi_{0.5}}{\xi_{1-\alpha} - \xi_{0.5}} \tag{2}$$

Similar to the method above, when the tolerance center and the center of quality characteristic value distribution are different, define skew process capability index C'_{PK} as below:

$$C'_{PK} = \frac{T-|\xi_{0.5}-M|}{\xi_{1-\alpha}-\xi_{\alpha}} \tag{3}$$

In Equation (3), $\xi_{0.5}$ is the median of Beta Distribution. M is the tolerance center. According to these Equation s, when quality characteristic value follows normal distribution, C'_P and C'_{PK} equal to C_P and C_{PK}, respectively.[3]

Analysis on relative estimate value of process capability under simulation analysis of statistical characteristics

The estimate value of C'_P、 C'_{PU} and C'_{PK}

$\hat{C}'_P$ is a natural estimate value of C'_P. The formula of $\hat{C}'_P$ is

$$\hat{C}'_P = \frac{T}{\xi_{0.99865}-\xi_{0.00135}} \tag{4}$$

In Equation (4), $\xi_{0.99865}$ is the 0.135% quantile of sample in actual producing process, and $\xi_{0.00135}$ is the 99.865% quantile of sample in actual producing process. When sample quantity is not very large, Bootstrap method can be took advantage to extend sample quantity. After sample quantity get large, centile of sample can be got. For instance, when sample quantity is 100, set Bootstrap as a base, get 100 selections. [4]In each selection, select 100 samples and then put them back. So get a sample which contains 10000 sample quantity. Then set the sequence of fuzzy set distance of each sample from low to high, take the mean value of 13th sample and 14th sample as $\xi_{0.00135}$, and the mean value of 9986th sample and 9987th sample as $\xi_{0.99865}$. Accordingly, get the formula of $\hat{C}'_{PU}$ and $\hat{C}'_{PK}$. $\hat{C}'_{PU}$ is the natural estimate value of C'_{PU}, and $\hat{C}'_{PK}$ is the natural estimate value of C'_{PK}.

$$\hat{C}'_{PU} = \frac{T-\xi_{0.5}}{\xi_{1-\alpha}-\xi_{0.5}} \tag{5}$$

$$\hat{C}'_{PK} = \frac{T-|\xi_{0.5}-M|}{\xi_{1-\alpha}-\xi_{\alpha}} \tag{6}$$

Analysis on relative estimate value under statistic model

Distribution Simulation of statistic value $\xi_{0.99865}$

Get some groups of random variable from Beta distribution. Make the number of group as m, and each group has a number of n samples. After calculation, get the observed value $\hat{a}$ and $\hat{b}$, which are maximized likehood statistics in each group. Thereby, statistic value $\xi_{0.99865}$ can be got. When m=10000, and get different value of n=20, 50, 100, 1000. Then get 4 frequency distribution charts of $\xi_{0.99865}$.

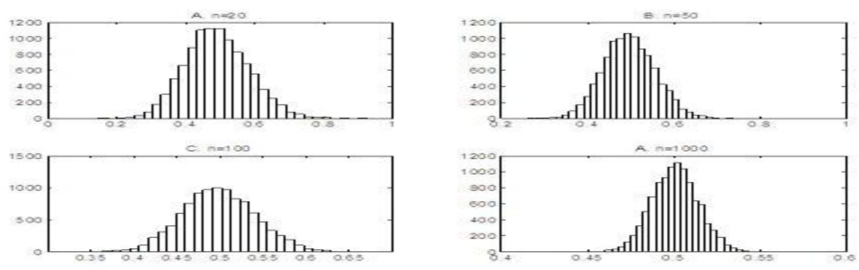

Fig.1. When n get different value, different frequency distributions of $\xi_{0.99865}$

In the above Figure, the distribution of $\xi_{0.99865}$ has the following 2 features. Firstly, when the sample quantity is little ($n=20$), the distribution is similar to symmetrical distribution. When the sample quantity is large($n=1000$), the distribution is similar to normal distribution.

Then, after analyzing each sub chart of Figure 1, the middle of distribution approach $\xi_{0.99865} = 0.501355$. Moreover, the larger the sample quantity is, the closer to $\xi_{0.99865} = 0.501355$ the middle of distribution is.

Distribution Simulation of statistic value $\hat{C}'_{P}$

Get the frequency distribution chart of $\hat{C}'_{P}$ by using similar way to get frequency distribution chart of $\xi_{0.99865}$.

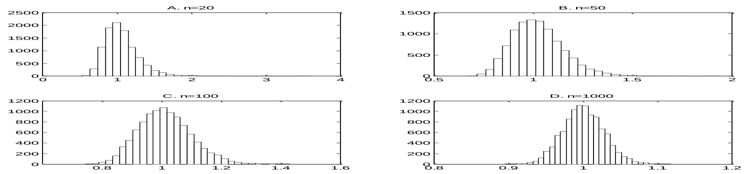

Fig.2. When n get different values, different frequency distributions of $\hat{C}'_{P}$

By using the same method, get frequency distribution of $\hat{C}'_{PU}$ and $\hat{C}'_{PK}$ relatively in Fig.3 and Fig.4.

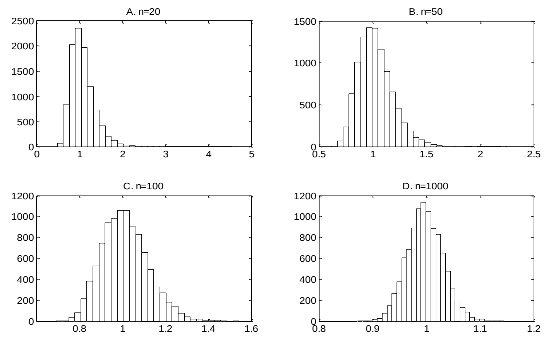

Fig.3.Frequency histogram of $\hat{C}'_{PU}$ for different n

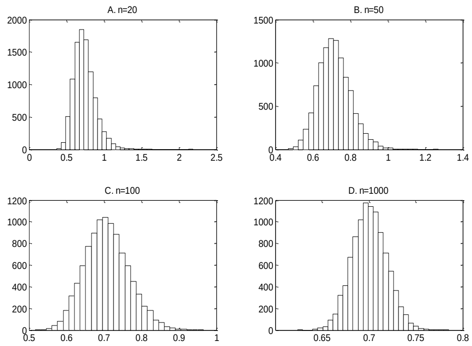

Fig.4.Frequency histogram of $\hat{C}'_{PK}$ for different n

Furthermore, after simulation, get the sample statistic value of $\hat{C}'_{P}$, $\hat{C}'_{PU}$, and $\hat{C}'_{PK}$, is shown in Table 1.

Table 1. Sample statistic value of parameter a's estimate value

Statistic	Mean	Median	Stande viation	skewness coefficient	Kurtosis	Mini	Maxi	JarqueBera	Probability
$\hat{\xi}_{0.99865}$ (n=20)	0.49179	0.48851	0.09173	0.230536	2.91384	0.22902	0.86431	91.67159	0.00000
$\hat{\xi}_{0.99865}$ (n=50)	0.49658	0.49547	0.05851	0.112843	2.97144	0.29155	0.73811	21.56242	0.00002
$\hat{\xi}_{0.99865}$ (n=100)	0.49814	0.49728	0.04176	0.075533	2.95197	0.35944	0.64729	10.46991	0.00533
$\hat{\xi}_{0.99865}$ (n=1000)	0.50105	0.50094	0.01341	-0.00397	3.03017	0.44537	0.552746	0.405440	0.81651
$\hat{C}'_{P}$ (n=20)	1.06189	1.02831	0.21676	0.974853	4.57097	0.57851	2.25435	2612.208	0.00000
$\hat{C}'_{P}$ (n=50))	1.02568	1.01278	0.12713	0.653427	3.80332	0.67745	1.75105	980.4982	0.00000
$\hat{C}'_{P}$ (n=100)	1.01439	1.00853	0.08798	0.457785	3.39333	0.77271	1.40713	413.7408	0.00000
$\hat{C}'_{P}$ (n=1000)	1.00134	1.00080	0.027296	0.174460	3.10742	0.90594	1.12863	55.53462	0.00000
$\hat{C}'_{PU}$ (n=20)	1.07769	1.02989	0.273762	1.144676	5.16855	0.50414	2.70052	4143.230	0.00000
$\hat{C}'_{P}$ (n=50))	1.03005	1.01163	0.15791	0.758401	4.05574	0.62450	1.98187	1423.032	0.00000
$\hat{C}'_{PU}$ (n=100)	1.01530	1.00679	0.10862	0.533210	3.52869	0.723627	1.51846	590.3162	0.00000
$\hat{C}'_{PU}$ (n=1000)	0.99844	0.997647	0.03347	0.198181	3.12645	0.88264	1.15702	72.12295	0.00000
$\hat{C}'_{PK}$ (n=20)	0.74066	0.72389	0.14259	0.827304	4.31304	0.38870	1.57933	1859.084	0.00000
$\hat{C}'_{PK}$ (n=50))	0.71707	0.71024	0.08494	0.503901	3.52968	0.44853	1.15525	540.0944	0.00000

$\hat{C}'_{PK}$ (n=100)	0.70913	0.70616	0.05928	0.377703	3.33335	0.53376	0.99427	284.0661	0.00000
$\hat{C}'_{PK}$ (n=1000)	0.70058	0.70016	0.01847	0.127817	3.05023	0.64028	0.78421	28.27968	0.00000

Under the analysis on Table 1 and Fig.2, Fig.3, and Fig.4, distributions of $\hat{C}'_{P}$, $\hat{C}'_{PU}$, and $\hat{C}'_{PK}$ has the following features. First of all, all distributions of these 3 statistic values lean to right. But as sample quantity gets large, distributions go to symmetrical gradually. Then, in sub charts of Chart 8, the modes of these distributions approach $1/(\xi_{0.00135}+\xi_{0.99865})=0.999929$. Moreover, when sample quantity gets larger, these modes get closer to $1/(\xi_{0.00135}+\xi_{0.99865})=0.999929$.Next, in sub charts of Fig.3, the modes of $\hat{C}'_{PU}$ distribution approach $(0.5-\xi_{0.5})/(\xi_{0.99865}-\xi_{0.5})=0.996624$. And, when sample quantity gets larger, these modes get closer to $(0.5-\xi_{0.5})/(\xi_{0.99865}-\xi_{0.5})=0.996624$.Furthermore, in sub charts of Fig.4, the modes of $\hat{C}'_{PK}$ distribution approach $(0.5-|\xi_{0.5}-0.25|)/(\xi_{0.99865}-\xi_{0.5})=0.699695$. Meanwhile, when sample quantity gets larger, these modes get closer to $(0.5-|\xi_{0.5}-0.25|)/(\xi_{0.99865}-\xi_{0.5})=0.699695$.At last, after analysis on $\hat{C}'_{P}$, $\hat{C}'_{PU}$, and $\hat{C}'_{PK}$ distribution, the fluctuation of $\hat{C}'_{PK}$ is the smallest, and its distribution approaches normal distribution the most. The fluctuation of $\hat{C}'_{PU}$ is the largest, and its distribution leaves away from normal distribution the most. Thus, according to statistical method, $\hat{C}'_{PK}$ and $\hat{C}'_{P}$ are good for leveling process capability.

Conclusion

This paper analyzes process capability by the features of random variable in fuzzy set distance. Based on reference, extend the definition of process capability index C_P and C_{PK}.[5] Then, get three process capability indexes C'_P, C'_{PK}, and C'_{PU}, which follow Beta distribution. However, after research, these indexes are restrictive for leveling process capability. Then, with optimized analysis on process capability of rate of rejection, product quality loss, and product value, get the way to optimize overall parameters of Beta distribution. At last, introduce $\hat{C}'_P$ and $\hat{C}'_{PK}$ as the estimate value of C'_P and C'_{PK} relatively. Due to complexity of these estimate values, Monte Carlo method is introduced to get distribution simulation, then take deep analysis on these distribution.

Acknowledgement

This research was financially supported by the Natural Science Foundation of Beijing ,China(NO.9144028); Project supported by Social Science Foundation of Beijing ,China (NO.14JGC105) .

References

[1] Y-Bai,Research in fuzzy quality control chart based on similarity index, international conference on mechanical science and engineering,2011, volume197:pp,50-54.

[2] S. Thomas Foster,Management Quality-An Integrative Approach, edited by China People's University Press, Beijing,2006.

[3] Taleb, H. Limam M,On Fuzzy and Probabilistic Control Charts. International Journal of Production Research, 40 (2002) 2849-2863.

[4] Wei-Zhu,The research on process capability index under the asymmetric deviation ,the China University of the Mining and Technology ,Beijing,(2007).

[5] Clements. J, Process Capability Calculations for Non-Normal Distributions,Quality Progress,1989, 22(9): 95~100.