Journal of the Society of Korea Industrial and Systems Engineering
Vol, 25, No. 6, pp.48－53, December 2002.

지수분포의 공정능력 평가

김홍준* · 김진수**

*대구산업정보대학 산업안전보건과 · **한밭대학교 산업경영공학과

Evaluation of Procss Capability for Exponential Distribution

Hong-Jun Kim* · Jin-Soo Kim**

*Dept. of Industrial Safety Health, Taegu Polytechnic College
**Dept. of Industrial and Management Engineering, Hanbat National University

품질보증 부분에서 대부분 통계적 수법들은 기지분포를 가정하여 데이터를 분석하고 있다. 예를 들면 공정능력에서는 정규분포를 가정하고 있고, 신뢰성 분석에서는 지수분포, 대수정규분포 또는 와이블분포 등을 가정하고 있다. 만약 이러한 것들이 가정하는 확률분포와 크게 편의될 때 도출된 결론은 그 유효성을 상실하게 된다. 따라서 공정이 정규분포를 하지 않는다면 정규분포에 기초한 공정능력지수의 사용은 부정확한 공정의 정보를 제시하게 된다. 이러한 문제를 해결하기 위하여 그 사례로 소표본일 경우 정규성의 유무를 Lilliefors 검정통계량으로 검정하였다. 그리고 최우추정량(MLE), 수정적률추정량(MME), 및 적률추정량(ME)을 사용하여 모수추정을 하여 그 각각의 경우에 공정능력지수로 평가하였다. 공정능력지수의 평가는 공정이 지수분포를 하는 경우에 적합한 공정능력지수(C_{pe})를 제안하고, 이것의 대안으로써 Pearson 시스템, Johnson 시스템 및 Burr 시스템과도 비교평가 하였다.

Keywords : process capability, process capability index based on exponential curves(C_{pe}), Pearson system, Johnson system, Burr system.variance method(WVM).

1. Introduction

A common, assumption in capability studies in that the individuls in the process being follow a normal distribution. If this is not the case, especially when the underlying probability distribution in heavily skewed, then the conclusions of the study are likely to be invalid. Yet, acceptable replacements for the process capability indices based on a normal distribution are available, if we could only determine the underlying distribution. To solve this problem, the first thing, exponential data is applied the Lilliefors test at statistic to the null hypothesis of normality.

The next exponential parameters is estimated in terms of the maximum likelihood estimators(MLE), a modification of the moment estimators(MME) and the moment estimators(ME) and then evaluated, respectively, process capabiliy index based on exponential curves(C_{pe}). We also consider process capability indices on Pearson system, Johnson system and Burr system.

The percentage nonconforming can be related to an equivalent capability indices for a process having a normal distribution. Thus, considering only process capability indices or percentage nonconforming separately is not a valid assessment of process capability, both criteria must be evaluated jointly.

2. The Lilliefors Test for the Exponential Distribution and parameter Estimatian

The two-parameter exponential distribution has (cdf, pdf) and hf

$$F(x;\theta,\gamma)=1-\exp(-\frac{x-\gamma}{\theta}),\ x>r \quad (1)$$

$$f(x;\theta,r)=\frac{1}{\theta}\exp(-\frac{x-r}{\theta}) \quad (2)$$

$$h(x;\theta,r)=\frac{1}{\theta} \quad (3)$$

Where θ>0 in a scale parameter and r is both a location and a threshold parameter. For $\gamma=0$ this in the well-known one-parameter exponential distribution.

The mean and variance of the exponential distribution are, respectively,

$$E(X)=r+\theta \quad (4)$$

$$Var(X)=\theta^2 \quad (5)$$

The P quantile of the exponential distribution is

$$x_p=r-\log(1-p)\theta \quad (6)$$

The exponential distribution in widely used in the field of reliability engineering as a model of the time of a component or system. In these application, the exponential distribution is a popular distribution for some kinds of electronic components as an example, capacitors or robust, high-quality integrated circuits.

But, this exponential distribution would not be appropriate for a population of electronic component having-causing quality defects.

The exponential distribution in usually inappropriate for modeling the life of mechanical components like bearing, subject to some combination of fatigue, corrosion, or wear. It is also usually inappropriate for electronic components that exhibit wearout properties over their technological life like lasers and filament devices[6].

2.1 The Lilliefors Test for the Exponential distribution

The data consist of a randome sample $X_1, X_2, \ldots, X_n$ of size n associated with some unknown distribution function, denoted by F(x). Compute the sample mean for use as an estimate of the unknown pasameter. For each X_i, compute Z_i, defined by

$$Z_i=X_i/\overline{X} \quad (7)$$

for use in conputing the test statistic.

First, the empirical distribution function $S(x)$ based on $Z_1, \ldots, Z_n$ in plotted on a graph. On the same graph the function $F^*(x)=1-e^{-x}$ is plotted for x>0 ; actually, only values at n point need to be determined, the points being at $x=Z_1$, $x=Z_2$, and so on. The maximum vertical distance between the two functions

$$T=\underset{x}{SUP}\mid F^*(x)-S(X)\mid \quad (8)$$

in the test statistic [3].

2.2 Parameter Estimation

2.2.1 Maximum Likelihood Estimator (MLE)

The density approximation to the likelihood generally provides an adequate approximation for the exponential distribution. For a sample consisting of only right-censored observation and observations reported as exact failure times, it is easy to show the MLE of θ is computed as

$$\hat{\theta}=\frac{TTT}{r} \quad (9)$$

Where $TTT=\sum_{i=1}^{\infty}$ in known as the total time on test, and t_i, $i=1,\ldots,n$, and the reported failure times for units that failed and the running (or censoring) time for the right-censored observations[6].

2.2.2 Moment Estimators (ME)

Moment Estimators (ME) of the exponential distribution are derived by equating sample moments to the corresponding distribution moment.

The first two moments are given by E_q.(2.4), E_q.(2.5) and the resulting ME equations are

$$\hat{\theta}=s,\ \hat{r}=\bar{x}-s \quad (10)$$

Here, $\bar{x}$ is the sample mean, and s is the sample standard deviation [11].

2.2.3 Modified Moment Estimators (MME)

Modified Moment Estimators[11] are variations of MLE and ME that employ the first-order statistic $X_{(1)}$, usually replacing the ME equation involving the third moment by one comparing the smallest observation with $X_{(1)}$

$$E(X_{(1)}) = r + \frac{\theta}{n} \quad \cdots\cdots (11)$$

and the MME equation are

$$\bar{x} = r + \theta, \; x_{(1)} = r + \frac{\theta}{n} \quad \cdots\cdots (12)$$

from which we get

$$\hat{\theta} = \frac{n(\bar{x} - x_{(1)})}{n-1}, \; \hat{r} = \frac{nx_{(1)} - \bar{x}}{n-1} \quad \cdots\cdots (13)$$

3. Process Capability Measures

3.1 Process Capability Indics

3.1.1 Process Capability Index based on Exponential distribution curvers (C_{pe})

The process capability index over the Weibull distribution process by Mukherjee[7] is given by

$$I = \frac{U-L}{R} = \frac{U-L}{F^{-1}(p_2) - F^{-1}(p_1)} = \frac{\theta^{-1/k}(U-L)}{[-\log(1-p_2)]^{1/k} - [-\log(1-p_1)]^{1/k}} \quad (14)$$

Here, We consider a quality(reliability) characteristic (x) having the exponential probability density function.

$$f(x\,;\,\frac{1}{\theta}) = \frac{1}{\theta} e^{-\frac{1}{\theta}} \quad \cdots\cdots (15)$$

Where ① U, L = upper and lower specification limits, respectively, for the quality characteristics(x).

② $\frac{1}{\theta}$=scale parameter of the exponential distribution.

③ $\mu =$ mean time to failure = θ

④ $\sigma =$ standard deviation of time to failure = θ

⑤ t_1, t_2 =lower and upper limites, respectively, of the process capability interval (also called the natural process interval)

⑥ $R = t_2 - t_1$

⑦ p_1, p_2 = areas under the exponential distribution curve to the left of t_1 and t_2. Thus, t_1 and t_2 are just the quantiles of the exponential distribution of order p_1 and p_2, respectively.

⑧ $C = p_2 - p_1$ = the area to be covered by the capability interval.

Similarly, process capability index based on exponential distribution curves (C_{pe}) may be obtained from the process capability index over the Weibull distribution (I)

$$I = \frac{U-L}{R} = \frac{U-L}{F^{-1}(p_2) - F^{-1}(p_1)} = \frac{\frac{1}{\theta}(U-L)}{[-\log(1-p_2)] - [-\log(1-p_1)]} \quad \cdots\cdots (16)$$

As an estimate of the process capability index, C_{pe} works out as

$$\hat{C}_{pe} = \frac{\frac{1}{\hat{\theta}}(U-L)}{[-\log(1-p_1-C)] - [-\log(1-p_1)]} = (U-L)G(\hat{\theta}^{-1}) \quad \cdots\cdots (17)$$

Where $\hat{\theta}^{-1}$ is MLE

3.1.2 Process capability Index based on the Pearson system

The Process capability Index are numerical quantities whose purpose is to indicate to what degree the output of a process is capable of staying within preassigned specifications, the so-called spec limits, when the process is in control. The condition of being in control is indicated on a control chart when all data points lie within certain control limits and no apparent trends or patterns are present.

Process are operated at some target, or nominal, value and have an upper limit *USL* and lower limit *LSL*. The standard process capability indices for normal distribution are

$$C_p = \frac{USL - LSL}{6\sigma} \quad \cdots\cdots (18)$$

and

$$C_{pk} = \min\left\{\frac{USL-\mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\} = \min(C_{pl}, C_{pu}) \quad (19)$$

Clenents[2] has developed a tecknique for adjusting C_p or C_{pk} non-normal situations based on Pearson curves, the so-called the percentile method is proposed. In this case, the formulas

$$C_p = \frac{USL - LSL}{P_{0.99865} - P_{0.00135}} \quad \cdots\cdots (20)$$

and

$$C_{pk} = \left\{ \frac{USL - x}{P_{0.99865} - x}, \quad \frac{x - LSL}{x - P_{0.00135}} \right\} = \min(C_{pu}, C_{pl}) \quad \cdots\cdots (21)$$

are usually used, although there are good reasons for replacing $\bar{x}$ here with the median $P_{0.5}$

In these formulas, $P_t = F^{-1}(t)$ is the t th percentile of the true distribution function F. when F in unknown, and assumed non-normal, E_q.(20) and (21) may be used to compute a reason approximation to F.

3.1.3 Process Capability Index based on the Johnson system

Johnson provided an alternation to the Pearson systems of curves for modeling non-normal distribution. His approach to start with a small set of curves capable of approximating the shape of wide specturm of probability distribution and then to find simple transformations that would convert these curve into the standard normal, or Z distribution. For S_L curves (lognormal), the Johnson transfornation can be witten

$$Z = r^* + \eta \ln(x - \varepsilon) \quad \cdots\cdots (22)$$

Farnum[6] is used Johnson curves to desribe non-normal process data

$$\hat{C}_{pk} = \min\left(\frac{Z_U}{3}, -\frac{Z_L}{3}\right) = \min(C_{pu}, C_{pl}) \quad \cdots\cdots (23)$$

where Z_U Z_L mean the specification limits *USL* and *LSL* values, respectively.

3.2 Percentage Nonconforming Measures

3.2.1 Pearson system

Karl Pearson introduced a system of distribution curves in which the frequency function $y = f(x)$ satisfies the following differential $Eq \cdot$ (24),

$$\frac{dy}{dx} = \frac{(x - c_0)y}{(c_1 + c_2 x + c_3 x^2)} \quad \cdots\cdots (24)$$

The method of matching moments plays an essential role the Pearson system, not only in the determination of parameters, but also in providing criteria for the selection of the appropriate type of curve.

3.2.2 Johnson system

Johnson[5] provided an alternative to the Pearson systems of curves for modeling non-normal distributions. His approach was to start with a small set of curves capable of approximating the shape of wide spectrum of probability distribution and then to find simple transformations that would convert these curves into the standard normal, or Z distribution. To obtain the density $G_i(x)$ of Johnson curves of type $i(i = 1, 2, 3)$, its distribution function.

$$G_i(x) = P\{z \le \gamma + \eta K_i(x, \lambda, \varepsilon)\} \quad \cdots\cdots (25)$$

in terms of the z distribution.

3.2.3 Burr system

The traditional approach in describing data sets has been to fit curves to frequency functions. In some cases, at least, this leads to troublesome integration when the cumulative distribution is desired, if only to calculate the theortical probabilities for comparison with the data. The question naturally arises whether the cumulative function could not be fit directly and the frequency function obtained, if needed, by differentiation. Such an approach is considered in this study which describes the Burr cumulative distribution[1].

$$F(x) = 1 - (1 + x^C)^{-K} \quad \cdots\cdots (26)$$

This function is fitted by the method of matching moments, using the cumulative moments instead of the ordinary moments.

4. Illustrative Example

To illustrative the use of the Lilliefors test for normality, as gien by Owen and Li[9] lead to exponential data.

0.029	0.046	0.133	0.194	0.265	0.287	0.322	0.433	0.441	0.464
0.483	0.528	0.606	0.789	0.940	1.681	1.766	2.014	3.088	3.279

To evaluate process capability, We will use upper specification limit *USL*=3 for these data.

4.1 The Lilliefors Test for the Exponential distribution

4.1.1 Hypotheses

H_0 : The random sample has the exponential distribution

$$F(x)=\begin{cases} 1-e^{x/t} & ,x>0 \\ 0 & ,x<0 \end{cases} \quad \cdots\cdots (27)$$

Where t is an unknown parameter

H_1 : The distribution of X is not exponential

The largest absolute deviation between $S(x)$ and $F^*(x)$ is seen to equal 0.1559.

The null hypothesis of an exponential distribution may be rejected $\alpha=0.05$ only if T exceeds 0.2345 (n=20, 1-α =0.95).

Since T=0.1559, the null hypothesis is accepted.

4.1.2 Parameter Estimation

① MLE

Fron E_q.(9), it is given an $\hat{\theta}$=0.8894

② ME

Fron E_q.(10), it is given an $\hat{\theta}$=0.9670

③ MME

Fron E_q.(13), it is given an $\hat{\theta}$=0.9057

4.2 Evaluation of Process Capability Indices

4.2.1 Process Capability index based on exponential distribution curves (C_{pe})

(1) MLE

From the table of Mukherjee et al.[8]

for k=1, the process capability index $C_{pe}=U/3.912\sigma$

Hence, the estimated process capability index is

$$\hat{C}_{pe}=\frac{3}{(3.912)(0.8894)}=0.86$$

It indicates that the process is nearly capable.

(2) ME

Similarly, the estimated process capability index is

$$\hat{C}_{pe}=\frac{3}{(3.912)(0.9670)}=0.79$$

(3) MME

Similarly, the estimated process capability index is

$$\hat{C}_{pe}=\frac{3}{(3.912)(0.9057)}=0.85$$

4.2.2 Process capability indices based on the Pearson system

As computed from example, We get the statistics of sample as follows :

$\hat{\mu}=0.8894$, median $P_{0.5}=0.369$, $P_{0.99865}=3.628$, $P_{0.00135}=0.153$, $\hat{\sigma}=0.9670$

Hence, the estimated process capability indices are $\hat{C}_p=0.86$ and $\hat{C}_{pk}=0.82$.

It indicates that the process is nearly capable.

4.2.3 Process capability index based on the Johnson system

To choose appropriate Johnson curve, a discriminant function by Slifker and Shapiro[10] applied in this example S_L distribution.

Then, on numerical computation, We get

$$\hat{Z}=0.5770+0.7973\ln(x-0.009)$$

<Table 1>. The evaluation of process capability measures

Population type	System (Measure of process capability)		$\hat{C}_{pk}$	$\hat{C}_{Npk}$	$\hat{C}_{pe}$	Nonconforming (PPM)	Remark ($E_q \cdot \hat{C}_{pk}$)	Cpability Rating
Normal	-		0.73			14,630	0.73	poor
Non-Nonmal	Exponential	MLE			0.86	36,430	0.60	poor
		ME			0.79	-	-	
		MME			0.85	36,430	0.60	
	Pearson			0.82		59,815	0.52	poor
	Johnson			0.48		73,530	0.48	very poor
	Burr		-			36,818	0.60	poor

Hence, the estimated process capability Index is

$$\hat{C}_{pl} = \frac{Z_u}{3} = \frac{1.45}{3} = 0.48$$

It indicates that the process is very poor.

The evaluation of process capability indices on the above-mentioned system are tabulated in the <table 1>.

4.3 Estimating the percentage nonconforming

The percentage nonconforming is illustrated in reference to a system with Pearson, Johnson and Burr in <table 1>.

5. Summary and Conclusions

The main objective of this paper to purpose a evaluating methods of process capability measures for exponential distributed quality characteristics. For correctly evaluating process capability, the first thing, exponential data is applied the Lilliefors test statistic to the null hypothesis of normality. The next, exponential parameters is estimated in terms of MLE, ME, MME and then evaluated, respectively, process capability index based on exponential curved (C_{pe}). We also consider process capability indices based on Pearson system, Johnson system and Burr system.

The use of process capability indices and the percentage nonconforming are illustrated simultaneously in reference to a system with Pearson, Johnson and Burr. Because, considering only process capability indices or percentage nonconforming separately is not a valid assessment of process capability, both criteria must be evaluated jointly.

Form calculated results in the <table 1>, it makes little difference the MLE, MME method and Pearson system. These vlaue indicates that the process in nearly capable. The ME method is not good in this example. With sample data, a suppose case, We will be accept the lognormal in favor of process capability for lognormal distribution is estimated $C_{pk}=0.48$, namely, Johnson system is underestimated than the others.

References

[1] Burr, I. W.; Cumulative Frequency Functions, Ann. Math. Stat., pp.215-232, 1942.

[2] Clements, J. A.; "Process Capability Calculations for Non-Normal Distributions", Quality progress, 9(2), pp. 329-336, 1989.

[3] Conover, W. J.; Practical Nonparametic Statistics, John Wiley & Sons, Inc., New York, 1998.

[4] Farnum, N. R.; "Using Johnson Curves to describe Non-Normal Process Data", Quality Engineering, 9(2), pp. 329-336, 1996~7.

[5] Johnson, N.; "System of Frequency Curves Generated by Translation", Biometrika, 36, pp.149-176, 1949.

[6] Meaker, W. Q. and Escobar, L. A.; Statistical Methods for Reliability Data. John Wiley & Sons, Inc., New York. 1997.

[7] Mukherjee, S. P.; "Process Capability Indices and Associated Inference Problems", in the proceedings of International Conference on Statistical Methods and Statistical Conputation for Quality and Reliability Improvement, 1995.

[8] Mukhejee, S. P. and Singh, N. K.; "Sampling properties of an Estimator of a New Process Capability Index for Weibull Distributed Quality Characteristics", Quality Engineering, 10(2), pp.291-294, 1997-98.

[9] Owen, D. B. and Li. H.; "The Starship for Point Estimation and Confidence Intervals on a Mean and for Percentiles", Communication in Statirtical -Simulation and Computation, 17, pp. 325-341, 1988.

[10] Slifker, J. F. and Shapiro, S. S.; "The Johnson System : Selection and Parameters Estimation", Technometrics, 22(2), pp. 239-246, 1980.

[11] ZWick, D.; "A Hybrid Method for Fitting Distribution to Data and Its use in Computing Process Capability Indices", Quality Engineering, 7(3), pp. 601-613, 1995.