

☒ 연구논문

## 공정능력지수에 대한 비평과 올바른 공정능력분석 절차

백재욱

한국방송대학교 응용통계학과

조진남

동덕여자대학교 전산통계학과

## Criticism and Guideline for the Use of Process Capability Index

JaiWook Baik

Korea National Open University

Jinnam Jo

Dongduk Women's University

### Abstract

Process capability indexes are widely used in industries. Nevertheless, important statistical issues that deal with the appropriate use of process capability indexes have only appeared in the technical literature and are thus unknown to many users. In this paper we will first review various kinds of process capability indexes and then try to understand how much uncertainty is associated with them in relation to the assumption of normal distribution, the estimate of  $\sigma$ , the sampling error, the size of number of observations available and the presence of data correlation. Finally, we will propose a proper use of process capability indexes in an effort to assess a complex process.

### 1. 서론

공정능력(process capability)이란 공정이 관리상태에 있을 때 그 공정에서 생산되는 제품의 품질변동이 어느 정도인지 나타내는 양이며, 이를 하나의 수치로 표현한 것이

공정능력지수이다. 공정능력지수는 공정에서 생산되는 제품의 품질을 평가하는 데 사용된다. 공정능력은 주로 4M(man, machine, material, method)에 의하여 영향을 받는다. 즉 작업자, 설비, 재료, 생산방식에 의하여 제품의 품질은 좌우된다. 따라서 제품의 품질을 높이기 위해서는 이들 요인중 품질변동에 영향을 주는 요소에 적절한 조치를 취함으로써 품질변동을 줄여야 한다. 공정에서 생산되는 제품의 품질변동이 작을수록 공정의 능력은 좋은 반면 품질변동이 클수록 공정의 능력은 나쁘다.

근래에 공정능력지수는 미국에서 6시그마 품질운동의 전개로 인하여 제조업체에서는 물론 영업 및 서비스 등의 비제조업 분야에서도 해당 프로세스의 능력을 평가하기 위해 많이 사용된다. 그런데 공정능력지수로 해당 공정의 능력을 올바르게 평가하기 위해서는 데이터가 서로 독립적이라든지, 정규분포를 따른다든지, 샘플링오차가 없다든지 등의 기본적인 가정이 충족되어야 한다. 그러나 현실적으로 위의 가정들은 잘 충족되지 않을 수 있다. 그런 경우 공정능력지수로 해당 공정의 능력을 평가하면 잘못된 판단을 할 수 있다(Rodriguez(1992) 참조). 이에 본 논문의 2절에서는 공정능력을 정량화하는 여러 가지 공정능력지수에 대해 알아보고, 3절에서는 공정능력지수  $C_p$ 와  $C_{pk}$  등이 공정의 능력을 측정하는 데 얼마나 불확실한 척도인지 살펴보고자 한다. 특히 데이터의 정규성 가정이 일치하지 않는 경우 공정능력지수에 대한 해석이 어떻게 달라지는지,  $\sigma$ 의 추정방법이 변하면 공정능력지수가 어떻게 변하는지, 샘플링 오차가 있다면 공정능력지수가 얼마나 불확실하게 나타나는지, 적은 수의 표본으로 공정능력지수를 추정하면 공정능력에 대한 해석이 어떻게 달라지는지, 데이터간 상관관계가 있다면 공정능력에 대한 해석이 어떻게 달라지는지에 대해서 알아보고자 한다. 아울러 4절에서는 복잡한 공정의 경우 공정능력은 실제로 어떤 과정을 거쳐 측정되어야 하는가를 살펴보고, 5절에서는 본 논문을 요약하기로 한다.

## 2. 공정능력의 정량화

본 절에서는 공정의 능력을 정량화하는 여러 가지 방법에 대해 알아보고자 한다.

### 2.1 $6\sigma$ 에 의한 공정능력

이 방법은 품질특성분포의  $\sigma$ 를 이용하여  $6\sigma$ 로 공정의 능력을 정한다. 물론  $\sigma$ 가 알려져 있지 않는 경우에는 추정해야 한다.  $\sigma$ 를 추정하는 방법은 여러 가지가 있다. 그 중 가장 많이 사용되는 방법은 다음의 두 가지이다. 첫 번째는 각 제품의 품질특성치를  $X_1, X_2, \dots, X_N$ 이라고 할 때 다음과 같이 표본표준편차를 구하는 것이다.

$$\hat{\sigma} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{N-1}}, \quad \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (2-1)$$

두 번째는 표본을 다음

$$\begin{array}{ll} X_{11}, X_{12}, \dots, X_{1n} & R_1 \\ X_{21}, X_{22}, \dots, X_{2n} & R_2 \\ \vdots & \\ X_{m1}, X_{m2}, \dots, X_{mn} & R_m \end{array}$$

과 같이 여러 개의 부분군(전체적으로는  $N=m \times n$ 개)으로 합리적으로 나눌 수 있는 경우, 우선 각 부분군에서 먼저 범위  $R_i$ 를 구하고, 이들의 평균을 이용하여  $\sigma$ 를 다음과 같이 추정한다.

$$\hat{\sigma} = \bar{R}/d_2 \quad (2-2)$$

여기서  $\bar{R} = (R_1 + R_2 + \dots + R_m)/m$ 이며,  $d_2$ 는 부분군의 크기  $n$ 의 함수로서 (2-2)의  $\hat{\sigma}$ 를  $\sigma$ 의 불편추정량으로 만드는 수정계수이다.

어느 방법으로  $\sigma$ 를 추정하든 결과적으로  $6\hat{\sigma}$ 를 사용하여 공정의 능력을 평가한다. 그런데 이 지수는 측정되는 단위에 의해서 영향을 받는다는 결점이 있다. 이러한 단점을 보완하고, 주어진 규격의 폭과 공정의 산포를 고려하여 다음과 같은 공정능력지수  $C_p$ 를 많이 사용한다.

## 2.2 $C_p$ 에 의한 공정능력(치우침이 없는 경우)

품질특성치  $X$ 의 기대치  $E(X) = \xi$ 라고 하는 경우  $C_p$ 의 정의는 다음과 같다.

$$\begin{aligned} C_p &= (USL - LSL)/(6\sigma), \quad \sigma = \sqrt{E(X - \xi)^2} \\ &= \frac{d}{3\sigma}, \quad d = (USL - LSL)/2 \end{aligned} \quad (2-3)$$

여기서 USL(upper specification limit)과 LSL(lower specification limit)은 각각 규격상한과 규격하한을 나타낸다. 식 (2-3)에서 규격상한과 규격하한간의 차이는 '허용할 수 있는 공정산포'를 나타내고, 그 값은 공정이 이루어지기 전에 이미 정해져 있는 값인 반면, '실제 공정산포'를 나타내는 공정표준편차  $6\sigma$ 는 공정의 능력에 따라 다른 값을 취하게 된다. 식 (2-3)의 공정능력지수는 단위에 영향을 받지 않으므로 서로 다른 공정간 능력을 비교할 수 있다. 공정분포의 중심이 규격상한과 규격하한의 중간에 있는 경우 식 (2-3)의 공정능력지수가 1이라는 것은 '허용할 수 있는 공정산포'와 '실제

공정산포'가 같다는 것을 의미하며, 2라고 하는 것은 '허용할 수 있는 공정산포'가 '실제 공정산포'의 두 배라는 것을 뜻한다.

데이터가 정규분포를 따르고, 공정분포가 규격한계의 중간에 위치하는 경우  $C_p$ 의 값이 1이라는 것은 2,700ppm(parts per million)[즉, 100만개의 제품중 2,700개가 규격한계선을 벗어난다는 것], 1.33이라는 것은 63ppm, 1.66이라는 것은 0.57ppm, 2라는 것은 0.002ppm을 각각 나타낸다. 구체적인  $C_p$ 의 값에 대한 해석에 대해서는 Kotz and Johnson(1993)을 참조하기 바란다.

다시 한 번 강조하거니와 앞에서와 같은 해석은 주어진 데이터가 정규분포를 따르고, 그 분포의 평균이 규격상한과 규격하한의 중간에 위치해 있을 때에만 의미가 있다. 즉, 이와 같은 가정이 충족되지 않는다면  $C_p$ 에 대한 추정치는 공정능력을 과대평가한다. 따라서 분포의 평균이 규격상한과 규격하한의 중간에 있지 않는 경우를 감안한 적절한 공정능력지수가 필요하다.

### 2.3 $C_{pk}$ 에 의한 공정능력(치우침이 있는 경우)

공정의 평균  $\xi$ 가 규격상한이나 규격하한으로부터 얼마나 멀리 떨어져 있는가를 고려하여 Sullivan(1984, 1985)과 Kane(1986a, 1986b)은 다음과 같은 공정능력지수  $C_{pk}$ 를 제안했다.

$$CPU = \frac{USL - \xi}{3\sigma} \quad (2-4)$$

$$CPL = \frac{\xi - LSL}{3\sigma} \quad (2-5)$$

$$C_{pk} = \frac{\min(USL - \xi, \xi - LSL)}{3\sigma} \quad (2-6)$$

$$\begin{aligned} &= \frac{d - |\xi - \frac{1}{2}(LSL + USL)|}{3\sigma} \\ &= (1 - k)C_p, \quad k = \frac{|\xi - \frac{1}{2}(LSL + USL)|}{d} \end{aligned} \quad (2-7)$$

$k$ 가 0이면 공정의 평균  $\xi$ 는 규격상한과 규격하한의 중간값과 같음을 의미하며, 1이면 공정의 평균  $\xi$ 가 규격상한값이나 규격하한값 중의 하나와 일치함을 의미한다.

만일 품질특성치  $X$ 의 분포가 정규분포를 따른다면 해당공정의 불량률은 다음과 같다.

$$\Phi\left(\frac{LSL - \xi}{\sigma}\right) + \left\{1 - \Phi\left(\frac{USL - \xi}{\sigma}\right)\right\} \quad (2-8)$$

여기서  $\Phi(\cdot)$ 는 표준정규분포의 누적확률을 의미한다.

## 2.4 $C_{pm}$ 및 $C_{pmk}$ 에 의한 공정능력

$C_p$ 나  $C_{pk}$ 는 공정의 평균  $\xi$ 가 목표치  $T$ 로부터 얼마나 떨어져 있는지를 고려하지 않는 지수이다. 이와 같은 점을 감안하여 Chang, Cheng and Spring(1988)은 다음과 같은 공정능력지수를 제안했다.

$$C_{pm} = \frac{USL - LSL}{6\sigma'}, \quad \sigma' = \sqrt{E(X - T)^2} = \sqrt{\sigma^2 + (\xi - T)^2} \quad (2-9)$$

여기서  $\sigma^2$ 과 마찬가지로  $\sigma'^2$ 도 추정해야 할 모수이다. 따라서  $N$ 개의 표본이 있다면  $\sigma'$ 가 알려져 있지 않는 경우 다음과 같은 식

$$\hat{\sigma}' = \sqrt{\frac{\sum_{i=1}^N (X_i - T)^2}{N-1}} \quad (2-10)$$

으로  $\sigma'$ 를 추정할 수 있고, 이를 이용하여  $C_{pm}$ 을 추정할 수 있다.

$C_{pm}$ 이  $C_p$ 와 다른 점은 식 (2-3)에서 공정의 평균  $\xi$  대신 목표치  $T$ 가 사용되고 있다는 점이다. 하지만 이 작은 변화가 공정의 능력을 '공정의 산포'와 '공정의 평균이 목표치와 떨어진 거리'라고 하는 두가지 측면에서 평가하는 데 유용하게 사용된다. 왜냐하면  $\sigma'^2 = E(X - \xi)^2 + (\xi - T)^2$ 이고  $E(X - \xi)^2 = \sigma^2$ 이기 때문이다. 따라서

$$\begin{aligned} C_{pm} &= \frac{USL - LSL}{6\sqrt{\sigma^2 + (\xi - T)^2}} \\ &= \frac{C_p}{\sqrt{1 + \frac{(\xi - T)^2}{\sigma^2}}} \end{aligned} \quad (2-11)$$

과 같이 쓸 수 있으며,  $C_{pm} \leq C_p$ 인 것을 쉽게 알 수 있다.

위의 식에서 공정분산  $\sigma^2$ 이 증가하면  $C_{pm}$ 은 감소하며, 공정의 평균  $\xi$ 가 목표치  $T$ 로부터 점점 멀어지면  $C_{pm}$ 은 감소한다.

Boyles(1991)는  $C_p$ 와  $C_{pm}$ 간의 관계를 파악하기 위해  $USL=65$ ,  $LSL=35$ ,  $T=50$ 이라고 가정하고, 다음과 같은 세 가지의 경우를 고려하였다.

A : $\xi=50$ ,	$\sigma=5$
B : $\xi=57.5$ ,	$\sigma=2.5$
C : $\xi=61.25$ ,	$\sigma=1.25$

이때 공정능력지수  $C_p$ ,  $C_{pk}$  및  $C_{pm}$ 은 각각 다음과 같이 계산된다.

공정	$C_p$	$C_{pk}$	$C_{pm}$
A	1	1	1
B	2	1	0.63
C	4	1	0.44

이 결과로부터  $C_p$ 와  $C_{pm}$ 은 서로 반대방향으로 움직인다는 것을 알 수 있다. 즉,  $C_p$ 에 의하면 공정의 능력은 C, B, A 순으로 좋으나  $C_{pm}$ 에 의하면 공정의 능력은 A, B, C 순으로 좋다. 한편,  $C_{pk}$ 에 의하면 어느 공정이나 그 능력은 똑같다. 이 예에서 알 수 있듯이 목표치 T가 규격상한과 규격하한의 중간값과 같지 않다면 식 (2-11)과 같은 공정능력지수는 공정의 능력을 재는 데 바람직한 도구라고 할 수 없다. 왜냐하면 공정의 평균  $\xi$ 가  $T-\delta$ 나  $T+\delta$ 인 경우 목표치  $T=(USL+LSL)/2$ 이 아니라면 공정의 불량률이 서로 다를지라도 두 공정에서  $C_{pm}$ 의 값은  $d(\sigma^2+\delta^2)^{-1/2}/3$ 으로 똑같이 나올 수 있기 때문이다. 이와 같은 관점에서  $T=(USL+LSL)/2$ 인 경우를 제외하고는  $C_{pm}$ 에 대해 논하는 것은 의미가 없다.

따라서 공정의 목표치 T가 규격상한과 규격하한의 중간값이 아닌 경우에도 Pearn, Kotz and Johnson(1992)은 공정의 능력을 평가하기 위해서 식 (2-11)을 다음과 같이 일반화시켰다.

$$\begin{aligned}
 C_{pmk} &= \frac{\min[USL-\xi, \xi-LSL]}{3\sigma'} \\
 &= \frac{C_{pk}}{\sqrt{1+\frac{(\xi-T)^2}{\sigma^2}}} \quad (2-12)
 \end{aligned}$$

이 식으로부터 공정의 평균  $\xi$ 가 규격상한이나 규격하한에 가까이 있을수록, 공정의 평균  $\xi$ 가 목표치 T로부터 멀어질수록  $C_{pmk}$ 는 작아진다는 것을 알 수 있다.  $\sigma'$ 가

알려지지 않은 경우에는 식 (2-10)을 사용하여 추정할 수 있으며,  $C_{pmk}$  또한 추정할 수 있다. 그런 의미에서  $C_{pmk}$ 는 여러 가지 상황에서 공정능력을 잴 수 있는 지표이다. 즉, 규격이 한쪽에만 있는 경우는 물론이고 공정의 목표치가 규격한계선의 중앙에 있지 않은 경우에도 공정의 능력을 측정할 수 있다. 물론 공정의 목표치가 규격상한과 규격하한의 중앙에 있는 경우  $C_{pmk}$ 와  $C_{pm}$ 은 같게 된다.

지금까지 식 (2-3), (2-6), (2-11), (2-12)와 같은 여러 가지의 공정능력지수를 소개하면서 각각의 지수가 공정의 어느 부분을 점검하고자 하는 것인지 설명했다. 아울러 식 (2-3), (2-6), (2-11), (2-12)와 같이 공정능력지수가 복잡해지면서 공정능력지수로 파악하고자 하는 것이 조금씩 추가되지만 그에 대한 해석이 용이하지 않다는 것을 알 수 있다.

다음은 공정능력지수가 데이터의 정규성 가정이라든지 데이터의 독립성 가정 또는 데이터의 독립성 가정 등이 성립하지 않는 경우 얼마나 불확실한 지수인지에 대해 논한다. 특히, 비판의 대상이 되는 공정능력지수는  $C_p$ 나  $C_{pk}$ 이다. 그러나 일반 산업현장에서는 해석의 편의상  $C_p$ 나  $C_{pk}$ 와 같은 간단한 지수를 많이 사용한다. 따라서 이들에 대한 비평을 중심으로 논의를 진행한다.

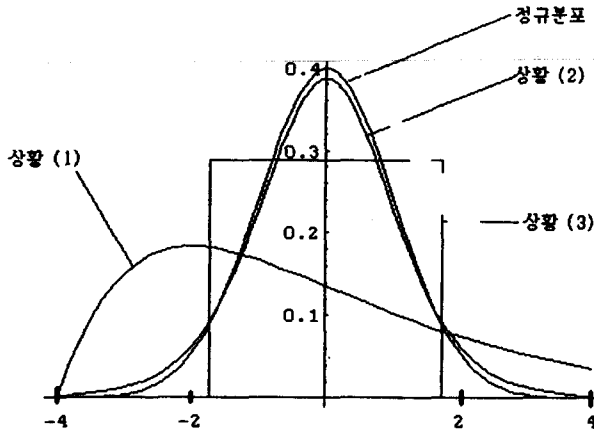
### 3. 공정능력지수에 대한 비평

#### 3.1 정규분포 가정에 대한 검토

어떤 공정에 대한 능력을 살펴보고자 할 때 공정으로부터 나오는 품질특성치의 분포가 편의상 정규분포를 따른다고 가정하고 공정능력지수를 구한다. 그런데 품질특성치의 분포가 정확히 정규분포를 따르기 어려운 경우가 많다. 다음은 Gunter(1989)에 나와있는 내용과 비슷한 것이다.

첫 번째 경우 : 한 쪽으로 경계가 있는 기울어진 분포 (<그림 1>의 상황 (1))

이와 같은 상황은 표면의 거칠음, 코팅 두께 등의 측정에서와 같이 그 측정값이 어떤 특정한 값보다 작을 수 없는 경우에 많이 나타난다. 비슷한 상황은 오염의 정도, 수명실험에서 고장날 때까지의 시간을 관측할 때에도 종종 나타난다. 이런 경우에는 정규분포보다 와이블분포 또는  $\chi^2$  분포와 같은 기울어진 분포가 현황분석에 주로 사용된다. 이 때에는 <그림 1>의 상황 (1)에서도 알 수 있듯이 공정데이터는 낮은 값에 많이 몰려있지만 큰 값이 종종 나타난다. 이와 같은 상황은 구멍뚫는 공정의 경우에도 발생할 수 있다. 즉, 구멍의 직경을 조사하는 경우 대부분 구멍의 직경은 드릴의 직경과 비슷하거나 그 보다 조금 클 것이다. 그러나 가끔 기계의 진동, 무던 도구, 잘못된 설치 등이 원인이 되어 뚫린 구멍의 직경이 예상외로 아주 커질 수 있다.



< 그림 1 > 정규분포,  $\chi^2$  분포, t 분포 및 일양분포의 경우

두 번째 경우 : 정규분포에 비해 꼬리가 길고 두툽한 분포(<그림 1>의 상황 (2))

이와 같은 분포는 공정이 완전히 관리되어 있지 않은 경우 많이 발생한다. 그런데 사실 공정이 완전히 관리되어 있지 않은 경우가 우리의 현실이므로 정규분포보다는 <그림 1>의 상황 (2)와 같은 분포가 실제 공정에 대한 상황을 더욱 잘 나타낸다고 할 수 있다.

정규분포는 어떤 중심값을 기준으로 주위에 많은 데이터들이 몰려있는 분포인데 반해, 실제 공정결과를 보면 그와 같은 현상은 좀처럼 나타나지 않는다. 기계간 또는 라인간의 아주 적은 차이, 기계설치상의 아주 적은 오차, 온도나 습도 등의 미미한 차이, 작업자간의 세세한 차이 등 소위 데밍이 말하는 자연적이며 불가피한 변동 (natural and unavoidable variation)으로 인하여 데이터가 정규분포를 따르기 보다는 <그림 1>의 상황 (2)에서와 같이 분포의 중심부분이 정규분포의 경우보다 좀더 낮으며 꼬리부분이 정규분포의 경우 보다 좀더 두툽한 형태를 띤다.

앞에서와 같은 분포는 공정이 대체로 좋은 상태로 유지되나 가끔, 그것도 정기적으로 좋은 상태를 벗어나는 경우에 종종 발생한다. 그런데 <그림 1>의 정규분포와 상황 (2)와의 미세한 차이를 히스토그램으로 파악하기는 어렵다. 이런 경우에는 정규확률지를 이용하여 주어진 데이터가 일직선을 어떤 형식으로 벗어나는가 확인해야 할 것이다. 분포에 대한 보다 정확한 정보를 얻기 위해서는 안정된 공정으로부터 수백 또는 수천 개의 측정치를 취하여 현황분석을 해야될 필요도 있다. 물론 주어진 데이터가 <그림 1>의 상황 (2)와 같은 분포를 따르는 경우 불량률의 추정치는 데이터가 정규분포를 따르는 경우보다 더 높아질 것이다.

세 번째 경우 : 꼬리가 매우 짧은 일양분포(<그림 1>의 상황 (3))

이와 같은 분포는 공정이 너무 불안정하여 분포의 꼬리 부분에 있는 나쁜 제품들이 모두 선별되어 제거되는 100% 검사시에 많이 발생한다. 그러나 그렇지 않은 경우에도



생길 수 있다. 예를 들어, 기계나 생산라인이 여러 개 있는 경우 각각은 일관성 있게 관리되어 있으나 평균이 똑같이 나오지 않는 경우가 있을 수 있다. 또 다른 예로, 공정산포는 아주 큰데도 일정한 치수 안에만 들면 합격이라고 보는 경우 이와 비슷한 현상이 발생한다. 예를 들어 작은 알갱이 형태의 환약이나 사탕이 일정한 크기의 구멍을 통과해야 합격품으로 보는 경우 <그림 1>의 상황 (3)과 비슷한 분포를 보게 된다. 이러한 분포를 갖는데도 불구하고 전체 데이터가 정규분포를 따르는 것처럼 공정능력지수를 계산하여 해석하면 공정의 능력은 과대 평가된다.

구체적으로 앞의 <그림 1>에서 정규분포는 평균  $\xi=0$ 이고 표준편차  $\sigma=1$ 인 정규분포를 나타내며, 상황 (1), (2), (3)은 각각

자유도가 4인  $\chi^2$  분포  
 자유도가 8인 t 분포  
 $-\sqrt{3}$ 에서  $\sqrt{3}$  사이의 일양 분포

를 나타낸다. 이런 경우 각각의 분포에서 평균이 0이고 표준편차가 1이 되도록 조정된 상태에서 관측치가 평균 0으로부터  $\pm 3$ 를 벗어나는 비율은 10,000개의 제품당

정규분포	: 27 개 (이중 1/2은 아래로, 1/2는 위로)
상황 1 ( $\chi^2$ 분포)	: 140개 (모두 위로)
상황 2 (t 분포)	: 85개 (이중 1/2은 아래로, 1/2는 위로)
상황 3 (일양분포)	: 0개

가 된다.

이와 같이 데이터가 정규분포를 따르지 않는데도 불구하고 정규분포를 따르는 것처럼 가정하고 공정능력지수를 구하면 공정의 능력에 대한 해석에 문제가 생길 수 있다. 그래서 Clements(1989)는 기존의 공정능력지수를 비정규분포에도 확장할 수 있는 다음과 같은 지수를 개발하기에 이르렀다.

$$C_{pk} = \min\left(\frac{USL - P_{0.5}}{P_{0.9987} - P_{0.5}}, \frac{P_{0.5} - LSL}{P_{0.5} - P_{0.0013}}\right) \quad (3-1)$$

여기서  $P_\alpha$ 는 분포의  $100\alpha\%$ 에 해당하는 백분위수(percentile)를 나타낸다. 그러면 데이터가 정규분포를 따르는 경우에도 평균 위아래로  $3\sigma$ 에 해당하는 점은 각각 99.87% 및 0.13%에 해당하는 백분위수이므로 식 (3-1)은 어떤 분포에나 적용시킬 수 있는 공정능력지수가 된다.

식 (3-1)에서와 같이  $C_{pk}$ 를 확장해서 정의하면 비대칭분포의 경우에도 의미가 있는 공정능력지수를 구할 수 있다. 하지만 문제는 분포의 꼬리에 해당하는  $P_{0.9987}$ 이나  $P_{0.0013}$ 을 명확히 추정하기 위해서는 표본의 수가 상당히 커야 한다는 점이다. 그렇지

만 이러한 어려움도 주어진 데이터에 대해 적절한 분포를 구하면 쉽게 극복할 수 있다.

### 3.2 공정능력지수의 불확실성

추정된 공정능력지수는 상수가 아닌 확률변수이다. 따라서 공정능력지수에 대한 보다 철저한 이해를 위해서는 추정된 공정능력지수에 대한 분포를 알아야 한다 (Kane (1986b), Chan, Cheng and Spiring(1988) 참조). 이때 물론 데이터가 정규분포를 따른다면 분석은 쉽다. 이런 점을 고려하여 본 논문에서도 데이터가 정규분포를 따른다고 가정하고  $\sigma$ 의 추정, 샘플링 오차, 적은 표본수, 품질특성치간 독립성의 결여와 관련하여 추정된 공정능력지수가 수반할 수 있는 불확실성의 정도에 대해서 살펴본다.

#### 3.2.1 $\sigma$ 의 추정과 관련된 공정능력지수의 불확실성

$\sigma$ 는 공정능력지수  $C_{pk}$ 의 분모에 있으므로 이 추정치가 작을수록  $C_{pk}$ 는 커지게 된다. 따라서 어떤 과정을 거쳐  $C_{pk}$ 를 계산할 것인지는  $C_{pk}$ 값을 구하는 사람이 제품의 공급자 또는 수요자에 따라서 달라질 수 있다.  $\sigma$ 의 추정치로서 제품의 수요자는 모든 변동(군내변동, 군외변동, 장기적인 변동 등)을 사용하려고 할 것이며, 제품의 공급자는 이상요인이 제거된 상태에서 군내변동만을 사용하려고 할 것이다.

공정관리에서는 평균의 관리를 위해서 평균관리도가 많이 쓰이며, 평균관리도로부터 산출되는 정보를 이용하여  $C_{pk}$ 를 종종 계산한다. 앞의 2.1에서와 같이 각 부분군의 크기가  $n$ 이고,  $m$ 개의 부분군이 있다고 하자. 이때 각각의 부분군으로부터 산술평균  $\bar{X}$ 와 범위  $R$ 을 계산할 수 있으며, 공정의 평균  $\mu$ 의 추정치로는 부분군의 산술평균  $\bar{X}$ 들을 다시 산술평균한  $\bar{\bar{X}}$ 를 쓴다. 하지만 공정의 표준편차  $\sigma$ 의 추정방법으로는 2.1에서 설명한 방법이외에 두 가지 방법이 더 있을 수 있다.

첫 번째 방법은  $N$ 개의 데이터를 군 구분없이 한 세트의 데이터로 보고 식 (2-1)에서와 같이 표본표준편차

$$\hat{\sigma} = S \quad (3-2)$$

로  $\sigma$ 를 추정하는 것이다. 두 번째 방법은 전체  $N$ 개의 데이터에서 범위  $R$ 을 구하여

$$\hat{\sigma} = R/d_2 \quad (3-3)$$

로  $\sigma$ 를 추정하는 것이다. 세 번째 방법은 식 (2-2)에서와 같이 각 부분군의 범위  $R_i$ 의 산술평균  $\bar{R}$ 를 이용한

$$\hat{\sigma} = \bar{R}/d_3 \quad (3-4)$$

로  $\sigma$ 를 추정하는 것이다. 네 번째 방법은 각 부분군에서 구한 표본표준편차  $S_j$ 의 산술평균

$$\hat{\sigma} = \bar{S} \quad (3-5)$$

로  $\sigma$ 를 추정하는 것이다.

그렇다면 이 네 가지 방법에 의한 공정능력지수, 예를 들어  $C_{pk}$ 의 추정치는 똑같은 결과를 나타낼까? 불행히도 그렇지 않다. 그렇다면 어느 것이 올바른 공정능력지수를 나타내는가? 그러나 여기에서는 어느 것이 옳고 그르다고 말할 수 없다. 왜냐하면 첫 번째나 두 번째 방법은 단기적인 변동은 물론 장기적인 변동까지 고려하여  $\sigma$ 를 추정한 것이고, 세 번째나 네 번째 방법은 단기적인 변동만을 고려하여  $\sigma$ 를 추정한 것이기 때문이다. 구체적으로 첫 번째나 두 번째 추정치인  $S$ 나  $R/d_2$ 는 군내변동은 물론 군간변동까지 고려한 총변동을 나타내는 반면, 세 번째나 네 번째 추정치인  $\bar{R}/d_2$ 나  $\bar{S}$ 는 부분군들의 단기변동(군내변동)들을 나타낸다.

실제 공정에서 소비자가 겪게 되는 변동이 군내변동과 군간변동을 합한 변동이라고 본다면  $\sigma$ 에 대한 첫 번째나 두 번째 추정치를 이용하여 공정능력지수를 구해야 할 것이다. 데이터의 수가 많은 경우 범위보다는 표준편차가 더 효율적이므로 첫 번째 추정치를 이용한 공정능력지수에 대해 더욱 이해할 필요가 있다. 하지만 때로는 평균 관리도를 그리면서 파악되는 단기간의 변동을 이용하여 공정능력지수를 많이 추정하므로 세 번째 방법(즉,  $\hat{\sigma} = \bar{R}/d_2$ )에 근거한 공정능력지수에 대하여 알아볼 필요가 있다(3.2.2 참조).

부분군이 합리적으로 형성되지 않는 경우에는 식 (3.2) 또는 (3.3)을 이용하여 해당 공정표준편차를 추정한다. 그러나 실제로는 이동범위를 이용하여 공정표준편차를 많이 추정한다. 즉,  $X_1, X_2, \dots, X_n$ 이 각 시점에서의 품질특성치,  $MR_j = |X_{j+1} - X_j|$ ,  $j=1, 2, \dots, n-1$  라고 정의하는 경우

$$\hat{\sigma} = \overline{MR}/d_2 \quad (3-6)$$

로  $\sigma$ 를 많이 추정한다. 여기서  $\overline{MR} = \sum_{j=1}^{n-1} MR_j / (n-1)$ 이다. 이 경우  $MR_j$ 들은 서로 독립이 아니므로 추정된 공정능력지수에 미치는 영향에 대해 유의할 필요가 있다(3.2.4 참조).

참고로, 일반 품질관리 현장에서는 데이터를 컴퓨터에 입력하면 공정능력지수가 자동적으로 계산된다. 그런데 사용하는 소프트웨어마다  $\sigma$ 의 추정방법이 조금씩 다르므로 공정능력지수의 계산 결과 또한 다르다. 따라서 컴퓨터를 이용하여 공정능력지수를 구하는 경우 모수의 추정이 어떤 방법으로 이루어지는지 파악해야 한다.

### 3.2.2 샘플링 오차와 관련된 공정능력지수의 불확실성

공정능력지수는 샘플링 오차(표본추출방법에 의해 발생하는 오차)에 의해서 그 값이 달라진다. 즉, 똑같은 추정방법으로 똑같은 상황하에서 여러 사람이 공정능력지수의 값을 구하기 위해 표본을 랜덤하게 뽑는 경우 추정된 값들은 서로 다르다. 따라서 불확실한 정도를 측정하기 위한 방법으로 공정능력지수에 대한 구간추정을 실시한다.

식 (3-2)에 의하여  $\sigma$ 를 추정하는 경우 주어진 데이터가 정규분포를 따른다면 표본의 수가  $N$ 일 때 다음의 통계량

$$\frac{(N-1) \hat{\sigma}^2}{\sigma^2} \quad (3-7)$$

은 자유도  $N-1$ 인  $\chi^2$  분포를 따른다. 여기서  $\hat{\sigma}^2 = \frac{1}{N-1} \sum_{j=1}^N (X_j - \bar{X})^2$ 이다. 따라서  $C_p$ 에 대한  $100(1-\alpha)\%$  신뢰구간은 다음과 같다.

$$\left( \frac{\chi_{N-1, \alpha/2}}{\sqrt{N-1}} \hat{C}_p, \frac{\chi_{N-1, 1-\alpha/2}}{\sqrt{N-1}} \hat{C}_p \right) \quad (3-8)$$

한편, 식 (3-4)에 의하여  $\sigma$ 를 추정하는 경우 Patnaik(1950)에 의하면 다음의 통계량

$$\frac{\nu \bar{R}^2}{(c^2 \sigma^2)} \quad (3-9)$$

은 대략적으로 자유도  $\nu$ 인  $\chi^2$  분포를 따른다. 여기서  $\nu$ 와  $c$ 는 부분군의 크기 및 부분군의 수에 의해 결정되는 값들이다. 따라서  $\bar{R}/d_2$ 로  $\sigma$ 를 추정하는 경우  $C_p$ 에 대한  $100(1-\alpha)\%$  신뢰구간은

$$\left( \frac{USL - LSL}{6\sqrt{\nu}} \frac{c}{R} \chi_{\nu, \alpha/2}, \frac{USL - LSL}{6\sqrt{\nu}} \frac{c}{R} \chi_{\nu, 1-\alpha/2} \right) \quad (3-10)$$

과 같이 나오며,  $C_{pk}$ 에 대한  $100(1-\alpha)\%$  신뢰구간은

$$\left( (1-k) \frac{USL - LSL}{6\sqrt{\nu}} \frac{c}{R} \chi_{\nu, \alpha/2}, (1-k) \frac{USL - LSL}{6\sqrt{\nu}} \frac{c}{R} \chi_{\nu, 1-\alpha/2} \right) \quad (3-11)$$

과 같이 나온다.

식 (3.10)과 (3.11)에서는  $C_p$ 나  $C_{pk}$ 에 대한 구간추정에 대해 설명했는데, 똑같은 통계량 (3.9)를 이용하여 어떤 공급자의 공정능력지수가 특정한 값인지 검정할 수 있다.

3.2.3 적은 표본에 기인한 공정능력지수의 불확실성

통상적으로 평균관리도 및 범위관리도로 공정의 평균 및 산포를 관리하면서 전체 공정의 능력을 살피기 위해 공정능력지수를 구한다. 이때 공정능력지수를 구하는 데 필요한 모수, 예를 들어 공정의 평균  $\xi$ 나 공정의 표준편차  $\sigma$ 는 평균관리도의 작성에 이용된 정보를 그대로 이용하는 경우가 많다. 그런데 최근 제품이 다품종 소량화되면서 예전과는 달리 데이터가 많지 않은데도 불구하고 관리도나 공정능력지수를 이용하여 공정의 능력을 평가한다. 자료가 많지 않은 상태에서 평균관리도의 관리한계선을 결정하여, 이를 공정관리에 이용하면 Quesenberry(1993)가 지적한 것처럼 제 1종의 오류를 범할 확률은 높아지고, 평균런길이(average run length)가 예상보다 길어지는 문제가 생긴다.

그렇다면 공정능력지수의 경우 적은 양의 자료로 모수를 추정하면 실제로 어떤 결과를 가져오는지 알아본다. 우선 다음과 같은 통상적인 조건이 성립한다고 가정한다. 첫째, 자료는 정규분포를 따르며, 둘째 해당 공정은 통계적으로 관리된 상태라고 가정한다.

그러면 식 (3-2)에서와 같이  $\hat{\sigma}=S$ 인 경우 Kotz and Johnson(1993)에 의하면

$$E(\hat{C}_p) = \frac{1}{b_f} C_p, \quad b_f = \sqrt{2/f} \frac{\Gamma(f/2)}{\Gamma((f-1)/2)}, \quad f=N-1 \quad (3-12)$$

이며, 몇 개의 f에 대한  $b_f$ 의 값을 나열하면 다음과 같이 나온다.

f와 $b_f$ 와의 관계								
f	4	9	14	24	29	39	49	59
$b_f$	0.798	0.914	0.945	0.968	0.974	0.981	0.985	0.987

따라서 데이터의 수가 적은 경우  $\hat{C}_p$ 는  $C_p$ 를 과대 평가하게 된다. 구체적으로 소량생산하에서 N=10개의 데이터로  $\hat{C}_p$ 를 구하는 경우  $C_p$ 는 9.4% 정도 과대평가된다. 따라서  $C_p$ 가 실제로 1 (2,700PPM임)인데도 1.094 (520PPM임)라고 말할 가능성이 많다.  $C_p$ 가 1이 아닌 다른 값, 예를 들어 0.5나 1.2인 경우 공정의 불량수준이 얼마나 된다고 잘못 판단하는지에 대해서는 다음과 같은 표를 보면 쉽게 알 수 있다.

표본의 크기가 10인 경우 원래의 공정 불량률 및 과장된 공정불량률

$C_p$	0.5	1	1.2
원래의 공정불량률	133,620PPM	2,700PPM	320PPM
과장된 공정불량률	101,000PPM	520PPM	80PPM

따라서 소표본에 의해 모수를 추정하는 경우에는 올바른 추정을 하기 위하여 불편 추정량

$$\hat{C}_p' = b_f \hat{C}_p \quad (3-13)$$

으로  $C_p$ 를 추정하면 좋을 것이다. 참고로,  $\hat{C}_p'$ 의 분산은 다음과 같이 나온다.

$$\text{Var}(\hat{C}_p') = \left[ \frac{fb_f^2}{f-2} - 1 \right] C_p^2 \quad (3-14)$$

다음으로 추정된  $C_{pk}$ 에 대한 기댓값은

$$E(\hat{C}_{pk}) = \frac{1}{b_f} C_{pk}, \quad (3-15)$$

이므로 소표본에 의한 잘못된 추정을 피하기 위해서는 불편추정량

$$\hat{C}_{pk}' = b_f \hat{C}_{pk} \quad (3-16)$$

으로  $C_{pk}$ 를 추정해야 한다. 참고로,  $\hat{C}_{pk}'$ 의 분산은 다음과 같다.

$$\text{Var}(\hat{C}_{pk}') = (1-k)^2 \left[ \frac{fb_f^2}{f-2} - 1 \right] C_p^2 \quad (3-17)$$

### 3.2.4 품질특성치간 독립성 결여로 인한 공정능력지수의 불확실성

지금까지의 논의는 품질특성치들이 서로 독립적인 경우에 적용된다. 하지만 제조공정의 경우 품질특성치간 독립이 아닌 경우가 많다. 특히 요즘과 같이 품질특성치가 계측기에 의해 자동적으로 측정되는 경우 측정치간 시간간격이 짧으므로 이들간 서로 종속적인 경우가 많다. 그런데도 불구하고 이들 품질특성치들이 서로 독립인 것처럼 기존의 방법대로 공정능력지수를 구하면 어떤 결과를 초래하는지 이 절에서 살펴보기로 한다.

일반적으로  $X_1, X_2, \dots, X_N$ 이 공통의 평균  $\xi$ 와 분산  $\sigma^2$ 을 가지며  $X_i$ 와  $X_j$  간 상관계수가  $\rho_{ij}$ 인 경우, Yang and Hancock (1990)에 의하면, 표본분산  $S^2$ 의 기댓값은

$$E(S^2) = (1 - \bar{\rho})\sigma^2 \quad (3-18)$$

이다. 여기서  $\bar{\rho}$ 란  $N(N-1)/2$ 개의  $X$ 들간의 짝에 대한 상관계수들을 평균한 것이다. 식 (3-18)로부터 상관계수들이 주로 양수인 경우  $\hat{C}_p = (USL - LSL)/6S$ 에서  $\sigma$ 를 과소 평가하게 되며, 따라서  $C_p$ 는 과대 평가된다고 하는 것을 알 수 있다. 특히 품질특성치간 양의 상관관계가 높으면 높을수록  $C_p$ 는 더욱더 과대 평가된다. 품질특성치간 종속적인 경우에는  $C_{pk}$ 도  $C_p$ 에서와 같이 잘못 평가된다.

구체적으로 매 시점의 특성치간 다음과 같은 1차 자기회귀(first order autoregressive) 관계

$$X_{i+1} = \phi X_i + a_{i+1}, \quad i = 1, 2, \dots, N \quad (3-19)$$

가 성립한다고 하자. 여기서  $-1 < \phi < 1$  이며  $a_i$ 는 서로 독립이고  $E(a_i) = 0$ ,  $Var(a_{i+1}) = \sigma_a^2$ ,  $Var(X_i) = \sigma_a^2 / (1 - \phi^2)$  (이것을  $\sigma^2$ 라고 하자)라고 가정한다. 식 (3-19)는 화학공정에서 품질특성치를 일정한 시간간격으로 측정하는 경우 나올 수 있는 모형이다. 이와 같이 부분군이 형성되기 어려운 경우에는 공정의 표준편차  $\sigma$ 를 추정하기 위해 서로 인접한 관측치들간의 이동범위의 평균을 이용한 식 (3-6)이 많이 사용된다. 그런 경우에는

$$\begin{aligned} E(\overline{MR}/d_2) &= \sqrt{1 - Corr(X_i, X_{i+1})} \sigma \\ &= \sqrt{1 - \phi} \sigma \end{aligned} \quad (3-20)$$

인 관계가 성립한다. 따라서 이로부터 연속적인 품질특성치간 양의 상관관계가 있다면  $\sigma$ 를 과소 평가하게 되어,  $C_p$ 는 과대 평가된다. 특히 연속적인 품질특성치간 양의 상관관계가 높을수록  $C_p$ 는 더욱더 과대 평가된다.

식 (3-19)보다 좀더 복잡한 2차 자기회귀모형

$$X_{i+2} = \phi_1 X_{i+1} + \phi_2 X_i + a_{i+2}, \quad i = 1, 2, \dots, N \quad (3-21)$$

이 성립하는 경우에는

$$E(\overline{MR}/d_2) = \sqrt{1 - \frac{\phi_1}{1 - \phi_2}} \sigma \quad (3-22)$$

인 관계가 성립하므로  $X_{i+2}$ 와  $X_i$ 간의 양의 상관관계가  $X_{i+2}$ 과  $X_{i+1}$ 간의 양의 상관관계와 더불어  $C_p$ 를 더욱더 과대 평가한다. (3-19)와 (3-21) 식에서와 같이 품질특성치간 종속적인 경우에는  $C_{pk}$ 도 마찬가지로 과대 평가된다.

$X$ 간의 종속관계는  $\hat{C}_p$ 에 보다는  $\hat{C}_{pm}$ 에 영향을 덜 미친다. 그 이유는  $\rho_{ij} = \text{Corr}(X_i, X_j)$ 이 어떤 값이더라도  $\frac{1}{N} E[\sum_{j=1}^N (X_j - T)^2] = \sigma^2 + (\xi - T)^2$ 이기 때문이다.

공구의 마모와 같이 계통적인 이상요인이 내재하는 경우 공정능력지수는

$$C_{pm}^* = \frac{USL - LSL}{6\sqrt{\sigma_i^2 + (\xi_i - T)^2}} \quad (3-23)$$

와 같이 정의하여 사용할 수도 있다. 여기에서  $\xi_i$  및  $\sigma_i$ 는  $i$ 번째 부분군의 평균 및 표준편차를 나타낸다. 물론  $C_{pm}^*$  이외의 다른 공정능력지수에서도  $\xi$ 나  $\sigma$  대신  $\xi_i$ 나  $\sigma_i$ 를 넣어 공정의 능력을 파악할 수 있다.

#### 4. 복잡한 공정의 경우 공정능력분석 절차

앞의 3절에서는 주어진 데이터가 어떤 분포를 따르는지 또는  $\sigma$ 가 어떻게 추정되는지 등에 따라 공정능력지수에 대한 해석이 얼마나 달라질 수 있는지 살펴보았다. 요즘과 같이 공정이 복잡한 경우 이에 대한 공정능력분석 또한 복잡하다. 예를 들어 제품들이 여러 로트에 들어있는 경우 이 제품들은 여러 생산라인에서 나오므로 제품간 변동이 있고, 각각의 생산라인내 제품들도 시간차 등에 따른 변동이 있으며, 일정한 시점에서 생산된 제품이라고 하더라도 제품간 및 제품내 변동 또한 있다. 따라서 이들중 품질의 특성치에 영향을 미치는 중요한 요인이 무엇인지 파악하여 공정에 대한 능력의 평가를 정확하게 할 필요가 있다.

다음은 복잡한 공정의 경우 품질변동에 영향을 미치는 요인이 무엇인지 살펴보고 공정을 안정화시키는데 유용한 도구들이다.

그림을 통한 분석 : 공정의 상황을 이해하는 데 가장 적절한 방법중의 하나가 그림이다. 특히 생산라인 별, 작업자 별 또는 시간대 별 품질특성치를 조사하여, 이를 그림으로 나타내면 중요 요인들에 특이한 점은 없는지 대략적으로 파악할 수 있다. 이 외



에도 관리도, 다변량 차트 등도 현상파악에 유용하게 이용될 수 있다.

**실험계획법 및 분산분석법** : 실험계획법은 공정의 반응변수에 영향을 미치는 독립변수들의 최적의 상태를 찾는 것 등에 이미 많이 이용되는 방법이며, 분산분석법은 전체 공정의 변동을 세부적인 요인으로 나누는 데 많이 이용된다. 이들 기법은 이미 공정관리에서 없어서는 안될 분석기법으로 자리를 잡고 있다.

**반응표면분석법** : 이 방법은 화학공정 등에 많이 이용되는 방법으로 여러 공정변수들이 수율이나 그 외 알고자 하는 품질특성치에 미치는 영향을 평가하는데 이용되며, 또한 최적의 결과를 낼 수 있도록 공정변수들의 적절한 수준을 효율적으로 찾는 데에도 이용된다.

**EVOP(Evolutionary Operation)법** : 이 방법은 현장에서 실제로 생산을 진행시켜 나가면서 공정의 최적 조건을 찾기 위하여 생산라인을 대상으로 실험할 수 있도록 짜여진 실험계획법이다. 이 실험계획법은 각 공정의 조건을 끊임없이 조금씩 변화시켜가면서 최적조건을 찾으려는 진화적인 공정조작법이다.

이들 기법들을 현장에 계속 적용하다보면 복잡한 공정이라도 문제점을 파악할 수 있으며, 그 결과 공정은 최적 상태에 도달할 수 있다. 이때 공정의 능력을 하나의 수치로 요약하기 위해서는 관심있는 품질특성치에 대하여 앞에서 기술한 공정능력지수를 구하면 된다.

그러면 처음으로 고안되는 복잡한 공정의 경우 어떤 과정을 거쳐 평가하는지 살펴본다 (Bemesderfer(1979) 참조). Bemesderfer에 의하면 다음 과정에 따라 공정을 평가하면 공정에 대한 문제점이 많이 줄어든다고 한다.

- (1) 각 공정변수들이 품질특성치에 미치는 영향을 구체적으로 파악할 수 있도록 수리적인 식을 이용하여 공정의 상황을 기술한다. 이때 유용한 도구로는 반응표면분석법이다.
- (2) 요구품질이 충족되도록 각 공정변수들의 최적공정조건을 결정한다. 어떤 공정이든 최적조건을 찾기 위해서는 잘 짜여진 실험계획이 필요하다.
- (3) 측정과정에 문제가 없는지 확인한다. 측정관리는 계측기 능력분석 (백재욱, 조진남(1996) 참조) 등의 방법을 이용하여 실시한다.
- (4) 공정의 능력을 결정한다. 복잡한 공정의 경우에는 일부 공정변수에 대한 수준이 조금만 달라져도 공정의 능력이 아주 다르게 나타난다. 따라서 일정하게 짜여진 계획에 의해 실험을 실시하고, 그 결과 나온 수치에 의해 공정의 능력을 파악한다.
- (5) 각 품질특성치를 조정할 수 있는 공정변수는 적어도 하나는 되도록 한다. 예를 들어, 유체 온도조절대를 설비에 종종 부착시키는데, 그 이유는 온도가 여러 품질특성치에 영향을 주기 때문이다.
- (6) 공정에 대한 설명서를 준비한다. 이 설명서에는 여러 공정변수를 어떻게 조정하는지 설명이 되어 있어야 한다.
- (7) 공정에 관한 정보를 현장요원에게 전달한다. 어떻게 공정이 이루어져야 하는지에 대해 간단한 소책자를 이용해서 알려줘도 되고, 필요하다면 일정한 틀을 갖춘 세미나

형식으로 알려줘도 된다.

마지막으로 공정능력을 알기 위해서는 다음과 같은 사항들을 확인할 필요가 있다.

- (1) 기계를 선정하는 경우 해당 공정을 대표할 수 있는 기계를 선정해야 한다. 예를 들어, 몇 대의 기계로 여타의 기계에 대한 능력을 유추하는 것이라면 선정된 기계가 다른 모든 기계를 대표할 수 있어야 한다. 한 대의 기계에서 여러 개의 부품이 생산되는 경우도 있으므로 기계의 선정에는 항상 조심해야 한다.
- (2) 공정의 조건을 명시해야 한다. 기계공정의 경우에는 기계의 속도, 냉각제, 사이클 시간 등을 명시해야 하며, 화학공정의 경우 온도, 압력 등을 명시해야 한다. 이와 같은 것이 명시되지 않는다면 공정능력지수는 쓸모 없는 것이 되고 만다. 왜냐하면 다른 공정조건하에서는 공정능력지수가 달리 나오기 때문이다.
- (3) 생산현장에서 일하는 사람들중 숙련자도 아니고 비숙련자도 아닌 보통 근로자를 선택해야 한다. 사람에 따라 일하는 능력에 차이가 있다면 사람에 대해 교육을 시켜 모두 일정한 수준에 도달하도록 해야 한다.
- (4) 원재료를 충분히 공급하여 공정능력에 관한 연구가 중단되지 않도록 해야 한다. 좀더 적은 수의 표본도 무난할 수도 있겠지만, 적은 표본의 수로는 품질특성을 제대로 파악하기 어려울 수 있다. 필요한 표본의 수는 약 50개 정도면 된다.
- (5) 측정관리가 제대로 잘 되었는지 확인해야 한다. 측정관리에 이용되는 여러 방법중 적절한 것 (예를 들어 GR&R 분석)을 선택하여 실시한다.
- (6) 어떤 작업이 이루어지는지 순서대로 기록할 수 있도록 준비한다.
- (7) 생산에 통상적으로 사용되는 재료나 부품을 공정에 투입한다.
- (8) 공정의 능력을 측정하기 위해서는 연구기간 동안 임의로 기계를 조작해서는 안된다.
- (9) 규격허용한계의 1/10까지의 정확도를 유지하면서 측정한다. 물론 0점 처리된 계측기로 측정한다.
- (10) 공정능력 연구기간 동안에는 불필요하게 작업을 중단하지 않는다.
- (11) 품질의 변동에 영향을 주는 것이라면 비정상적인 조건하에서 운용되는 것이라고 하더라도 생산현장을 주시한다.
- (12) 생산조건을 대표할 수 있는 공구를 사용한다.
- (13) 비정상적인 조건 (예를 들어, 공정시작 후 워밍업 기간)하에서 공정이 어떻게 운용되는가 관찰하는 것이 아니라면 정상적인 조건하에서 공정이 운용되고 있다는 것을 확인한다.

## 5. 요약

공정능력지수는 공정이 통계적으로 안정된 상태에서 잴 수 있는 지표이다. 왜냐하

면 통계적으로 관리되어 있지 않은 상태에서 공정능력지수를 구하면 그 당시의 일시적인 공정 상태에 대한 판단은 될지언정 장기적인 관점에서 공정에 대한 판단은 될 수 없기 때문이다. 대표적인 공정능력지수에 대해서는 앞의 2절에서 간단히 설명하고 있다.

공정능력지수는 하나의 숫자로 공정의 상태를 파악할 수 있다는 매력 때문에 인기를 얻어왔다. 사실 공정능력지수 만큼 공정의 향상을 위해 쉽고 간단하게 사용할 수 있는 도구는 없다. 하지만 앞의 3절의 설명에서 알 수 있듯이 여러 가지 통계적인 문제가 해결되지 않으면 공정능력지수의 유용성은 격감된다.

3절에서는 데이터의 정규성의 가정이 들어맞지 않으면 공정능력지수에 대한 해석이 어떻게 달라 지는지,  $\sigma$ 의 추정방법이 변하면 공정능력지수가 어떻게 변하는지, 샘플링오차가 있으면 공정능력지수가 얼마나 불확실하게 나타나는지, 적은 수의 표본으로 공정능력지수를 추정하면 공정능력에 대한 해석이 어떻게 달라지는지, 데이터간 상관관계가 있으면 공정능력에 대한 해석이 어떻게 잘못 나올 수 있는지에 대해 간단히 살펴보았다.

종종 관리자들은 공정관리를 하기 위해 여러 가지 노력을 제대로 하기 보다는 원하는 수준의 공정능력지수 값을 구하기 위하여 헛수고를 하기도 한다. 예를 들어, 공정능력지수의 값이 원하는 수준보다 낮게 나오면 표본을 달리 선택하거나 또는 편의있는 측정을 하여 공정능력지수의 값을 끌어올리려고 한다. 그래서 결국 공정능력지수를 구해 해당 공정의 능력을 평가하는 것이 불행하게도 품질개선에 장애요인이 되기도 한다. 따라서 이러한 폐단을 없애기 위해서는 앞의 4절에서 제시된 일정한 절차에 따라 공정의 능력을 측정해야 한다. 결국 3절에서 언급된 문제점들을 이해한 후 4절에 있는 공정능력분석절차를 따른다면 공정에 대한 보다 근본적인 문제를 파악할 수 있을 것이며, 따라서 공정능력지수를 이용한 공정의 향상도 이루어질 수 있을 것으로 본다.

## 참고문헌

- [1] 박성현, 박영현(1994), 「통계적 품질관리」, 민영사.
- [2] 박창순(1995), 「통계적 품질관리」, 울곡출판사.
- [3] 백재욱, 조진남(1996), “계측기 능력분석과 실험계획법,” 「품질경영학회지」 24권, 3호, pp. 145-159.
- [4] Bemserderfer, J. L.(1979), “Approving Verification of Manufacturing Planning,” *Journal of Quality Technology*, Vol. 11, No. 1, pp. 1-12.
- [5] Boyles, R. A.(1991), “The Taguchi Capability Index,” *Journal of Quality Technology*, Vol. 23, No. 1, pp. 17-26.
- [6] Chan, L. J., Cheng, S. W. and Spring, F. A.(1988), “A New Measure of

- Process Capability:  $C_{pm}$ ," *Journal of Quality Technology*, Vol. 20, No. 3, pp. 162-175.
- [7] Clements, J. A.(1989), "Process Capability Calculations for Non-Normal Distributions," *Quality Progress* Vol. 22, pp. 98-100.
- [8] Deming, W. E.(1982), *Quality, Productivity and Competitive Position*, MIT Center for Advanced Engineering Study, Cambridge, MA.
- [9] Floyd, D. A. and Laurent, C.(1995), "Gauging: An underestimated consideration in the application of statistical process control." *Quality Engineering*, Vol. 8, No. 1, pp. 13-29.
- [10] Gunter, B. H.(1989), "The use and abuse of  $C_{pk}$ ," *Quality Progress*, Vol. 22, January (pp. 72-73), March(pp. 108-109), May(pp. 79-80), July(pp. 86-87).
- [11] Herman, J. T.(1989), "Capability Index-Enough for Process Industries?," *Proceedings, ASQC 43<sup>rd</sup> AQC*.
- [12] Kane, V. E.(1986a), "Process Capability Indices," *Journal of Quality Technology*, Vol. 18, No. 1, pp. 41-52.
- [13] Kane, V. E.(1986b), "Corrigenda," *Journal of Quality Technology*, Vol. 18, No. 4, p. 265.
- [14] Kotz S. and Johnson, N. L.(1993), *Process Capability Indices*, Chapman and Hall.
- [15] Patnaik, P. B.(1950), "The Use of Mean Range as an Estimator of Variance in Statistical Tests," *Biometrika*, Vol. 37, No. 1, pp. 78-87.
- [16] Pearn, W. L., Kotz, S. and Johnson, N. L.(1992), "Distributional and Inferential Properties of Process Capability Indices," *Journal of Quality Technology*, Vol. 24, No. 4, pp. 216-231.
- [17] Quesenberry, C. P.(1993), "The effect of sample size on estimated limits for  $\bar{X}$  and X control charts," *Journal of Quality Technology*, Vol. 25, pp. 237-247.
- [18] Rodriguez, R. N.(1992), "Recent Development in Process Capability Analysis," *Journal of Quality Technology*, Vol. 24, No. 4, pp. 176-187.
- [19] Sullivan, L. P.(1984), "Reducing Variability: A New Approach to Quality," *Quality Progress*, Vol. 17, No. 1, pp. 15-21.
- [20] Sullivan, L. P.(1985), Letters, *Quality Progress*, Vol. 18, No. 1, pp. 7-8.
- [21] Yang, K. and Hancock, W. M.(1990), "Statistical Quality Control for Correlated Samples," *International Journal of Production Research*, Vol. 28, pp. 595-608.