

## 2차원 벡터 공정능력지수 $C_p$ 와 $C_{pk}$ 의 근사 신뢰영역<sup>1)</sup>

박병선\*, 이충훈\*\*, 조중재\*\*\*

\*충북대학교 전자계산학과, \*\*LG전자(주) DRM 사업부, \*\*\*충북대학교 통계학과

### On the Confidence Region of Vector-valued Process Capability Indices $C_p$ & $C_{pk}$

Byoung-sun Park\*, Choong-Hoon Lee\*\*, Joong-Jae Cho\*\*\*

\*Dept. of Computer Science, Chungbuk National University, \*\*LG DRM,

\*\*\*Dept. of Statistics, Chungbuk National University

**Key Words** : Process Capability Index, Asymptotic distribution, Bivariate normal distribution, Confidence region, Plug-in estimator, Monte-carlo experiment

#### Abstract

In this paper we study two vector-valued process capability indices  $C_p = (C_{px}, C_{py})$  and  $C_{pk} = (C_{pkx}, C_{pky})$  considering process capability indices  $C_p$  and  $C_{pk}$ . First, we derive two asymptotic distributions of plug-in estimators  $\widehat{C}_p = (\widehat{C}_{px}, \widehat{C}_{py})$  and  $\widehat{C}_{pk} = (\widehat{C}_{pkx}, \widehat{C}_{pky})$  under some proper conditions. Second, we examine the performance of asymptotic confidence regions of our process capability indices  $C_p = (C_{px}, C_{py})$  and  $C_{pk} = (C_{pkx}, C_{pky})$  under  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$

### 1. 서론

공정능력(process capability)이란 제조공정이 제품의 설계과정에서 설정한대로 얼마나 균일한 제품을 생산할 수 있는지를 반영하는 공정의 고유능력 즉, 균일성을 의미한다. 이처럼 고유능력을 평가하기 위해서 다양한 통

계적 기법들이 제안되어 왔으며, 이를 공정의 변동과 제품의 규격한계 등으로 공정능력을 평가하는 것을 공정능력분석(process capability analysis)이라 하고, 이를 정량적으로 표현한 것이 공정능력지수(process capability index)이다.

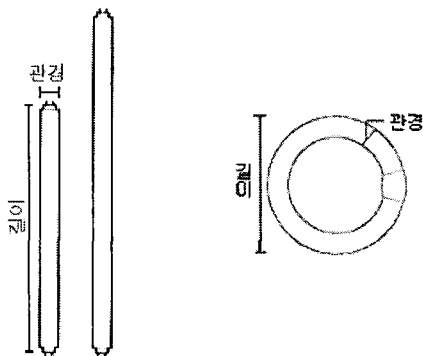
공정능력분석은 통계적 품질관리(statistic-

1) 본 연구는 한국과학재단 목적기초연구(R05-2001-000-00082-0)지원으로 수행되었음

al quality control), 품질경영(total quality management), 나아가 최근 많은 기업에 도입하여 성공을 거두고 있는 6시그마(six sigma)에서도 매우 중요한 영역이라고 할 수 있다.

공정능력분석을 위해 일차원적인 여러 가지 공정능력지수들과 비공정능력지수들에 대하여 활발하게 연구되고 있다(Pearn, Kotz와 Johnson(1992), Kotz와 Johnson(1993), Greenwich, M.와 Jahr-Schaffrath, B.L.(1995)).

일반적으로 하나의 제품은 여러 개의 특성을 갖고 있으며, 이러한 특성들을 기초로 그 제품의 품질을 판단하게 된다. 예를 들면, <그림 1>과 같이 조명에 사용되는 광램프의 경우에는 관경 뿐만 아니라 길이, 무게, 소비전력, 금속의 형상, 스위치특성, 광속유지율, 수명, 광원색, 외관 등 하나의 제품이지만 여러 개의 품질특성을 갖고 있으며, 각 특성들에 대하여 설정된 규격을 만족시킬 경우 그 제품의 품질이 우수하다고 할 수 있을 것이다.



<그림 1> 직관형 형광 램프와 환형 형광 램프

하지만 제품의 품질은 그들 특성 각각의 성능보다는 오히려 결합된 성능에 영향을 받는

다. 실제로 이들 특성을 하나 하나 측정하지만, 특성치들간에 독립적이라기 보다는 오히려 관련이 있어서 결합적으로 분포되어 있다고 가정하는 것이 보다 현실적일 것이다. 이러한 경우에 일변량인 경우의 변동조차 하나의 지수로 표현하는 것은 매우 위험한 일이라고 할 수 있다. 이 때 벡터 공정능력지수를 사용하는 것이 매우 자연스럽고 바람직할 것이다. 또한 자동 검사 시스템에 의해 여러 특성들을 하나 하나 쉽게 측정할 수 있는 오늘날에 다변량 품질관리의 중요성은 더욱 크다고 할 수 있겠다. Alt와 Smith(1988)는 현재 이용가능한 다변량 공정관리 기법들을 훌륭하게 정리·연구하였다. 최근에는 여러 특성들을 기초로 하여 공정능력을 올바르게 평가하는 문제가 중요한 이슈가 되었다. Kocherlakota, S와 Kocherlakota, K(1991)는 이변량 정규분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$  하에서 가장 간단한 공정능력지수의 추정량들인  $\hat{C}_{px}$ 와  $\hat{C}_{py}$ 의 결합 확률분포함수를 계산하고 통계적 추론 방향을 제시하였다. Hubele et al.(1991)는 또 다른 형태의 벡터 공정능력지수를 정의하였는바, 3가지 성분을 가지고 정의된 뜻있는 유용한 생각이다. 그러나 벡터 공정능력지수들에 관한 로버스트한 추정 문제에 대한 연구 결과가 거의 없는 실정이다.

본 논문에서는 두 가지 벡터 공정능력지수들에 대한 통계적 추론을 위한 근사 신뢰영역 설정 문제에 관하여 연구하였다. 먼저 2절에서는 일반적으로 현장에서 많이 이용되는 일차원 공정능력지수들인  $C_p$ 와  $C_{pk}$ 의 개념을 2차원 공정능력지수  $C_p = (C_{px}, C_{py})$ 와

$C_{pk} = (C_{pkx}, C_{pky})$ 로 확장하여 이 지수의 플러그-인 추정량들과 관련된 극한 확률분포

를 유도하고 근사 신뢰영역을 제시하였다. 그리고 3절에서는 이변량 정규분포 하에서 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 95% 신뢰영역에 관하여 다양한 모의실험을 수행함으로써, 본 논문에서 제시한 근사 신뢰영역에 대한 효율성을 포괄적으로 조사·연구하였다.

## 2. 벡터 공정능력지수와 근사 신뢰영역

공정능력분석에 관한 연구는 공정능력지수(Process Capability Index), 비공정능력지수(Process Incapability Index) 그리고 다변량 공정능력지수(Multivariate Process Capability Index) 분야로 나눌 수 있는데, 공교롭게도 가장 단순한 일변량 공정능력지수들과 관련된 연구에 집중되어 있다. 이러한 연구결과 들로는 수없이 많이 있다. 특히, Pearn, Kotz와 Johnson(1992)는 공정능력지수들의 추정량들에 대해 분포적인 문제들을 연구하였다. 또한 Kotz와 Johnson(1993)은 기존 연구결과 들을 로버스트 문제, 다변량 공정능력지수들 까지 체계적으로 정리하였다. Chan, Xiong와 Zhang(1990)는 몇 가지 추정된 공정능력지수들과 관련된 극한분포 결과를 유도하였다. 생산 현장에서 적용할 때의 특수성을 고려하더라도 다변량 공정능력지수, 특히 벡터 공정능력지수 분야에 대한 연구는 활발하지 못한 상황이다.

다행히 Kocherlakota,S.와 Kocherlakota,K.(1991) 등의 벡터 공정능력지수 분야에 관한 연구가 있지만 아직 미흡한 실정이라고 할 수 있다. 즉, 그들은 정규 이변량 자료들을 바탕으로 한 기본적인 이차원 벡터 공정능력지수

$C_p = (C_{px}, C_{py})$ 와 이 지수에 대하여 한 쪽 규격만을 고려한 이차원 벡터 공정능력지수

$C_{pu} = (C_{pux}, C_{puy})$ 의 추정량들의 확률분포함수와 간단한 통계적 추론문제에 대하여 연구하였지만, 너무 복잡하고 제한된 결과로 현장에서 적용하기엔 한계가 있다고 생각된다. 그리고 Wierda,S.(1992) 등의 여러 학자들이 다변량 공정능력지수들에 대하여 연구하였다.

### 2.1 벡터 공정능력지수의 추정

이론적으로나 생산 현장에서의 응용을 고려할 때, 가장 기본적이고 많이 사용되고 있는 일차원 공정능력지수인

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

와

$$C_{pk} = \frac{\text{Min}(\mu - LSL, USL - \mu)}{3\sigma} = \frac{d - |\mu - M|}{3\sigma}$$

들을 기초로 하여 확장한 2차원 벡터 공정능력지수  $C_p = (C_{px}, C_{py})$ 와  $C_{pk} = (C_{pkx}, C_{pky})$ 에 대하여 점추정량과 관련된 극한 분포이론 등을 연구하여, 이를 바탕으로 보다 간편하고 훨씬 유용한 근사 신뢰영역에 대하여 논의하게 될 것이다.

우선, 각각의 평균이  $\mu_x$ 과  $\mu_y$ 이고 분산이  $\sigma_x^2$ 과  $\sigma_y^2$ 이며, 4차중심적률  $\mu_{4x} = E(X - \mu_x)^4$ 과  $\mu_{4y} = E(Y - \mu_y)^4$ 이 유한인 이변량 모집단으로부터의 공정특성 확률표본  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 에 대하여, 대표적인 이변량 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 은 다음과 같이 정의된다(Kocherlakota,S.와 Kocherla-

kota,K.(1991) 등).

$$C_p = (C_{px}, C_{py}) = \left( \frac{d_x}{3\sigma_x}, \frac{d_y}{3\sigma_y} \right)$$

$$C_{pk} = (C_{pkx}, C_{pky})$$

$$= \left( \frac{d_x - |\mu_x - M_x|}{3\sigma_x}, \frac{d_y - |\mu_y - M_y|}{3\sigma_y} \right)$$

단,  $USL_x, LSL_x$ 와  $USL_y, LSL_y$ 는 각각 이변량 공정  $(X, Y)$ 에 대한 규격상한과 규격하한을 나타내며,  $d_x$ 와  $d_y$  및  $M_x$ 와  $M_y$ 는 각각 다음을 나타낸다.

$$d_x = \frac{USL_x - LSL_x}{2}, \quad d_y = \frac{USL_y - LSL_y}{2},$$

$$M_x = \frac{USL_x + LSL_x}{2}, \quad M_y = \frac{USL_y + LSL_y}{2}$$

그리고 이들의 추정량  $\widehat{C}_p$ 와  $\widehat{C}_{pk}$ 은 플러그인 방법(Plug-in method)에 의해 다음과 같이 추정될 수 있을 것이다.

$$\widehat{C}_p = (\widehat{C}_{px}, \widehat{C}_{py})$$

$$= \left( \frac{d_x}{3S_x}, \frac{d_y}{3S_y} \right)$$

$$\widehat{C}_{pk} = (\widehat{C}_{pkx}, \widehat{C}_{pky})$$

$$= \left( \frac{d_x - |\bar{X} - M_x|}{3S_x}, \frac{d_y - |\bar{Y} - M_y|}{3S_y} \right)$$

단, 표본평균  $\bar{X}$ 와  $\bar{Y}$ 는

$$\hat{\mu}_x = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$\hat{\mu}_y = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

로 정의하고, 표본분산  $S_x^2$ 와  $S_y^2$ 은 다음과 같이 정의한다.

$$\hat{\sigma}_x^2 = S_x^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

$$\hat{\sigma}_y^2 = S_y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

## 2.2 추정량과 관련된 극한분포

먼저 벡터 공정능력지수  $C_p = (C_{px}, C_{py})$ 와  $C_{pk} = (C_{pkx}, C_{pky})$ 들의 추정량과 관련된 극한분포 이론을 확립하기 위하여 중요한 보조정리를 소개하면 다음과 같다.

### [보조정리1]

만약 4차 중심적률  $\mu_{4x} = E(X - \mu_x)^4$ 과  $\mu_{4y} = E(Y - \mu_y)^4$ 이 존재하면, 표본의 크기  $n$ 이  $n \rightarrow \infty$ 일 때, 다음의 결과가 성립한다.

$$(Z_{1n}, Z_{2n}, Z_{3n}, Z_{4n})$$

$$= (\sqrt{n}(\bar{X} - \mu_x), \sqrt{n}(\bar{Y} - \mu_y), \sqrt{n}(S_x^2 - \sigma_x^2), \sqrt{n}(S_y^2 - \sigma_y^2))$$

$$\xrightarrow{d} (Z_1, Z_2, Z_3, Z_4) \stackrel{d}{=} MN(\mathbf{0}, \Sigma_{4 \times 4})$$

단, 공분산 행렬  $\Sigma_{4 \times 4} = (\sigma_{ij})_{4 \times 4}$ 은 다음과 같은 각각의 성분을 갖는다.

$$\sigma_{11} = Var(X) = \sigma_x^2,$$

$$\sigma_{12} = Cov(X, Y),$$

$$\sigma_{13} = E(X - \mu_x)^3 = \mu_{3x},$$

$$\sigma_{14} = Cov(X, (Y - \mu_y)^2),$$

$$\sigma_{22} = Var(Y) = \sigma_y^2,$$

$$\sigma_{23} = Cov((X - \mu_x)^2, Y),$$

$$\sigma_{24} = E(Y - \mu_y)^3 = \mu_{3y},$$

$$\sigma_{33} = \mu_{4x} - \sigma_x^4,$$

$$\sigma_{34} = Cov((X - \mu_x)^2, (Y - \mu_y)^2),$$

$$\sigma_{44} = \mu_{4y} - \sigma_y^4$$

**증명** Cho & Park(2002) 참조 □

물론, Kocherlakota,S와 Kocherlakota,K (1991)는 상관계수  $\rho$ 를 갖는 이변량 정규분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 하에서 가장 간단한 공정능력지수의 추정량들인  $\hat{C}_{px}$ 와  $\hat{C}_{py}$ 의 결합 확률분포함수를 계산하여, 벡터 공정능력지수  $C_p$ 에 대한 정확한 신뢰영역을 구하고자 노력하였지만, 현실적으로 활용가능성이 적은 신뢰영역이라 사료된다. 또한 벡터 공정능력지수  $C_{pk}$ 에 대한 정확한 신뢰영역을 구하는 것은 불가능한 일이라 생각된다.

보다 간편하고 바람직한 신뢰영역을 설정하기 위하여 위의 [보조정리1]을 기초로 벡터 공정능력지수  $C_p$ 의 플러그-인 방법에 의한 점추정량  $\hat{C}_p$ 와 관련된 극한분포를 소개하면 다음과 같다.

**[정리1]**

만일 4차 중심적률  $\mu_{4x} = E(X - \mu_x)^4$ 과  $\mu_{4y} = E(Y - \mu_y)^4$ 이 존재한다면, 표본의 크기  $n$ 이  $n \rightarrow \infty$ 일 때, 다음의 결과가 성립한다.

$$\begin{aligned} & \sqrt{n} (\hat{C}_p - C_p) \\ &= \sqrt{n} (\hat{C}_{px} - C_{px}, \hat{C}_{py} - C_{py}) \\ & \xrightarrow{d} \left( -\frac{d_x}{6\sigma_x^3} Z_3, -\frac{d_y}{6\sigma_y^3} Z_4 \right) \equiv BN(0, V_p) \end{aligned}$$

단,  $(Z_3, Z_4) \sim N(0, \Sigma_{22})$ ,  $\Sigma_{22} = \begin{pmatrix} \sigma_{33} & \sigma_{34} \\ \sigma_{34} & \sigma_{44} \end{pmatrix}$ 이고,

분산-공분산행렬  $V_p = \begin{pmatrix} \sigma_{p11} & \sigma_{p12} \\ \sigma_{p21} & \sigma_{p22} \end{pmatrix}$ 의 각 성분

은 다음과 같이 계산 정리된다.

$$\begin{aligned} \sigma_{p11} &= \frac{d_x^2}{36\sigma_x^6} (\mu_{4x} - \sigma_x^4), \\ \sigma_{p12} &= \sigma_{p21} \end{aligned}$$

$$= \frac{d_x d_y}{36\sigma_x^3 \sigma_y^3} Cov((X - \mu_x)^2, (Y - \mu_y)^2))$$

$$\sigma_{p22} = \frac{d_y^2}{36\sigma_y^6} (\mu_{4y} - \sigma_y^4)$$

**증명** Cho & Park(2002)참조 □

또 하나의 중요한 벡터 공정능력지수  $\hat{C}_{pk}$ 와 관련된 극한분포 이론을 유도하기 위해서는  $\mu_x$ 와  $M_x$ 의 크기 및  $\mu_y$ 와  $M_y$ 의 크기에 따라 다음과 같은 9가지의 경우를 고려해야 할 것이다.

- ㉠  $\mu_x = M_x, \mu_y < M_y$     ㉡  $\mu_x = M_x, \mu_y = M_y$
- ㉢  $\mu_x = M_x, \mu_y > M_y$     ㉣  $\mu_x < M_x, \mu_y < M_y$
- ㉤  $\mu_x < M_x, \mu_y = M_y$     ㉥  $\mu_x < M_x, \mu_y > M_y$
- ㉦  $\mu_x > M_x, \mu_y < M_y$     ㉧  $\mu_x > M_x, \mu_y = M_y$
- ㉨  $\mu_x > M_x, \mu_y > M_y$

위의 각 경우에 대하여 유도하여 정리하면 다음과 같다.

**[정리2]**

만일 4차 중심적률  $\mu_{4x} = E(X - \mu_x)^4$ 과  $\mu_{4y} = E(Y - \mu_y)^4$ 이 존재한다면, 표본의 크기  $n$ 이  $n \rightarrow \infty$ 일 때, 다음의 결과가 성립한다.

$$\begin{aligned} & \sqrt{n} (\hat{C}_{pk} - C_{pk}) \\ &= \sqrt{n} (\hat{C}_{pkx} - C_{pkx}, \hat{C}_{pky} - C_{pky}) \end{aligned}$$

$\xrightarrow{d}$

$$\left\{ \begin{array}{l} \left( -\frac{d_x Z_3}{6\sigma_x^3} - \frac{|Z_1|}{3\sigma_x}, \right. \\ \left. \frac{(|M_y - \mu_y| - d_y)Z_4}{6\sigma_y^3} + \frac{\rho(M_y, \mu_y, Z_2)Z_2}{3\sigma_y} \right), \\ \mu_x = M_x \\ \left( \frac{(|M_x - \mu_x| - d_x)Z_3}{6\sigma_x^3} + \frac{\rho(M_x, \mu_x, Z_1)Z_1}{3\sigma_x}, \right. \\ \left. -\frac{d_y Z_4}{6\sigma_y^3} - \frac{|Z_2|}{3\sigma_y} \right), \\ \mu_y = M_y \\ \left( \frac{(|M_x - \mu_x| - d_x)Z_3}{6\sigma_x^3} + \frac{\text{sgn}(M_x - \mu_x)Z_1}{3\sigma_x}, \right. \\ \left. \frac{(|M_y - \mu_y| - d_y)Z_4}{6\sigma_y^3} + \frac{\text{sgn}(M_y - \mu_y)Z_2}{3\sigma_y} \right), \\ \mu_x \neq M_x \ \& \ \mu_y \neq M_y \end{array} \right.$$

단,

$$\rho(M, \mu, Z) = \begin{cases} -\text{sgn}(Z), & M = \mu \\ \text{sgn}(M - \mu), & M \neq \mu \end{cases}$$

**증명** 위에서 언급된 9가지 경우의 각각에 대하여  $\sqrt{n}(\widehat{C}_{pk} - C_{pk})$ 의 극한 확률분포를 유도하기로 하자.

우선  $\mu_x = M_x$ 이고  $\mu_y < M_y$  인 경우에 대하여 유도하면 다음과 같다.

$$\begin{aligned} & \sqrt{n}(\widehat{C}_{pk} - C_{pk}) \\ &= \sqrt{n}((\widehat{C}_{pkx} - C_{pkx}), (\widehat{C}_{pky} - C_{pky})) \\ &= \left( \sqrt{n} \left( \frac{d_x - |\bar{X} - M_x|}{3S_x} - \frac{d_x}{3\sigma_x} \right), \right. \\ & \quad \left. \sqrt{n} \left( \frac{d_y - |\bar{Y} - M_y|}{3S_y} - \frac{d_y - (M_y - \mu_y)}{3\sigma_y} \right) \right) \\ &= \left( \frac{\sqrt{n}(-d_x(S_x - \sigma_x) - \sigma_x|\bar{X} - \mu_x|)}{3S_x\sigma_x}, \right. \\ & \quad \left. \frac{\sqrt{n}(-d_y(S_y - \sigma_y) - \sigma_y|\bar{Y} - M_y| + S_y(M_y - \mu_y))}{3S_y\sigma_y} \right) \\ &= \left( \frac{\sqrt{n}}{3S_x\sigma_x} \left\{ -d_x \left( \frac{S_x^2 - \sigma_x^2}{S_x + \sigma_x} \right) - \sigma_x |\sqrt{n}(\bar{X} - \mu_x)| \right\}, \right. \\ & \quad \left. \frac{\sqrt{n}}{3S_y\sigma_y} \left\{ -d_y \left( \frac{S_y^2 - \sigma_y^2}{S_y + \sigma_y} \right) + |\bar{Y} - M_y|(S_y - \sigma_y) \right. \right. \\ & \quad \left. \left. - S_y(|\bar{Y} - M_y| - (M_y - \mu_y)) \right\} \right) \end{aligned}$$

$$\begin{aligned} &= \left( \frac{\sqrt{n}}{3S_x\sigma_x} \left\{ -d_x \left( \frac{S_x^2 - \sigma_x^2}{S_x + \sigma_x} \right) - \sigma_x |\sqrt{n}(\bar{X} - \mu_x)| \right\}, \right. \\ & \quad \left. \frac{\sqrt{n}}{3S_y\sigma_y} \left\{ -d_y \left( \frac{S_y^2 - \sigma_y^2}{S_y + \sigma_y} \right) + |\bar{Y} - M_y| \left( \frac{S_y^2 - \sigma_y^2}{S_y + \sigma_y} \right) \right. \right. \\ & \quad \left. \left. - S_y \frac{(\bar{Y} - \mu_y)(\bar{Y} + \mu_y) - 2M_y(\bar{Y} - \mu_y)}{|\bar{Y} - M_y| + (M_y - \mu_y)} \right\} \right) \end{aligned}$$

위의 확률변수는 Slutsky 정리에 의해 다음의 극한 확률분포를 갖는다. 즉,

$$\sqrt{n}(\widehat{C}_{pk} - C_{pk})$$

$$\rightarrow \left( -\frac{d_x Z_3}{6\sigma_x^3} - \frac{|Z_1|}{3\sigma_x}, \frac{(M_y - \mu_y - d_y)}{6\sigma_y^3} Z_4 + \frac{Z_2}{3\sigma_y} \right)$$

다음으로,  $\mu_x < M_x$ 이고  $\mu_y < M_y$ 인 경우에 대하여 유도하면 다음과 같다.

$$\begin{aligned} & \sqrt{n}(\widehat{C}_{pk} - C_{pk}) \\ &= \sqrt{n}((\widehat{C}_{pkx} - C_{pkx}), (\widehat{C}_{pky} - C_{pky})) \\ &= \left( \sqrt{n} \left( \frac{d_x - |\bar{X} - M_x|}{3S_x} - \frac{d_x - (M_x - \mu_x)}{3\sigma_x} \right), \right. \\ & \quad \left. \sqrt{n} \left( \frac{d_y - |\bar{Y} - M_y|}{3S_y} - \frac{d_y - (M_y - \mu_y)}{3\sigma_y} \right) \right) \\ &= \left( \frac{\sqrt{n}(-d_x(S_x - \sigma_x) - \sigma_x|\bar{X} - \mu_x| + S_x(M_x - \mu_x))}{3S_x\sigma_x}, \right. \\ & \quad \left. \frac{\sqrt{n}(-d_y(S_y - \sigma_y) - \sigma_y|\bar{Y} - M_y| + S_y(M_y - \mu_y))}{3S_y\sigma_y} \right) \\ &= \left( \frac{\sqrt{n}}{3S_x\sigma_x} \left\{ -d_x \left( \frac{S_x^2 - \sigma_x^2}{S_x + \sigma_x} \right) + |\bar{X} - M_x|(S_x - \sigma_x) \right. \right. \\ & \quad \left. \left. - S_x(|\bar{X} - M_x| - (M_x - \mu_x)) \right\}, \right. \\ & \quad \left. \frac{\sqrt{n}}{3S_y\sigma_y} \left\{ -d_y \left( \frac{S_y^2 - \sigma_y^2}{S_y + \sigma_y} \right) + |\bar{Y} - M_y|(S_y - \sigma_y) \right. \right. \\ & \quad \left. \left. - S_y(|\bar{Y} - M_y| - (M_y - \mu_y)) \right\} \right) \\ &= \left( \frac{\sqrt{n}}{3S_x\sigma_x} \left\{ -d_x \left( \frac{S_x^2 - \sigma_x^2}{S_x + \sigma_x} \right) + |\bar{X} - M_x| \left( \frac{S_x^2 - \sigma_x^2}{S_x + \sigma_x} \right) \right. \right. \\ & \quad \left. \left. - S_x \frac{(\bar{X} - \mu_x)(\bar{X} + \mu_x) - 2M_x(\bar{X} - \mu_x)}{|\bar{X} - M_x| + (M_x - \mu_x)} \right\}, \right. \\ & \quad \left. \frac{\sqrt{n}}{3S_y\sigma_y} \left\{ -d_y \left( \frac{S_y^2 - \sigma_y^2}{S_y + \sigma_y} \right) + |\bar{Y} - M_y| \left( \frac{S_y^2 - \sigma_y^2}{S_y + \sigma_y} \right) \right. \right. \\ & \quad \left. \left. - S_y \frac{(\bar{Y} - \mu_y)(\bar{Y} + \mu_y) - 2M_y(\bar{Y} - \mu_y)}{|\bar{Y} - M_y| + (M_y - \mu_y)} \right\} \right) \end{aligned}$$

위의 확률변수는 역시 Slutsky 정리에 의해 다음의 극한확률분포를 갖는다. 즉,

$$\sqrt{n} (\widehat{C}_{pk} - C_{pk})$$

$$\rightarrow \left( -\frac{(M_x - \mu_x - d_x)}{6\sigma_x^3} Z_3 + \frac{Z_1}{3\sigma_x}, \right.$$

$$\left. \frac{(M_y - \mu_y - d_y)}{6\sigma_y^3} Z_4 + \frac{Z_2}{3\sigma_y} \right)$$

그리고 나머지 7가지 경우에 대해서도 비슷한 방법으로 계산·정리하면 [정리2]의 결과를 얻을 수 있다. □

### 2.3 벡터 공정능력지수의 근사 신뢰영역

Kocherlakota,S와 Kocherlakota,K(1991)는 이변량 정규분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$  하에서 가장 간단한 공정능력지수의 추정량들인  $\widehat{C}_{px}$ 와  $\widehat{C}_{py}$ 의 결합 확률분포함수를 계산하고 통계적 추론 방향을 제시하였다. 하지만 이를 기초로한 신뢰영역의 설정 과정은 매우 복잡하다고 할 수 있으며, 또한 이변량 정규분포 가정 하에서 조차도 공정능력지수  $\widehat{C}_{pkx}$ 와  $\widehat{C}_{pky}$ 의 결합 확률분포함수를 유도하여 벡터 공정능력지수  $C_{pk}$ 에 대한 신뢰영역을 설정하는 문제는 불가능하리라고 사료되어 비효율적이라고 할 수 있다. 따라서 이 절에서는 앞 절의 극한분포 결과를 이용한 기본적인 2차원 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 보다 바람직한  $100(1-a)\%$  근사 신뢰영역에 대하여 구체적으로 설정하여 제시하고자 한다.

만일 4차 중심적률  $\mu_{4x} = E(X - \mu_x)^4$ 과  $\mu_{4y} = E(Y - \mu_y)^4$ 이 존재하는 적당한 이변량 확률분포로부터 확률표본  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 을 얻었다고 가정할 때, 앞에서 연구된 결과로부터 두 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 근사 신뢰영역을 다음과 같이 설정할 수 있을 것이다.

첫째, 이차원 벡터 공정능력지수  $C_p$ 에 대한  $100(1-a)\%$  신뢰영역은 다음과 같다.

$$n(\widehat{C}_p - C_p)'(\widehat{V}_p)^{-1}(\widehat{C}_p - C_p) \leq \chi^2(2; a)$$

단,  $\chi^2(2; a)$ 는 자유도가 2인  $\chi^2$ -분포의 하위  $100(1-a)$  백분위수를 나타내고, 또한 추정행렬  $\widehat{V}_p$ 는 행렬  $V_p$ 의 추정량으로 다음과 같다.

$$\widehat{V}_p = \begin{pmatrix} \widehat{\sigma}_{p11} & \widehat{\sigma}_{p12} \\ \widehat{\sigma}_{p21} & \widehat{\sigma}_{p22} \end{pmatrix}$$

단,

$$\widehat{\sigma}_{p11} = \frac{d_x^2}{36 \widehat{\sigma}_x^6} (\widehat{\mu}_{4x} - \widehat{\sigma}_x^4),$$

$$\widehat{\sigma}_{p12} = \frac{d_x d_y}{36 \widehat{\sigma}_x^3 \widehat{\sigma}_y^3} \widehat{Cov}((X - \mu_x)^2, (Y - \mu_y)^2)$$

$$\widehat{\sigma}_{p22} = \frac{d_y^2}{36 \widehat{\sigma}_y^6} (\widehat{\mu}_{4y} - \widehat{\sigma}_y^4).$$

추정행렬  $\widehat{V}_p$ 에 있는 각 성분들은 플러그인 방법에 의한 추정량으로 쉽게 계산된다.

둘째, 이차원 벡터 공정능력지수  $C_{pk}$ 에 대한  $100(1-a)\%$  신뢰영역은 다음과 같다.

$$n(\widehat{C}_{pk} - C_{pk})'(\widehat{V}_{pk})^{-1}(\widehat{C}_{pk} - C_{pk}) \leq \chi^2(2; a)$$

단, 추정행렬  $\widehat{V}_{pk}$ 는 행렬  $V_{pk}$ 의 추정량으로 앞서 제시한 9가지 경우에 따라 구체적으로 표현될 수 있을 것이다.

예를 들어, 공정 품질특성치들이 이변량 정규분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 로 잘 적합된다면,  $\mu_x = M_x$ 이고  $\mu_y < M_y$ 인 경우에 [정리2]에서 다음의 극한 확률변수

$$\left( -\frac{d_x Z_3}{6\sigma_x^3} - \frac{|Z_1|}{3\sigma_x}, \frac{(|M_y - \mu_y| - d_y) Z_4}{6\sigma_y^3} + \frac{Z_2}{3\sigma_y} \right)$$

를 고려하여 추정행렬  $\widehat{V}_{pk}$ 는 행렬  $V_{pk}$ 의 추정량으로 다음과 같이 계산된다.

$$\widehat{V}_{pk} = \begin{pmatrix} \widehat{\sigma}_{pk11} & \widehat{\sigma}_{pk12} \\ \widehat{\sigma}_{pk21} & \widehat{\sigma}_{pk22} \end{pmatrix}$$

단,

$$\begin{aligned} \widehat{\sigma}_{pk11} &= \widehat{Var}\left(-\frac{d_x Z_3}{6\widehat{\sigma}_x^3} - \frac{|Z_1|}{3\widehat{\sigma}_x}\right) \\ &= \frac{d_x^2}{36\widehat{\sigma}_x^6} (\widehat{\mu}_{4x} - \widehat{\sigma}_x^2) + \frac{\pi-2}{9\pi} \\ &= \frac{d_x^2}{36\widehat{\sigma}_x^6} 2\widehat{\sigma}_x^2 + \frac{\pi-2}{9\pi} \\ &= \frac{d_x^2}{18\widehat{\sigma}_x^2} + \frac{\pi-2}{9\pi} \end{aligned}$$

$$\begin{aligned} \widehat{\sigma}_{pk22} &= \widehat{Var}\left(\frac{(|M_y - \mu_y| - d_y)}{6\widehat{\sigma}_y^3} Z_4 + \frac{Z_2}{3\widehat{\sigma}_y}\right) \\ &= \frac{(M_y - \bar{Y} - d_y)^2}{36\widehat{\sigma}_y^6} 2\widehat{\sigma}_y^4 + \frac{1}{9} \\ &= \frac{(M_y - \bar{Y} - d_y)^2}{18\widehat{\sigma}_y^2} + \frac{1}{9} \end{aligned}$$

$$\begin{aligned} \widehat{\sigma}_{pk12} &= \widehat{\sigma}_{pk21} \\ &= -\frac{d_x(M_y - \bar{Y} - d_y)}{36\widehat{\sigma}_x^3\widehat{\sigma}_y^3} \widehat{Cov}((X - \mu_x)^2, (Y - \mu_y)^2) \\ &\quad - \frac{d_x}{18\widehat{\sigma}_x^3\widehat{\sigma}_y} \widehat{Cov}((X - \mu_x)^2, Y) \\ &= -\frac{d_x(M_y - \bar{Y} - d_y)}{36\widehat{\sigma}_x^3\widehat{\sigma}_y^3} 2\widehat{\rho}^2\widehat{\sigma}_x^2\widehat{\sigma}_y^2 \\ &= -\frac{d_x(M_y - \bar{Y} - d_y)}{18\widehat{\sigma}_x\widehat{\sigma}_y} \widehat{\rho}^2 \end{aligned}$$

물론 분산  $\sigma_x^2$ 과  $\sigma_y^2$ 의 추정량은 앞에서 정의된 표본분산  $S_x^2$ 과  $S_y^2$ 이고, 4차 중심적률  $\mu_{4x}$ 의 추정량은 다음과 같다.

$$\widehat{\mu}_{4x} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4$$

또한 상관계수  $\rho$ 의 추정량은 표본상관계수  $\widehat{\rho}$ 을 의미한다.

비슷한 방법으로  $\mu_x < M_x$ 이고  $\mu_y < M_y$ 인 경우에 대한 극한 확률변수

$$\left( \frac{(M_x - \mu_x - d_x)Z_3}{6\widehat{\sigma}_x^3} + \frac{Z_1}{3\widehat{\sigma}_x}, \frac{(M_y - \mu_y - d_y)Z_4}{6\widehat{\sigma}_y^3} + \frac{Z_2}{3\widehat{\sigma}_y} \right)$$

를 고려하여 추정행렬  $\widehat{V}_{pk}$ 는 행렬  $V_{pk}$ 의 추정량으로 다음과 같이 계산된다.

$$\widehat{V}_{pk} = \begin{pmatrix} \widehat{\sigma}_{pk11} & \widehat{\sigma}_{pk12} \\ \widehat{\sigma}_{pk21} & \widehat{\sigma}_{pk22} \end{pmatrix}$$

단,

$$\begin{aligned} \widehat{\sigma}_{pk11} &= \frac{1}{9} + \frac{(M_x - \bar{X} - d_x)^2}{18\widehat{\sigma}_x^2}, \\ \widehat{\sigma}_{pk12} &= \widehat{\sigma}_{pk21} \\ &= \frac{(M_x - \bar{X} - d_x)(M_y - \bar{Y} - d_y)}{18\widehat{\sigma}_x\widehat{\sigma}_y} \widehat{\rho}^2 + \frac{\widehat{\rho}}{9}, \\ \widehat{\sigma}_{pk22} &= \frac{1}{9} + \frac{(M_y - \bar{Y} - d_y)^2}{18\widehat{\sigma}_y^2}. \end{aligned}$$

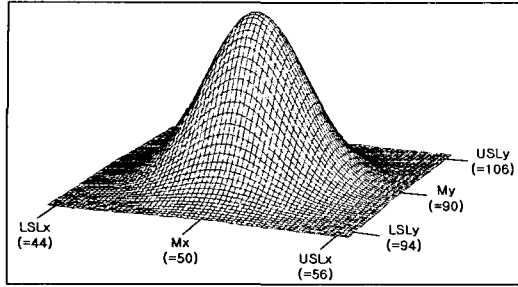
### 3. 모의실험 연구

본 절에서는 2차원 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 신뢰영역 추정 문제를 특히 실제 현장에서 일반적으로 가정되고 있는 이변량 정규분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$  하에서 다양한 유형의 데이터를 발생, 적용함으로써 2차원 벡터 공정능력지수들의 현장 적용 가능성과 실용성을 고려하고자 한다.

#### 3.1 모의실험 설계

2차원 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 신뢰영역 추정 문제에 관한 모의실험을 수행함에 있어, 먼저 이변량 공정  $(X, Y)$ 에 대하여 설정된 규격은 <그림 2>와 같다.





<그림 2> (X, Y)에 대한 규격 설정

그리고 공정능력지수와 시그마 수준과의 관계를 고려하여 이변량 공정 (X, Y)의 시그마 수준이 각각 3, 4.5 및 6시그마 수준의 조합이 되도록 평균  $\mu_x$ 와  $\mu_y$  및 표준편차  $\sigma_x$ 와  $\sigma_y$ 을 선택하여 설계하였다. 또한 상관계수에 따른 다양한 공정 상태를 고려하기 위하여 상관계수  $\rho$ 의 값을 0,  $\pm 0.3$ ,  $\pm 0.6$ ,  $\pm 0.9$ 로 하였으며, 각 쌍의  $(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$ 에 대하여, 표본의 크기  $n$ 은 20, 30, 60을 사용하여 다양한 조건별로 모의실험을 수행하였다.

특히 2차원 벡터 공정능력지수  $C_{pk}$ 에 대해서는 앞 절에서 설명한 9가지 경우 중에서 대표적으로  $\mu_x < M_x$ 이고  $\mu_y < M_y$ 인 경우와  $\mu_x = M_x$ 이고  $\mu_y < M_y$ 인 경우에 대해서만 모의실험을 수행하였다. 2차원 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 세부적인 모의실험 설계는 다음의 <표 1>~<표 3>과 같다.

<표 1>  $C_p$ 에 대한 모의실험 설계

$(\mu_x, \sigma_x)$	$C_{px}$	$(\mu_y, \sigma_y)$	$C_{py}$	(X, Y)의 시그마수준
(50, 2.0)	1.0	(100, 2)	1.0	(3.0, 3.0)
(50, 2.0)	1.0	(100, 1)	2.0	(3.0, 6.0)
(50, 1.0)	2.0	(100, 2)	1.0	(6.0, 3.0)
(50, 1.0)	2.0	(100, 1)	2.0	(6.0, 6.0)

<표 2>  $C_{pk}$ 에 대한 모의실험 설계 ( $\mu_x < M_x$ 이고  $\mu_y < M_y$ 인 경우)

$(\mu_x, \sigma_x)$	$C_{pkx}$	$(\mu_y, \sigma_y)$	$C_{pky}$	(X, Y)의 시그마수준
(47.0, 1.0)	1.0	(97.0, 1.0)	1.0	(4.5, 4.5)
(47.0, 1.0)	1.0	(98.5, 1.0)	1.5	(4.5, 6.0)
(48.5, 1.0)	1.5	(97.0, 1.0)	1.0	(6.0, 4.5)
(48.5, 1.0)	1.5	(98.5, 1.0)	1.5	(6.0, 6.0)

<표 3>  $C_{pk}$ 에 대한 모의실험 설계 ( $\mu_x = M_x$ 이고  $\mu_y < M_y$ 인 경우)

$(\mu_x, \sigma_x)$	$C_{pkx}$	$(\mu_y, \sigma_y)$	$C_{pky}$	(X, Y)의 시그마수준
(50, 2.0)	1.0	(97.0, 1)	1.0	(4.5, 4.5)
(50, 2.0)	1.0	(98.5, 1)	1.5	(4.5, 6.0)
(50, 4/3)	1.5	(97.0, 1)	1.0	(6.0, 4.5)
(50, 4/3)	1.5	(98.5, 1)	1.5	(6.0, 6.0)

본 모의실험은 SAS Ver 8.1을 사용하여 수행하였으며, 특히 행렬연산을 위하여 SAS/IML 및 시각적 효과를 위한 그래프 작성을 위하여 SAS/GRAPH를 이용하였다.

### 3.2 모의실험 절차

이변량 정규분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 의 각 모수들과 표본크기  $n$ 에 의해서 이루어지는 한 번의 모의실험 절차는 다음과 같다.

**[1단계]** 상관계수가  $\rho$ 인 이변량 정규공정 분포  $BN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 로부터 크기가  $n(=20, 30, 60)$ 인 원래의 이변량 표본들을 생성한다.

**[2단계]** [1단계]에서 생성된 표본들을 이용하여 벡터 공정능력지수  $C$ 에 대한 극한분포를 활용한 95% 신뢰영역을 설정한다.

[3단계] [2단계]에서 설정된 신뢰영역 안에 벡터공정능력지수  $C$ 의 참값이 포함되는 지를 결정한다.

위의 3단계로 이루어진 한 번의 모의실험 절차를  $N(=1,000)$ 번 반복한다. 이로부터 95% 신뢰영역 안에 실제 벡터 공정능력지수  $C$ 의 참값이 포함되는 횟수의 비율을 계산 할 것이다.

### 3.3 모의실험 결과 분석

앞 절에서 소개한 모의실험 절차에 의하여 2차원 벡터 공정능력지수  $C_{\beta}$ 와  $C_{pk}$ 에 대한 95% 신뢰영역의 포함비율을 구한 결과가 <표 4>~<표 6>에 제시되어 있다. 결과에서 사용된 '\*' 표시는 포함비율에 대한 99% 신뢰구간에 해당하는 범위 (0.933, 0.967)을 벗어나

는 경우를 의미한다.

<표 4>는  $C_{\beta}$ 에 대한 95% 신뢰영역의 포함비율 수행결과이다. 먼저  $\rho$ 의 변화에 따른 결과를 살펴보면,  $\rho=-0.3$ 인 경우에 대하여 일관성 있게 기대값 0.95에 가까운 값을 갖는 것으로 나타났다. 실제로 이 경우에 대한 12개의 포함비율 모두가 범위 내에 있는 것으로 나타나 수행결과가 매우 좋음을 알 수 있다. 다음으로는  $\rho$ 의 값인 -0.6, 0, 0.3, 0.6, 0.9, -0.9인 경우의 순으로 수행결과가 좋은 것으로 나타났다. 특히  $\rho=\pm 0.9$ 인 경우, 이에 대한 24개의 포함비율 중 13개가 범위를 벗어나고 있는 것으로 나타났다. 다음으로 표본크기  $n$ 에 따른 수행결과를 살펴보면,  $n$ 이 증가할수록 기대값 0.95에 가까운 값을 갖는 것으로 나타났다. 실제로  $n=20$ 인 경우에 대한 28개의 포함비율 중 15개가 범위를 벗어나고 있는 반

<표 4>  $C_{\beta}$ 에 대한 95% 신뢰영역의 포함비율

(X, Y)의 시그마수준	$n \backslash \rho$	$\rho$							
		-0.9	-0.6	-0.3	0	0.3	0.6	0.9	
(3.0,3.0)	20	0.929*	0.938	0.946	0.946	0.935	0.931*	0.922*	
	30	0.940	0.936	0.944	0.945	0.936	0.933	0.930*	
	60	0.953	0.947	0.954	0.946	0.938	0.947	0.942	
(3.0,6.0)	20	0.927*	0.921*	0.935	0.930*	0.932*	0.921*	0.921*	
	30	0.931*	0.938	0.948	0.950	0.931*	0.929*	0.926*	
	60	0.958	0.952	0.952	0.942	0.945	0.942	0.945	
(6.0,3.0)	20	0.925*	0.934	0.941	0.935	0.938	0.927*	0.917*	
	30	0.930*	0.939	0.943	0.944	0.941	0.946	0.948	
	60	0.950	0.945	0.954	0.959	0.954	0.950	0.949	
(6.0,6.0)	20	0.925*	0.934	0.941	0.935	0.938	0.927*	0.917*	
	30	0.930*	0.939	0.943	0.944	0.941	0.946	0.948	
	60	0.950	0.945	0.954	0.959	0.954	0.950	0.949	

면,  $n=60$ 인 경우에 대한 28개의 포함비율 모두가 범위 내에 있는 것으로 나타났다.

<표 5>는  $\mu_x < M_x$ 이고  $\mu_y < M_y$ 인 경우의  $C_{pk}$ 에 대한 95% 신뢰영역의 포함비율 결과이다. 먼저  $\rho$ 의 변화에 따른 결과를 살펴보면,  $\rho = -0.6, -0.3, 0, 0.3$ 인 경우에 대하여 일관성있게 기대값 0.95에 가까운 값을 갖는 것으로 나타났다. 실제로 이 경우에 대한 48개의 포함비율 모두가 범위 내에 있는 것으로 나타나 수행결과가 매우 좋음을 알 수 있다. 다음으로는  $\rho$ 의 값이  $-0.9, 0.6, 0.9$ 인 경우의 순으로 수행결과 좋은 것으로 나타났다. 특히  $\rho = \pm 0.9$ 인 경우에 대한 24개의 포함비율 중 4개가 범위를 벗어나고 있는 것으로 나타났다. 다음으로 표본크기  $n$ 에 따른 수행결과를 살펴보면, 역시  $n$ 이 증가할수록 기대값 0.95에 가까운 값을 갖는 것으로 나타났다. 실제로  $n=20$

인 경우에 대한 28개의 포함비율 중 3개가 범위를 벗어나고 있는 것으로 나타났으며, 반면  $n=60$ 인 경우에 대한 28개의 포함비율 모두가 이 범위 내에 있는 것으로 나타났다.

<표 6>은  $\mu_x = M_x$ 이고  $\mu_y < M_y$ 인 경우의  $C_{pk}$ 에 대한 95% 신뢰영역의 포함비율 결과이다. 먼저  $\rho$ 의 변화에 따른 결과를 살펴보면,  $\rho$ 의 값이  $-0.3$ 과  $0$ 인 경우에 대하여 일관성있게 기대값 0.95에 가까운 값을 갖는 것으로 나타났다. 실제로 이 경우에 대한 24개의 포함비율 중 단지 2개만이 범위를 벗어나고 있을 뿐 수행결과가 매우 좋음을 알 수 있다. 다음으로는  $\rho$ 의 값이  $-0.6, 0.3, 0.6, -0.9, 0.9$ 인 경우의 순으로 수행결과가 좋은 것으로 나타났다. 특히  $\rho$ 의 값이  $-0.9$ 와  $0.9$ 인 경우에 대한 24개의 포함비율 중 19개가 범위를 벗어나는 것으로 나타났다. 다음으로 표본크기  $n$

<표 5>  $C_{pk}$ 에 대한 95% 신뢰영역의 포함비율( $\mu_x < M_x$ 이고  $\mu_y < M_y$ 인 경우)

(X, Y)의 시그마수준	$n \backslash \rho$	$\rho$						
		-0.9	-0.6	-0.3	0	0.3	0.6	0.9
(4.5,4.5)	20	0.951	0.948	0.945	0.951	0.945	0.934	0.924*
	30	0.932*	0.933	0.945	0.940	0.935	0.932*	0.931*
	60	0.947	0.947	0.953	0.948	0.940	0.941	0.947
(4.5,6.0)	20	0.934	0.946	0.955	0.949	0.938	0.931*	0.937
	30	0.935	0.949	0.950	0.954	0.945	0.935	0.936
	60	0.954	0.955	0.958	0.954	0.951	0.949	0.948
(6.0,4.5)	20	0.933	0.943	0.943	0.946	0.949	0.936	0.921*
	30	0.947	0.936	0.943	0.950	0.950	0.949	0.945
	60	0.952	0.955	0.955	0.954	0.952	0.951	0.951
(6.0,6.0)	20	0.952	0.948	0.947	0.953	0.944	0.934	0.937
	30	0.953	0.950	0.959	0.952	0.940	0.932*	0.934
	60	0.943	0.942	0.940	0.942	0.948	0.936	0.938

<표 6>  $C_{pk}$ 에 대한 95% 신뢰영역의 포함비율 :  $\mu_x = M_x$ 이고  $\mu_y < M_y$ 인 경우

(X, Y)의 시그마수준	n \ ρ	ρ						
		-0.9	-0.6	-0.3	0	0.3	0.6	0.9
(4.5,4.5)	20	0.929*	0.935	0.942	0.936	0.937	0.933	0.922*
	30	0.918*	0.930*	0.943	0.945	0.943	0.940	0.918*
	60	0.925*	0.939	0.938	0.935	0.933	0.942	0.932*
(4.5,6.0)	20	0.907*	0.926*	0.932*	0.932*	0.924*	0.922*	0.907*
	30	0.915*	0.940	0.938	0.938	0.931*	0.924*	0.911*
	60	0.932*	0.944	0.947	0.940	0.937	0.936	0.919*
(6.0,4.5)	20	0.920*	0.939	0.939	0.941	0.939	0.929*	0.926*
	30	0.936	0.933	0.945	0.950	0.948	0.942	0.937
	60	0.936	0.948	0.957	0.957	0.949	0.942	0.934
(6.0,6.0)	20	0.934	0.949	0.953	0.959	0.951	0.942	0.928
	30	0.942	0.949	0.951	0.950	0.934	0.933	0.924*
	60	0.928*	0.941	0.940	0.945	0.950	0.938	0.933

에 따른 수행결과를 살펴보면,  $n$ 이 증가할수록 기대값 0.95에 가까운 값을 갖는 것으로 나타났다. 실제로  $n=20$ 인 경우에 대한 28개의 포함비율 중 14개가 범위를 벗어나고 있는 것으로 나타났으며, 반면  $n=60$ 인 경우에 대한 28개의 포함비율 중 단지 6개만이 범위를 벗어나고 있는 것으로 나타났다.

#### 4. 결론 및 향후 연구과제

본 논문에서는 이변량 공정분포에 대하여 현실적으로 불가능하리라고 생각되는 두 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 정확한 신뢰영역 대신에 근사적인 신뢰영역에 대하여 다음의 연구결과들을 얻었다.

첫째, 이변량 공정분포에 대하여 적당한 조건 하에서 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 의 추정량과 관련된 극한분포 결과를 유도·증명하고, 이를 기초로 두 벡터 공정능력지수  $C_p$ 와  $C_{pk}$ 에 대한 근사적인 신뢰영역을 설정·제시하였다.

둘째, 이변량 정규공정분포 하에서, 본 논문에서 제시한 이론적인 결과를 기초로 설정된  $C_p$ 와  $C_{pk}$ 의 95% 신뢰영역에 대한 모의실험을 수행한 결과 만족스러운 결과를 얻었다. 즉, 모의실험을 통하여 조사해 본 결과, 대체적으로 표본의 크기가 30 이상인 경우에는 명목 포함확률 95% 수준을 유지하고 있는 것으로 나타났다. 하지만 상관관계가 높아질수록 명목 포함확률 95%를 벗어나는 경우가 존재하는 것으로 드러났지만, 현장에서 적용하기

에는 충분히 의미가 있으리라고 사료된다.

향후 연구과제로 적당한 조건 하에서 벡터 공정능력지수  $C_p$ 와  $C_{pm}$  그리고  $C_{pk}$ 의 플러그인 추정량과 붓스트랩 추정량들에 관련된 극한분포 결과를 유도·증명하여 붓스트랩 방법의 일치성을 확립하고, 이를 기초로 벡터 공정능력지수  $C_p$ 와  $C_{pm}$  그리고  $C_{pk}$ 에 대한 몇 가지 근사 신뢰영역을 제안, 연구하는 문체(대표본 이론연구)를 고려할 수 있겠다. 그리고 여러 가지 이변량 공정 분포 하에서, 붓스트랩 방법을 포함하여 본 논문에서 제시한 이론적인 결과를 기초로 설정될 수 있는 벡터 공정능력지수  $C_p$ 와  $C_{pm}$  그리고  $C_{pk}$ 에 대한 몇 가지 형태의 신뢰영역에 대한 모의실험을 포괄적으로 수행하여 효율적으로 분석하는 보다 실제적인 연구가 가능하리라 사료된다(소표본 모의실험연구).

## 참고문헌

- [1] Alt, F.B. and N.D. Smith (1988). Multivariate Process Control, in P.R. Krishnaiah and C.R. Rao, Editors, *Handbook of Statistics*, Volume 7, North-Holland, Amsterdam, pp. 333-351.
- [2] Beran, R.J. (1984). Bootstrap Methods in Statistics. *Jahresbericht der Deutschen Mathematiker Vereinigung*, 86, pp. 14-30.
- [3] Chan, L.K., Xiong, Z. and Zhang, D. (1990). On the Asymptotic Distributions of Some Process Capability Indices. *Communications in Statistics : Theory and Methods*, 19(1), pp. 11-18.
- [4] Cho, J.J., Han, J.H. and Jo, S.H. (1997). Bootstrapping Unified Process Capability Index, *Journal of the Korean Statistical Society*, 26(4), pp. 543-554.
- [5] Cho, J.J., Kim, J.S. and Park, B.S. (1999). Better Nonparametric Bootstrap Confidence Interval for Process Capability Index  $C_{pk}$ , *Korean Journal of Applied Statistics*, 12(1), pp. 45-65.
- [6] Cho, J.J. and Park, B.S. (2002). Unpublished paper.
- [7] Diccio, T.J. and Tibshirani, R. (1987). Bootstrap Confidence Intervals and Bootstrap Approximations, *Journal of the American Statistical Association*, 82, pp. 163-170.
- [8] Efron, B. (1979). Bootstrap Methods : Another look at the jackknife, *Annals of Statistics*, 7, pp. 1-26.
- [9] Franklin, L.A. and Wasserman, G.S. (1992). Bootstrap Lower Confidence Interval Limits for Capability Indices, *Journal of Quality Technology*, 24, pp. 196-210.
- [10] Greenwich, M. and Jahr-Schaffrath, B.L. (1995). A Process Incapability Index, *International Journal of Quality & Reliability Management*, 12(4), pp. 58-71.
- [11] Hall, P. (1988). Theoretical Comparison of Bootstrap Confidence Intervals. *Annals of Statistics*, 16, pp. 927-953.
- [12] Han, J.H., Cho, J.J. and Lim, C.S. (2000). Bootstrap Confidence Limits for Wright's  $C_s$ , *Communications in Statistics : Theory and Methods*, 29(3), pp. 485-505.

- 
- [13] Hubele,N.F., Shahriari,H. and Cheng, C.S.(1991). A bivariate process capability vector, in Statistical Process Control in Manufacturing (J.B. Keats and D.C. Montgomery, eds.) M. Dekker : New York, pp. 299-310.
- [14] Kocherlakota,S. and Kocherlakota,K.(1991). Process capability indices : Bivariate Normal distribution, *Communication in Statistics : Theory and Methods*, 20, pp. 2529-2547.
- [15] Kotz,S. and Johnson,N.L.(1993). Process Capability Indices, 1st ed., Chapman & Hall.
- [16] Pearn,W.L., Kotz,S. and Johnson,N.L (1992). Distributional and Inferential Properties of Process Capabiblity Indices, *Journal of Quality Technology*, 24, pp. 216-231.
- [17] Wierda,S.J.(1992). A multivariate process capability index, *Proceedings 9th International Conference: Israel Society of Quality Assurance*, Jerusalem, Israel, pp. 517-522.
-