

6 Sigma 현장적용 적용 시 초기 데이터 분석에 대한 고찰

이상복 · 최은향

서경대학교 경영대학원 경영학과, 6시그마 전공

Study on analysis of initial Data on 6 Sigma application in real fields

Key Words : 6시그마, Gage R&R, Histogram, Box-plot, PDF, Box-Cox 변화

Abstract

본 연구에서는 현장에서 6 시그마를 활용 시 통계분석의 기초가 되는 초기 데이터 분석에 대한 고찰이다. 통계의 가장 기본이 되는 데이터가 잘못되었으면 나머지 모두 문제가 된다. 이에 데이터 초기에 발생할 수 있는 여러 오류의 가능성을 살펴보고 각각에 대해 해결책을 제시하였다. 여기서 활용하는 방법들은 계측기 선정, Gage R&R, Histogram, Box-plot, PDF, Box-Cox 변화 등이다.

1. 서 론

본 연구에서는 현장에서 6시그마 적용 시 초기 데이터 오류에 대한 분석을 하였다. 부분적으로는 이와 관련된 많은 논의는 되어왔고 대학교재 등에서 언급하고 있지만 체계적으로 정리한 것이 없다. 현장에 6시그마가 활발히 사용되고 있음에도 불구하고 데이터가 정확한지는 의문이 많다. 대부분 6 시그마 교재에서 이 문제를 거론하지 않고 있으며, 6 시그마 과제 해결과정(로드맵)에서도 이 문제에 대한 검증 단계가 없다. 하지만 통계에서 초기의 데이터가 정확하지 않으면 나머지 단계는 오류이다. 본 연구에서는 이점을 중시하여 아래와 같이 해결책을 제시하였다.

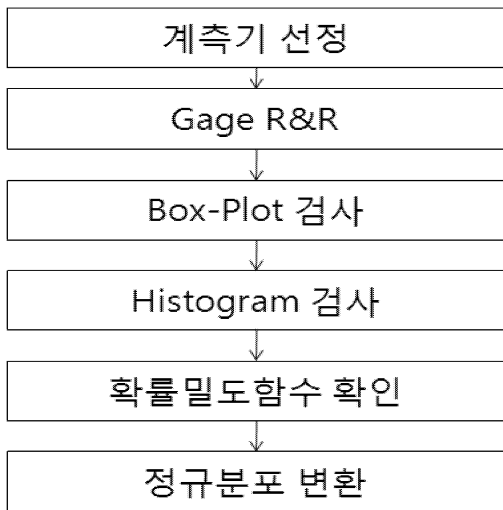
본 연구에서 다루는 내용은 크게 6가지이다. 계측기 선택, Gage R&R, 이상치 여부 검토, 테

이터 형태 검사, 확률밀도함수 확인, 정규분포로 변화 등이다.

이러한 과정을 거쳐 데이터를 분석해야 올바른 결론을 도출할 수 있다. 본 논문에서는 검증 순서를 설명하고 이에 맞는 예를 제시하였다. 본 논문에서 제시하는 내용들은 기존 문헌에서 부분적으로 언급하였으나 이들 내용들을 활용하기 쉽게 정리한 데 의의가 있다.

2. 초기 데이터 검증 절차 제공

본 연구에서는 데이터를 얻는 초기 단계인 게이지 선택 문제부터 데이터의 정규분포 변환까지 6 단계를 제시하였다<그림 1>.

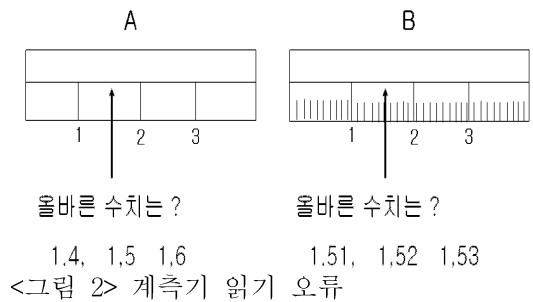


<그림 1> 데이터 검정 순서

1) 계측기 선정

데이터는 계측기를 통하여 얻는다. 이 단계에서 계측기 선정이 잘못되었다면 나머지 데이터는 오류일 것이다. 계측기 선정이 올바른지 검토한다. 이 단계에서 오류는 관리하는 수준보다 1/10 단계 정밀한 계측기를 사용해야 한다는 것이다.

계측기 읽는 원칙은 계측기에 나타난 숫자만 읽어야 한다. <그림 2>의 A자를 통해서 읽는 수치는 2이다. A자를 계측기로 사용했으면 1, 2 등으로 읽어야 하므로, 1단위 수치는 정확한 값이 아닌 예측한 값으로, 1단위 관리는 할 수 없다. B 자는 1.5가 정확한 값이다. B 자의 눈금은 1/10 cm로 나타났기 때문이다. 처음 데이터를 읽는 것을 검토하고 관리수치보다 1/10 정밀한 계측기를 선택한다.



<그림 2> 계측기 읽기 오류

2) Gage R&R

6시그마 문제풀이 순서에 Gage R&R을 다루고 있어 이 문제는 대부분 해결되고 있다. 계측기 읽는 사람의 변동성 문제를 검증해야 한다. 이 부분은 계측기 자체의 문제와 활용 문제로 나누어 볼 수 있다.

계측기 자체의 문제는 계측기 사용시에 계측기 검교정 확인서를 확인한다. 유효날자를 확인하여 검교정 받은 인증서가 있으면 된다. 검교정을 받은 계측기를 사용하는 사람의 반복성과 재현성 검정은 미니탭에서 제공한 프로그램을 이용하면 쉽다. 계량치인 경우 반드시 샘플 시료를 2개 이상, 반복실험, 실험자 2명 이상으로 실험하여 “미니탭>통계분석>품질도구>Gage R&R”을 이용하면 <그림 3>과 같은 검사결과가 나타난다.

게이지 R&R			
출처	표준		
	편차(SD)	연구변동(6 * SD)	%연구변동(%SV)
총게이지 R&R	0.066615	0.39969	32.66
반복성	0.035940	0.21564	17.62
재현성	0.056088	0.33653	27.50
측정자	0.030200	0.18120	14.81
측정자*Part	0.047263	0.28358	23.17
부품-대-부품	0.192781	1.15668	94.52
총변동	0.203965	1.22379	100.00

구별되는 범주의 수 = 4

<그림 3> Gage R&R 결과

위와 같은 결과에서 %연구변동이 30% 이상이면 본 측정 데이터는 믿을 수 없다. 다시 측정해야한다. 특히 위 그림에서 재현성이 문제있는 것으로 나타났다. <그림 3>의 아래측에 구별되는 범주의 수가 5 이상되어야 본 측정 데이터를 신뢰할 수 있다. 이 과정을 거치면 Gage R&R은 문제 없다고 본다.

3) Box-Plot 검사

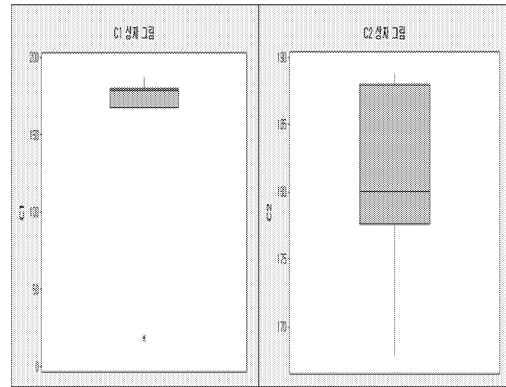
계측기를 통하여 얻은 데이터를 그대로 믿어서는 안된다. 혹 읽는 과정이 올바르더라도 이기 작업 등에서 오류가 발생할 수 있다. 이런 문제는 상자 그림을 그리면 쉽게 해결할 수 있다.

올바로 측정된 데이터를 컴퓨터에 입력하는 과정에서 소수점을 잘못 기재할 수 있다. 혹은 관련 없는 샘플이 측정되는 경우도 있다. 예로

아래 데이터는 남성 30대 키를 측정한 것이다.

167.8 180.1 188.0 178.6 180.1 177.6 18.89

마지막 데이터는 소수점 한자리를 잘못 기재한 경우이다. 이를 제거하지 않으면 평균이 155.9이고 이를 18.89 -> 188.9로 수정한 평균은 180.16이다. 이는 엄청난 차이이다. 데이터가 많으면 이러한 오류는 눈으로 봐서 알 수 없다. “미니탭>그래프>상자그림”을 그리면 <그림 4>와 같이 확실하게 구별된다. <그림 4>의 왼쪽 그림에서 이상치는 “*”로 표시되고 이런 데이터는 이유를 찾아서 해결하여 다시 그리면 <그림 4>의 오른쪽과 같이 그려진다. 오른쪽 그림과 같이 그려지면 문제가 없는 것이다.

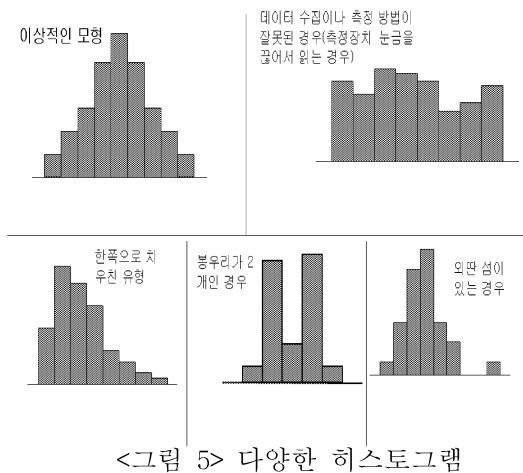


<그림 4> 상자 그림 예

4) Histogram 검사

상자 그림으로 데이터 이기 작업에서 문제없으면 데이터를 가져온 집단이 동일한지 검증해야 한다. 만약 해결하고자하는 문제 범위 밖의 데이터를 가져왔는지, 관리상태가 아닌 데이터를 가져왔는지는 히스토그램을 그리면 쉽게 알 수 있다. 또한 데이터의 전체적인 상황을 알아야 통계량 사용을 정확히 할 수 있다.

6시그마 문제 해결과정 시 데이터 수집은 같은 조건하에 있는 샘플을 구해야 한다. 같은 조건 하에서 관리가 되고 있다면 데이터는 자연스런 정규분포를 이룬다고 본다. 구해진 데이터를 히스토그램을 그리면 쉽게 이를 알 수 있다. <그림 5>와 같이 다양한 히스토그램이 그려지면 데이터를 얻은 조건을 검토해야 한다. 관리가 되지 않는 환경에서 얻은 데이터는 다음에 샘플링하면 전혀 다른 데이터를 얻을 수 있다. 혹은 두 집단이 섞여 있는 경우도 있다.

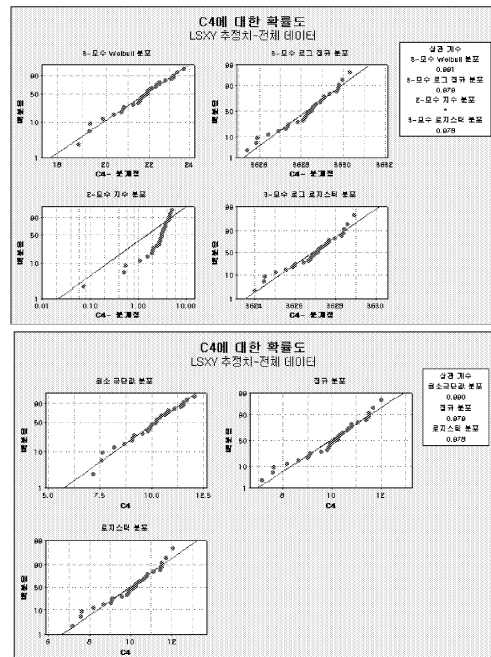
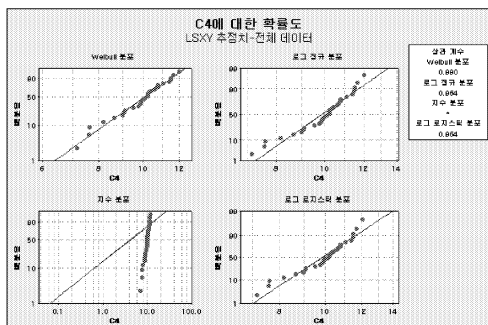


<그림 5> 다양한 히스토그램

5) 확률밀도함수 확인

데이터 수집을 마쳤으면 평균이나 분산 등 통계량을 활용하기 전에 적당한 분포를 알아야 된다. 예를 들면 지수분포함수를 정규분포처럼 사용하면 결론을 믿을 수 없다.

“미니탭>통계분석>신뢰성분석>분포분석(우측 관측중단)”을 활용하면 주어진 데이터에 대해 <그림 6>과 같이 다양한 분포 10개와 맞는 정도를 계산한다. 그래프에서 대각선과 일치하는 것이 해당 확률분포함수이다.



<그림 6> 다양한 분포함수들

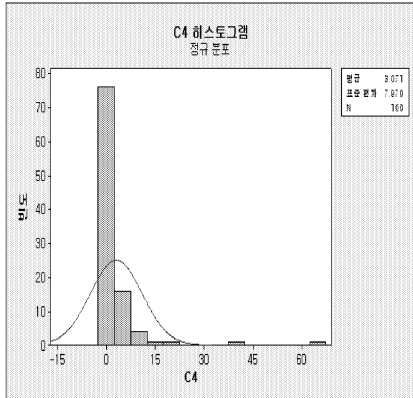
데이터가 정규분포가 아닌데 정규분포로 간주하여 계산하는 것은 불합리하다. 데이터를 얻었으며 정규분포인지 검토하고, 아니면 다양한 분포함수와 비교하여 정확한 분포를 찾는 것이 중요하다.

6) 정규분포 변화

통계적 가설검정의 기본은 정규분포이다. 정규분포가 아닌 것을 정규분포로 가정하여 통계값을 사용하면 현실과 맞지 않는다. 단계 5에서도 적당한 확률분포를 찾을 수 없는 경우도 많다. 적당한 분포함수가 없는 경우는 히스토그램을 그려보면 심하게 치우친 데이터는 Box-Cox 변환을 하여 정규분포에 가깝게 만들어 사용해야 한다. Box-Cox 변환으로 모든 비정규분포가 정규분포로 변환할 수는 없다. Box-Cox 변환은 정규분포에 가깝게 하려는 것이다.

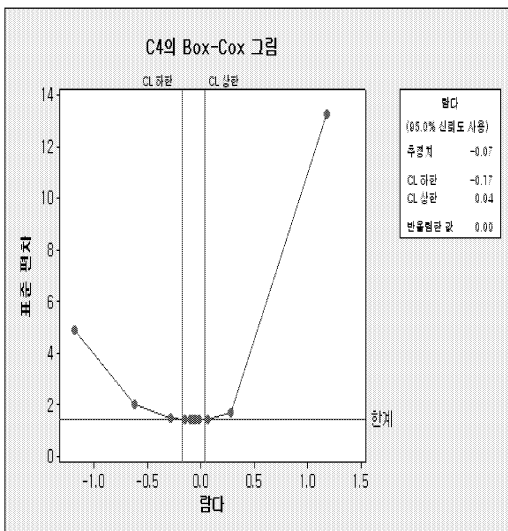
<그림 7>과 같이 초기 데이터를 얻었다면 이

데이터는 알려진 분포함수는 없는 경우이다.



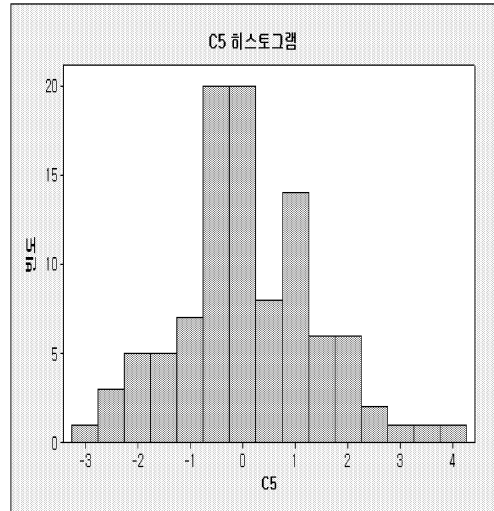
<그림 7> 초기 데이터의 히스토그램

“미니맵>통계분석>관리도>Box-Cox 변환”을 선택하면 <그림 8>과 같이 Box-Cox 그림이 나타나 자동으로 변환을 해준다.



<그림8> Box-Cox 결과값

자동 변환한 데이터를 히스토그램으로 그리면 <그림 9>와 같이 정규분포에 가깝게 그려졌다.



<그림9> Box-Cox변환된 데이터의 히스토그램

3. 적용 사례

본 절에서는 2절에서 제시한 방법들의 적용사례를 제시한다. 사료에 사용되는 분말을 만드는 A 기업은 생산되는 분말을 일정하게 관리하고자 한다. 분말 규격은 100 ± 3 g이다.

1) 계측기 선정

데이터는 전자저울을 사용한다. 전자저울은 소숫점 2자리까지 읽을 수 있다. 계측기는 최근에 검교정을 받아 인증서가 첨부되었다. 문제가 없다.

2) Gage R&R

검사원 2명이 샘플 3개로 2회 반복 실험하였다.

샘플	작업자	측정값
1	A	100.1
1	A	100.2
1	B	100.1
1	B	100.1
2	A	110.1
2	A	110.3

2	B	110.2
2	B	110.1
3	A	120.1
3	A	120.2
3	B	120.2
3	B	120.2

측정값에 대한 Gage R&R

출처	표준 편차 (SD)	연구 변동 (6*SD)	%연구 변동 (%SV)
총 Gage R&R	0.0722	0.4330	0.72
반복성	0.0722	0.4330	0.72
재현성	0.0000	0.0000	0.00
작업자	0.0000	0.0000	0.00
부품-대-부품	10.0249	60.1497	100.00
총 변동	10.0252	60.1512	100.00

고유한 범주의 수 = 195

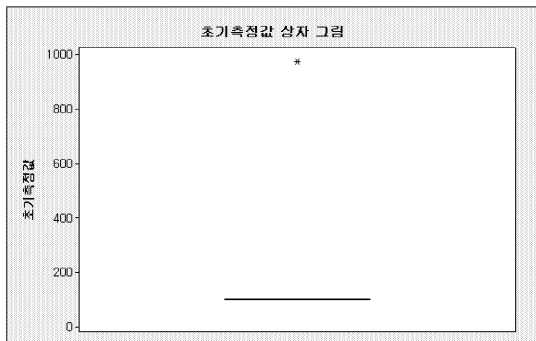
위의 표와 같이 Gage R&R은 충분하다.

3) Box-Plot 검사

초기에 측정된 데이터는 다음과 같다.

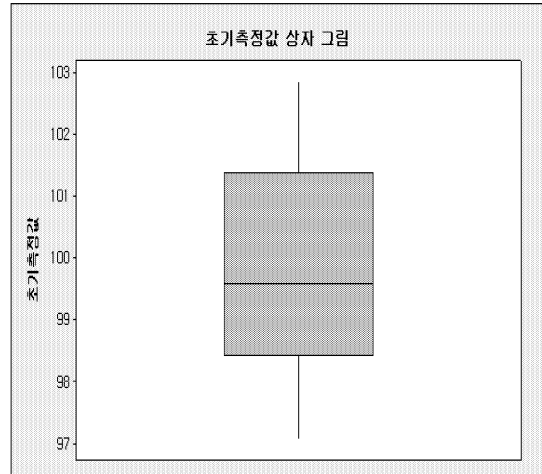
102.4	101.2	99.4	98.2	102.5	98.9
101.9	99.0	102.8	97.5	99.1	102.3
98.5	101.0	97.5	97.7	99.5	102.8
98.1	99.6	102.0	100.5	973.0	97.1
100.8	100.6	101.2	99.9	99.0	99.3

Box plot을 그렸더니 다음과 같다.



<그림10> 초기 데이터 상자그림

데이터 하나가 심하게 큰 수치가 있다. 973.0이다. 이는 소수점을 잘못 썼다. 이를 97.3으로 수정하여 다시 그렸다.

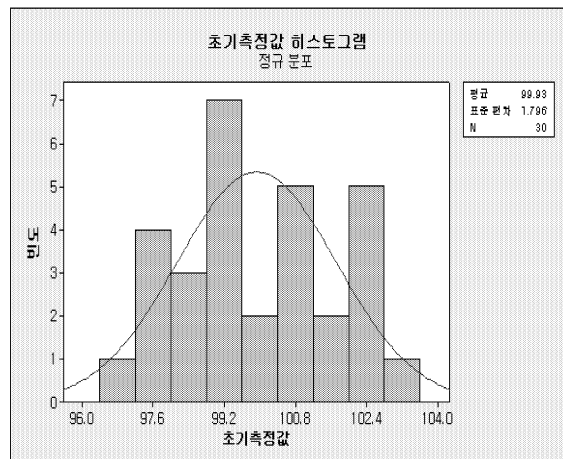


<그림11> 수정된 상자그림

이상이 없다.

4) Histogram 검사

이를 히스토그램을 그렸다.

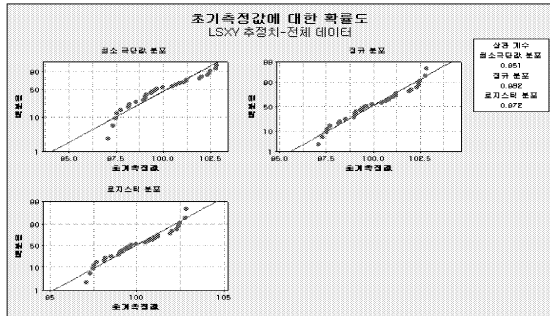


<그림12> 초기 히스토그램

이상 데이터가 없고, 관리한계선을 크게 벗어나지 않는 점 등을 볼 때, 문제없다.

5) 확률밀도함수 확인

데이터 상태로 확률분포를 검정하였다.



<그림 13> 확률밀도함수 검토

정규분포에 가깝다. 정규성 검정한 결과 p값이 0.259로 정규분포로 볼 수 있다.

6) 정규분포 변화

정규분포로 확인되어 별도로 정규분포로 변환할 필요는 없다.

위의 데이터는 정규분포로 인정하고 나머지 개선 작업을 진행하면 좋다.

4. 결론

본 연구에서는 현장에서 6시그마 개선 작업에서 초기 데이터 활용 시에 발생 가능한 오류에 대해 분석하였다. 이에 관련하여 대응방법을 제시하였다.

본 연구는 새로운 이론이나 방법을 제시하기보다 알려진 방법들을 현장에서 활용할 수 있도록 단계별로 정리하였다.

본 내용이 현장에서 쉽게 적용하는데 의의가 있다. 현장에 많은 도움이 되기를 기원한다.

참고문헌

- (1) 이상복(2001) "예제중심의 실험계획법(분산분석, 요인배치, 반응표면분석, 혼합물분석)" 이레테크
- (2) 이상복(2004), 품질이야기, 이레테크
- (3) 이상복,염경철(2006), 샘플링검사실무(ISO 2859) 새로운규격에 맞춘, 이레테크
- (4) 이상복(2007), Minitab을 이용한 관리도 설계 및 활용, 이레테크