

계측기 능력분석과 실험계획법

백재욱

한국방송대학교 응용통계학과

조진남

동덕여자대학교 계산통계학과

Gauge Capability Analysis and Designed Experiments

Jaiwook Baik

Korea National Open University

Jinnam Jo

DongDuck Women's University

Abstract

In today's organization, measurement plays a crucial role in helping improve process or quality. In this paper, we review the measurement error study, classical gauge repeatability and reproducibility study, and designed experiment suited for the determination of the measurement capability.

Measurement error study is very simple to use but is rather crude. Hence, it should be used as a preliminary study to determine whether further study is necessary.

Classical gauge repeatability and reproducibility (GR&R) study is the most common technique for evaluation of gauge capability. It calculates a percentage that relates the repeatability, reproducibility, and overall R&R to the specification range for the parameter measured. Hence, the individual repeatability and reproducibility statistics will indicate the area on which to concentrate. However, GR&R study only gives a point estimate of each component, which leaves a room for improvement. It is always good to integrate the use of control charts to ascertain the statistical stability of the measurement process.

The tools of statistical experimental design are available for a comprehensive design and analysis of the measurement process. Hence, in this paper, we present gauge capability analysis as an experimental design problem and compare it with the classical GR&R study.

1. 서론

계측은 기술을 선도한다. 그 이유는 무엇인가를 제대로 측정하기 위해서는 그 대상을 완전히 파악하지 않으면 안되기 때문이다. 계측은 종종 초정밀, 초미량 또는 아주 복잡할 수 있으며, 계측대상 또한 여러 가지일 수 있다. 통계적 공정관리에서는 계측대상이 곧 제품의 품질특성치이다.

어느 공정에서든 품질특성의 과거 및 현재 수준이 제대로 파악되어야 품질의 향상이 이루어진다. 즉 제조공정에서는 공정이 구체적으로 어떻게 이루어지는지, 그리고 주어진 공정을 계속적으로 어떻게 향상시킬 것인지 파악될 때에 비로서 좋은 품질의 제품이 만들어진다. 이로 인해, 주어진 공정으로부터 나오는 정보에 대한 적절한 통계처리는 더욱 필요해졌으며, 이는 곧, 공정의 향상은 물론 기업의 생존에 필수적인 요소가 되었다.

그런데 생산공정의 운영과 관리에 사용되는 자료의 질은 주어진 공정의 상황을 파악하는데 이용되는 측정도구 즉 계측기에 의해 크게 좌우된다. 데밍(1982)은 제품의 품질향상을 위해서는 공정의 변동(process variability)에 대한 이해가 필요하다고 하는데, 이 변동은 크게 측정과정 중에서 생기는 변동(measurement variability)과 제품 자체의 변동(product variability)의 합으로 구성된다고 한다. 특히 측정과정 중에 생기는 변동이 전체 공정의 변동에 어느 정도의 비율을 차지하는지 아는 것은, 어떤 특정한 공정향상 정책을 수립하여 실시했을 때 그것이 제대로 효과를 발휘했는지 파악하는데 필요하다. 더구나 근래와 같이 첨단기기가 동원되는 측정현장에서는 효율적인 계측기 능력분석(gauge capability analysis)이 전체 공정의 이해에 더욱 중요한 위치를 차지한다.

측정결과에 영향을 미치는 여러 가지 요인을 보다 잘 이해하기 위해서는 측정자료가 왜 항상 같게 나오지 않는가에 대한 실증적인 연구가 이루어져야 한다. 따라서 계측기 능력분석의 목적은 계측과정 도중 측정결과에 영향을 미치는 요인을 사전에 예측하여, 이에 대해 보다 잘 관리하여 측정과정의 향상은 물론 공정 자체의 향상을 꾀하는데 있다. 이때 여러 가지 실험이 필요한데, 그 이유는 여러 실험을 통하여 통계적으로 적절한 모형을 파악하며, 이를 이용하여 측정과정에 대한 적절한 해석을 내릴 수 있기 때문이다.

계측자료에 대한 가장 간단한 분석도구는 Floyd와 Laurent(1995)가 설명하는 measurement error study이다. 그런데 이 방법은 간단하여 유용하지만 계측결과에 문제가 있을 때 계측과정중 어느 부분에 문제가 있는지는 전혀 알 수가 없다. 또한 계측기의 능력을 단지 점추정치 하나로만 요약하고 있으므로 그에 대한 신빙성 또한 문제가 있다.

계측자료에 대한 다른 분석도구로는 Montgomery(1991)와 Montgomery와 Runger(1993a)가 설명하는 Gauge repeatability and reproducibility (GR&R) study가 있다. 이 방법은 앞의 방법보다 좀더 포괄적인 방법으로서, 계측자료의 변동을 계측기

자체에 의한 변동(repeatability를 말함)과 작업자 때문에 생기는 변동(reproducibility를 말함)으로 나누어, 계측결과가 만족스럽지 못할 경우 어느 요인이 문제인지 파악할 수 있는 방법이다. 그러나 이 방법도 각각 요인에 의한 변동 등을 하나의 점추정치로만 나타내므로 통계적 해석이 부족하다는 단점이 있다.

마지막으로 계측자료의 분석에 많이 쓰이는 통계 분석도구로 Deutler(1991), Montgomery와 Runger(1993a), Montgomery와 Runger(1993b), Tsai(1988)가 설명하는 실험계획법이 있다. 특히 실험계획법에서 이용되는 반복, 블록화, 랜덤화 등의 여러 가지 기본 원리들이 계측기 능력분석과 관련된 연구에서 아주 많이 이용될 수 있다. 따라서 실험계획법의 기법을 보다 잘 이해하고, 계측기 실험에 이를 제대로 활용할 수 있다면 다른 어느 방법을 사용하는 것보다 훨씬 더 포괄적인 계측기 능력분석을 실시할 수 있을 것이다.

이 논문에서는 계측기 능력분석을 기존의 계측기 능력분석 방법인 measurement error study와 GR&R study의 측면에서 먼저 살펴보고, 다음으로 실험계획법의 측면에서도 또한 살펴보고자 한다. 실험계획법에서도 여기까지 모형이 있지만 여기에서는 이 원배치모형에 대하여만 살펴보고자 한다. 아울러 이러한 계측자료의 분석에서는 관리도 등의 그림이 유용하게 사용되므로, 이들 그림으로는 어떤 것들이 있는지 살펴본다.

2. Measurement error study

<표 1>은 Floyd와 Laurent(1995)에 나오는 것으로 measurement error study에서는 이와 비슷한 자료들이 많이 인용된다. 즉 한 명의 작업자가 하나의 계측기를 이용하여 여러 번 품질특성치를 측정한다. 이때 계측결과가 만족할 만한지 측정하는 방법으로 measurement error ratio(MER)가 있으며, <표 1>의 자료의 경우 MER은 0.2835로 나온다. Measurement error study에서는 이 값이 0.1보다 크므로 계측과정에 문제가 있다고 판단을 내린다.

그러나 앞에서와 같은 일률적인 해석에는 문제가 있다고 본다. 왜냐하면 똑같은 상황에서 똑같은 수만큼 측정을 한다고 해도 그 결과는 항상 같게 나올 수는 없기 때문이다. 그러므로 MER에 대한 구간추정을 해야될 필요가 있다.

신뢰구간을 구하는 방법 또한 여러 가지가 있을 수 있으나 계측자료가 정규분포를 따른다고 가정하고(사실 표 1의 자료는 따르지 않지만) 카이제곱 분포를 이용하여 대략적인 신뢰구간을 구할 수 있다. 그 결과 MER에 대한 95% 신뢰구간은 $0.221 < \text{MER} < 0.394$ 와 같이 나오며, 따라서 MER은 최소한 0.2보다는 크므로 계측과정에 정말로 문제가 있다고 말할 수 있다.

주어진 자료가 어떤 특정한 분포를 따르지 않는 경우 MER에 대한 신뢰구간을 구하는 방법으로 resampling 방법을 생각할 수 있다. 그 방법의 기본 원리는 표본으로부터 엮을 수 있는 불확실성을 MER에 대한 신뢰구간의 추정에 그대로 반영하자는 것이다

< 표 1 > Measurement error study의 예

The measurement error must be limited so that the total variation reflects as much of the actual product variation as possible.

Procedure

Check for a set master/part master * OK (100)

Check for the last calibration OK

1. Measure 1 part 25 times.

<u>100</u>	<u>101</u>	<u>100</u>	<u>100</u>	<u>100</u>
<u>100</u>	<u>100</u>	<u>100</u>	<u>100</u>	<u>100</u>
<u>100</u>	<u>100</u>	<u>99</u>	<u>100</u>	<u>100</u>
<u>101</u>	<u>101</u>	<u>100</u>	<u>100</u>	<u>101</u>
<u>100</u>	<u>101</u>	<u>100</u>	<u>100</u>	<u>100</u>

2. Calculate the average of the 25 measurements.

$$\bar{X} = \underline{100.16}$$

3. Calculate the standard deviation of the 25 measurements.

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}} = 0.47258$$

4. s is the measurement error of the measuring process.

5. Calculate the Measurement Error Ratio.

$$\text{Measurement error ratio} = \frac{6 \times sd}{\text{Total tolerance}} = \frac{2.835}{10}$$

6. Check one the following :

Measurement Error Ratio > 0.10

The measurement process is not adequate.

The method/equipment should be evaluated and replaced before proceeding any further.

Measurement Error Ratio ≤ 0.10

The measurement process is adequate.

7. Compare the \bar{X} to the known measurement check for accuracy

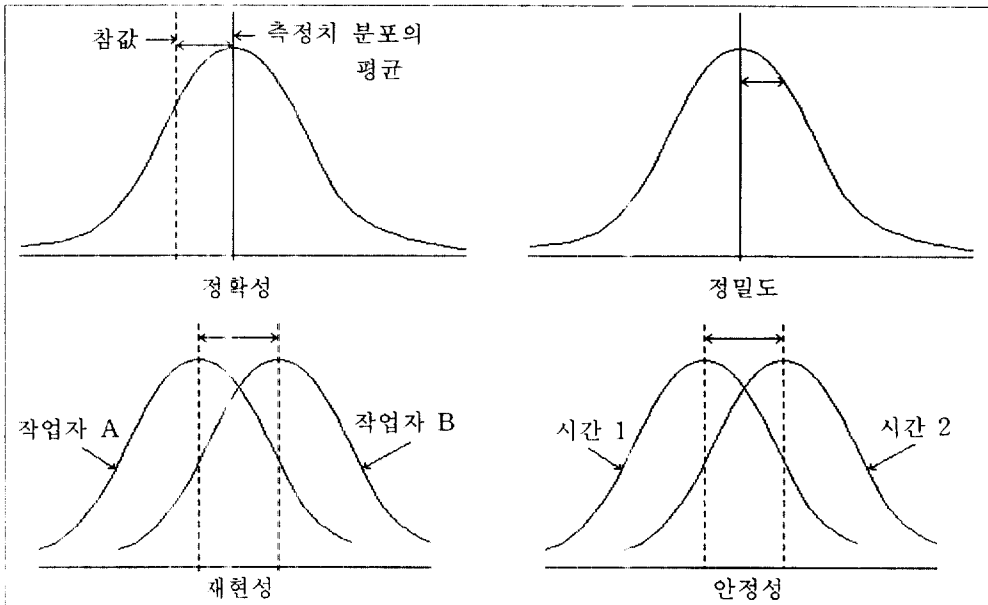
(see * above) 100.16 - 100 = 0.16 deviation

8. Comments _____

앞의 <표 1>의 자료의 경우 1,000 번 resampling한 결과 MER에 대한 95% 신뢰구간은 $0.166 < MER < 0.360$ 으로 나온다. 따라서 MER은 최소한 0.1보다는 크므로 계측과정에 문제가 있다고 말할 수 있다. 그렇다면 측정과정중 어느 요인이 문제가 되는 것인가? 그렇지만 이에 대한 해답은 measurement error study를 통해서서는 구할 수 없다.

3. GR&R 분석

다음은 GR&R 분석에서 많이 나오는 Montgomery(1991)와 Montgomery와 Runger(1993a)에 있는 예이다. 지금 많은 작업자들중 랜덤하게 뽑힌 3명의 작업자가 어느 공정으로부터 랜덤하게 뽑힌 20개의 제품에 대해 그 특성치를 각각 2 번씩 반복 측정하였다. 이때 각 작업자는 20개의 제품을 랜덤한 순서로 측정하였으며 그 결과는 <표 2>와 같다. <표 2>의 자료에서 보듯이 계측기로부터 나온 측정치들은 대부분 상이한 값들을 갖는다. 다음은 상이한 계측값들과 관련하여 계측기가 가져야 할 바람직한 성질들이며, 이를 그림으로 나타내면 <그림 1>과 같다.



< 그림 1 > 계측기의 바람직한 성질들

< 표 2 > GR&R 분석

PART NO.	Susan				Bill				Tom			
	MEASUREMENTS				MEASUREMENTS				MEASUREMENTS			
	1	2	\bar{x}	R	1	2	\bar{x}	R	1	2	\bar{x}	R
1	21	20	20.5	1	20	20	20.0	0	19	21	20.0	2
2	24	23	23.5	1	24	24	24.0	0	23	24	23.5	1
3	20	21	20.5	1	19	21	20.0	2	20	22	21.0	2
4	27	27	27.0	0	28	26	27.0	2	27	28	27.5	1
5	19	18	18.5	1	19	18	18.5	1	18	21	19.5	3
6	23	21	22.0	2	24	21	22.5	3	23	22	22.5	1
7	22	21	21.5	1	22	24	23.0	2	22	20	21.0	2
8	19	17	18.0	2	18	20	19.0	2	19	18	18.5	1
9	24	23	23.5	1	25	23	24.0	2	24	24	24.0	0
10	25	23	24.0	2	26	25	25.5	1	24	25	24.5	1
11	21	20	20.5	1	20	20	20.0	0	21	20	20.5	1
12	18	19	18.5	1	17	19	18.0	2	18	19	18.5	1
13	23	23	24.0	2	25	25	25.0	0	25	25	25.0	0
14	24	24	24.0	0	23	25	24.0	2	24	25	24.5	1
15	29	30	29.5	1	30	28	29.0	2	31	30	30.5	1
16	26	26	26.0	0	25	26	25.5	1	25	27	26.0	2
17	20	20	20.0	0	19	20	19.5	1	20	20	20.0	0
18	19	21	20.0	2	19	19	19.0	0	21	23	22.0	2
19	25	23	25.5	1	25	24	24.5	1	25	25	25.0	0
20	19	19	19.0	0	18	17	17.5	1	19	17	18.0	2

$$\bar{x}_1 = 22.30 \quad \bar{R}_1 = 1.00 \quad \bar{x}_2 = 22.28 \quad \bar{R}_2 = 1.25$$

$$\bar{x}_3 = 22.60 \quad \bar{R}_3 = 1.20$$

$$\sigma_{total}^2 = \sigma_{product}^2 + \sigma_{gauge}^2$$

Repeatability

Reproducibility

$$\bar{R} = (\bar{R}_1 + \bar{R}_2 + \bar{R}_3)/3$$

$$\bar{x}_{max} = \max(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 22.60$$

$$= \frac{1}{3} (1.00 + 1.25 + 1.20)$$

$$\bar{x}_{min} = \min(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 22.28$$

$$= 1.15$$

$$R_{\bar{x}_i} = \bar{x}_{max} - \bar{x}_{min} = 22.60 - 22.28 = 0.32$$

$$\hat{\sigma}_{repeatability} = \frac{\bar{R}}{d_2} = \frac{1.15}{1.128} = 1.02$$

$$\hat{\sigma}_{reproducibility} = \frac{R_{\bar{x}_i}}{d_2} = \frac{0.32}{1.698} = 0.19$$

$$\hat{\sigma}_{gauge}^2 = \hat{\sigma}_{repeatability}^2 + \hat{\sigma}_{reproducibility}^2$$

$$= (1.02)^2 + (0.19)^2 = 1.08$$

정확도(accuracy) : 측정치 분포의 평균이 참값과 차이가 없을 때 정확도가 높다고 한다.

정밀도(precision 또는 repeatability) : 동일 제품의 품질특성치를 무한히 많이 측정하여도 그 관측값들은 항상 똑같지 않고 일정한 산포를 갖게 되는데, 이때 산포가 작으면 작을수록 정밀도는 높다고 한다.

재현성(reproducibility) : 둘 또는 세명의 작업자가 동일 계측기로 동일 제품에 대해 품질특성치를 측정해도 그들의 평균간에는 보통 차이가 있게 되는데, 이때 그들의 차이가 크면 클수록 계측기의 재현성은 떨어진다고 한다.

안정성(stability) : 시간의 변화에 따라 계측결과가 달라질 수 있는데, 이들 평균치들은 서로 차이가 나지 않을수록 좋으며, 이때 안정성이 높다고 한다.

R&R : 계측기에 의한 변동(repeatability)과 계측자에 의한 변동(reproducibility)의 합으로서, 이 값이 공정의 허용한계(total tolerance = 규격상한치 - 규격하한치)에 비해 아주 작으면(예를 들어 0.1보다 작으면) 계측기는 정밀도와 재현성 면에서 관리되어 있다고 본다.

GR&R 분석에서는 계측기의 정밀도 및 재현성을 <표 2>의 아래 부분에 나와 있는 바와 같이 구한다. 그 결과 계측기의 정밀도 및 재현성은 각각

$$\hat{\sigma}_{repeatability} = \frac{\overline{R}}{d_2} = \frac{1.15}{1.128} = 1.02 \text{와}$$

$$\hat{\sigma}_{reproducibility} = \frac{R \overline{\bar{x}_i}}{d_2} = \frac{0.32}{1.693} = 0.19$$

로 나오며, 계측기의 분산은

$$\begin{aligned} \hat{\sigma}_{gauge}^2 &= \hat{\sigma}_{repeatability}^2 + \hat{\sigma}_{reproducibility}^2 \\ &= 1.08 \end{aligned}$$

로 나온다.

그러면 total tolerance가 주어진 경우 $6 \times \hat{\sigma}_{gauge} / \text{total tolerance}$ 를 구하여, 이 값이 0.1보다 크다면 측정과정에 문제가 있다고 보고 측정과정중 어느 요인이 문제인지 점검한다. 이때 $\hat{\sigma}_{repeatability}$ 와 $\hat{\sigma}_{reproducibility}$ 의 크기를 비교하여 큰 쪽에 문제가 있다고 본다.

지금까지 살펴본 바와 같이 GR&R 분석에서는 어느 요인에 의한 자료의 변동을 점 추정치 하나로만 나타낸다. 그러나 아무리 똑같은 계측실험이라 하더라도 항상 같은 결과가 나오는 것은 아니다. 따라서 각 변동에 대한 구간추정을 구하는 것도 의미가 있을 것이다. 주어진 자료가 정규분포를 따르는 경우 $\hat{\sigma}_{repeatability} = \overline{R}/d_2$ 의 기대값과

분산은 각각 $\sigma_{repeatability}$ 및 $(d_3^2 \sigma_{repeatability}^2) / (op d_2^2)$ 로 나온다. 따라서 앞의 <표 2>의 자료를 이용하여 정밀도 $\sigma_{repeatability}$ 에 대한 95% 신뢰구간을 중심극한정리에 의하여 대략적으로 구해보면 $0.8248 < \sigma_{repeatability} < 1.2152$ 와 같이 나온다. 이와 같이 다른 요인에 의한 자료의 변동에 대해서도 95% 신뢰구간을 구할 수 있다면, 주어진 측정과정에 대해 보다 정확한 해석을 내릴 수 있을 것이다.

정밀도나 재현성에 문제가 있는지 파악하는 또다른 방법으로 관리도 등의 그림을 빼놓을 수 없다. 다음 <그림 2>는 <표 2>의 자료에 대한 \bar{X} 와 R 관리도이다. 여기에서 관리한계선은 세 작업자가 측정시 같은 수준의 오차를 낸다는 가정하에 구한 것이다. 구체적으로 \bar{X} 관리도의 경우 관리상한선, 중심선, 관리하한선은 각각

$$UCL = \bar{\bar{X}} + A \times \hat{\sigma}_e / \sqrt{n}$$

$$CL = \bar{\bar{X}}$$

$$LCL = \bar{\bar{X}} - A \times \hat{\sigma}_e / \sqrt{n}$$

이다. 여기서 $\bar{\bar{X}}$ 는 전체 자료의 평균이고, A는 확률한계를 나타내는 계수(보통 σ 으로 많이 잡으나 이 논문에서는 확률한계 1%를 만족시키는 2.575로 정한다)이며, $\hat{\sigma}_e$ 는 $\hat{\sigma}_{repeatability}$ 를 나타내고, n은 각각의 작업자가 각각의 제품에 대하여 실시한 반복 측정수를 나타낸다. R 관리도의 경우 관리상한선, 중심선, 관리하한선은 각각

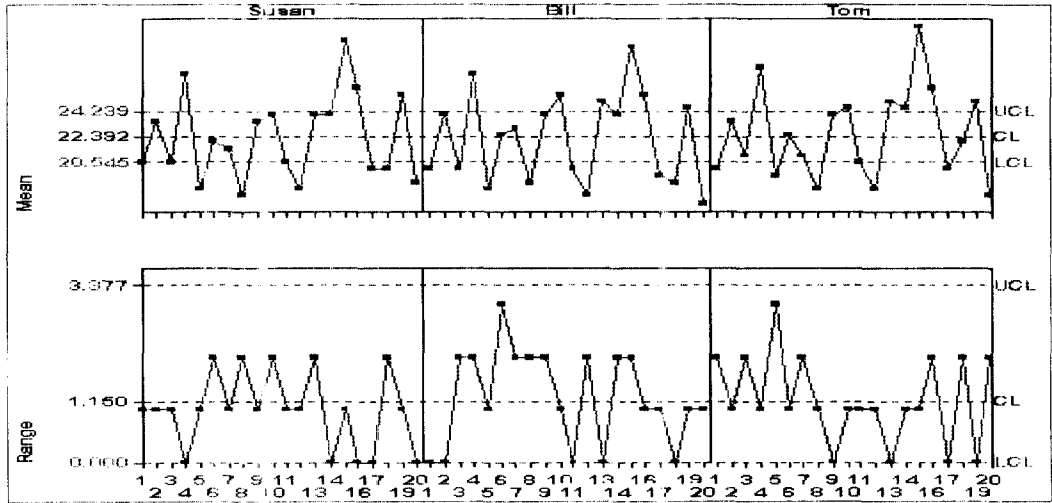
$$UCL = \bar{\bar{R}} + A \times \hat{\sigma}_e d_3$$

$$CL = \bar{\bar{R}}$$

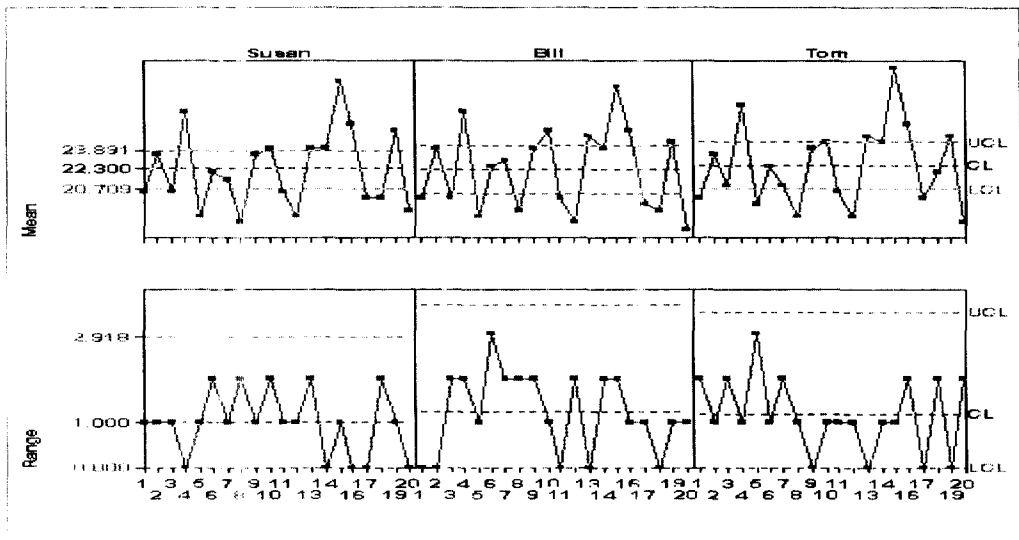
$$LCL = \bar{\bar{R}} - A \times \hat{\sigma}_e d_3$$

이다. 여기서 $\bar{\bar{R}}$ 는 범위들의 평균을 나타내며, d_3 는 n의 함수로서 Montgomery (1991)의 부록에서 그 값을 쉽게 발견할 수 있다.

<그림 2>의 아래쪽 그림은 계측기의 정밀도를 나타내며, 각각의 값들이 관리한계선 안에 있으므로 정밀도는 관리되어 있다고 본다. 반면에 위쪽 그림에서는 관측값들이 관리한계선을 많이 벗어나 있다. 그러나 별 문제될 것은 없을 것 같다. 그 이유는 그림에서 관리한계선은 측정오차에 근거를 두고 결정된 것이고, 각 제품간 편차(이 표준 편차가 제품의 품질관리용 관리도에 이용되는 것임)는 이런 측정오차보다는 아마 훨씬 더 클 것이기 때문이다.



< 그림 2 > \bar{X} 와 R 관리도 (세 작업자에게 동일한 관리한계선 적용)



< 그림 3 > \bar{X} 와 R 관리도 (세 작업자에게 각기 다른 관리한계선 적용)

<그림 3>은 앞의 <그림 2>와 똑같으나 다만 관리한계선이 각 작업자별로 따로 결정된 경우이다. 즉 <그림 3>은 세 작업자의 측정치의 표준편차가 모두 다르다고 가정하고 그린 그림이다. 구체적으로 \bar{X} 관리도의 경우 i 번째 작업자의 관리상한선

중심선, 관리하한선은 각각

$$UCL = \bar{\bar{X}}_i + A \times \hat{\sigma}_{ei} / \sqrt{n}$$

$$CL = \bar{\bar{X}}_i$$

$$LCL = \bar{\bar{X}}_i - A \times \hat{\sigma}_{ei} / \sqrt{n}$$

이다. 여기서 $\bar{\bar{X}}_i$ 는 i 번째 작업자의 평균을 나타내고, $\hat{\sigma}_{ei}$ 는 i 번째 작업자의 정밀도를 나타낸다. <표 2>의 자료의 경우 σ_{ei} 는 $1.00/1.128=0.8865$ 로 추정된다. 다음으로 R 관리도의 경우 i 번째 작업자의 관리상한선, 중심선, 관리하한선은 각각

$$UCL = \bar{\bar{R}}_i + A \times \hat{\sigma}_{ei} d_3$$

$$CL = \bar{\bar{R}}_i$$

$$LCL = \bar{\bar{R}}_i - A \times \hat{\sigma}_{ei} d_3$$

이다. 여기서 $\bar{\bar{R}}_i$ 는 i 번째 작업자의 범위들의 평균을 나타낸다. <그림 3>의 경우를 봐서도 계측기의 정밀도는 관리되어 있는 것으로 보인다.

<그림 4>는 계측자료의 재현성을 점검하는 그림으로서 평균을 관리하는 관리상한선, 중심선 및 관리하한선은 각각

$$UCL = \bar{\bar{\bar{X}}} + A \times \hat{\sigma}_e / \sqrt{pn}$$

$$CL = \bar{\bar{\bar{X}}}$$

$$LCL = \bar{\bar{\bar{X}}} - A \times \hat{\sigma}_e / \sqrt{pn}$$

으로 주어지며(여기서 p는 측정되는 제품의 수를 말함), 범위를 관리하는 범위관리도의 경우 관리상한선, 중심선 및 관리하한선은 각각

$$UCL = \bar{\bar{R}}^* (1 + A d_2 / d_3)$$

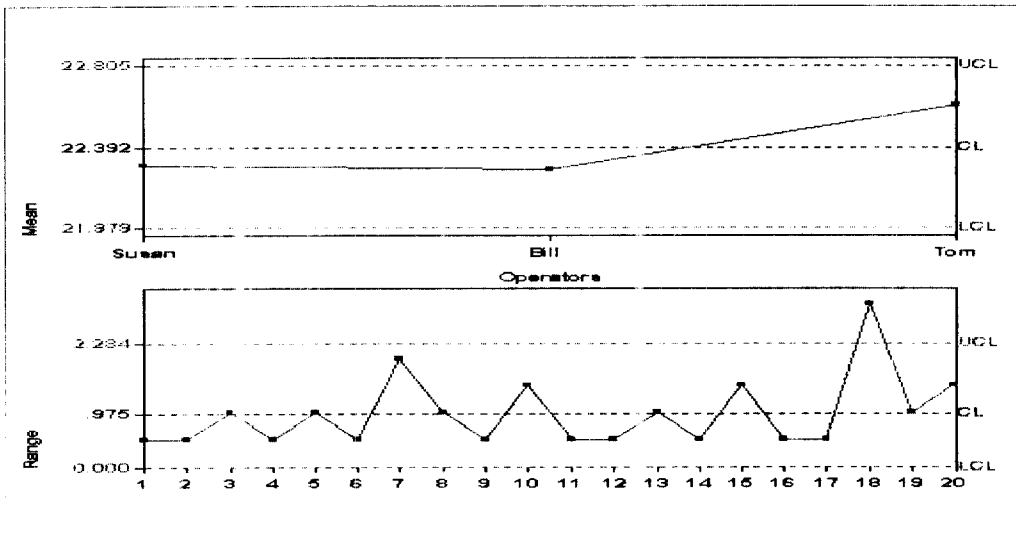
$$CL = \sum_{j=1}^k R_j^* (= \bar{\bar{R}}^*)$$

$$LCL = \bar{\bar{R}}^* (1 - A d_2 / d_3)$$

으로 주어진다. 여기서 R_j^* 는 j 번째 제품에서 3명의 평균측정치에 대한 범위를 말한다. 한 예로서 <표 2>의 경우 R_1^* 은 $20.5-20.0=0.5$ 이다.

<그림 4>의 아래 그림은 각 부품에 대한 계측자들의 계측결과가 제품에 따라 얼마

나 서로 상이한가를 보여주는 범위관리도이다. 이 그림에서 대부분의 점들이 관리한계선 안에 있는 반면 18번째의 제품만이 작업자들간 측정결과가 많이 다르다는 것을 알 수 있다. 따라서 이에 대한 점검이 필요하다. <그림 4>의 위쪽 그림은 각 측정자들의 평균치를 나타내는데, 이 그림을 봐서도 각 작업자들간 큰 차이가 없다는 것을 알 수 있다.



< 그림 4 > <표 2>의 자료에 대한 재현성 점검

4. 실험계획법

이제 실험계획법적인 사고방식을 이용하여 <표 2>의 자료를 분석해보기로 한다. 다음의 모형에서 X_{ijk} 는 작업자 i 가 제품 j 에 대해 k 번째로 측정한 관측값이다. 그러면 <표 2>의 자료에 대해서 다음과 같은 이인자 요인배치 모형을 적용할 수 있다.

$$X_{ijk} = \mu + O_i + P_j + (OP)_{ij} + R_{k(ij)}$$

여기서 μ , O_i ($i = 1, 2, \dots, o$), P_j ($j = 1, 2, \dots, p$), $(OP)_{ij}$, $R_{k(ij)}$ ($k = 1, 2, \dots, n$)는 각각 전체 모평균, 작업자의 주효과, 제품의 주효과, 작업자와 제품간 교호작용 효과 그리고 마지막으로 오차항을 나타낸다. 계측기 능력분석에서 보통 작업자 효과, 제품 효과, 작업자와 제품간 교호작용 효과 및 오차는 서로 독립적인 랜덤효과를 가지며, 평균이 0이고 분산이 각각 σ_o^2 , σ_P^2 , σ_{OP}^2 , σ_R^2 인 정규분포를 따른다고 가정한다.

특히 제품 효과와 작업자 효과는 랜덤효과로 취급된다. 왜냐하면 GR&R 분석에서 제품이나 작업자는 각각 많은 제품이나 또는 많은 작업자들로부터 랜덤하게 뽑히기 때문이다. 하지만 작업자의 수가 항상 똑같이 2내지 3명으로 고정된 경우 작업자 효과는 고정효과로 취급해야 한다.

이인자 요인배치모형에서 상정하는 가정들이 제대로 맞는다면 분산분석을 실시할 수 있으며, Montgomery와 Runger(1993a)의 <표 3>에는 그 결과가 있다. 그 표로부터 계측기의 정밀도, $\sigma^2_{repeatability}$ 는 0.99로 나오며, 이는 GR&R 분석에서 구한 1.02와 비슷하다는 것을 알 수 있다.

한편, 계측기의 재현성은 작업자와 관련된 측정자료의 변동이므로 우선 $\hat{\sigma}^2_o$ 로 추정할 수 있을 것이다. 그러나 작업자와 제품간 교호작용이 있을 수 있다. 예를 들어 어떤 작업자는 제품을 고정시킨다거나 다루는데 서툴러 작업 초기에 재는 제품의 품질특성치가 작업 말기에 재는 제품의 품질특성치와 서로 다를 수도 있다. 이러한 경우에는 두 요인간 교호작용 효과도 작업자에 의한 측정오차에 넣어야 한다. 따라서 계측기의 재현성은

$$\sigma^2_{reproducibility} = \sigma^2_o + \sigma^2_{op}$$

으로 표시할 수 있다. 그러나 <표 2>의 자료의 경우 σ^2_o 와 σ^2_{op} 을 추정하고, 이를 이용하여 계측기의 재현성을 추정하면 그 값은 이상하게도 0보다 작은 $\hat{\sigma}_{reproducibility} = -0.125$ 로 추정된다.

그런데 이와 같은 현상은 작업자와 제품간 교호작용 효과가 적은 경우 자주 일어난다. 이와 같은 현상을 극복하는 방법중의 하나는 어떤 요인의 분산이 0보다 작게 추정되면, 이를 0으로 두는 것이다. 그러나 Searle(1971)이 지적하듯이 이때는 다른 요인에 대한 분산추정치에 편의(bias)를 갖게 된다. 따라서 다음과 같은 수정된 분산 추정방법(Milliken과 Johnson(1984))이 더 적절하다고 생각된다. 이 방법은 자료가 주어진 경우 자료의 분산을 잘 설명하는 요인들만을 모두 골라 모형에 포함시키는 것이다. 구체적으로 설명하면, 이 방법은 모든 요인이 다 들어가 있는 모형으로부터 각 요인의 주효과 및 교호작용 효과에 대한 가설검정을 각각 실시하여, 만일 모든 요인이 주어진 유의수준에서 유의하다면 각 요인에 대한 분산을 추정하고, 그렇지 않고 어느 요인이든지 주어진 유의수준에서 유의하지 않다면 이를 모형에서 제외시키는 것이다. 이와 같은 작업은 마침내 모든 요인이 주어진 유의수준에서 대체적으로 유의하다고 판단될 때까지 계속된다. 이와 같은 방법으로 적절한 모형을 찾아 각 요소에 대한 분산추정치를 구하면 이제 그 추정치는 0보다 크게 된다.

그러면 <표 2>의 자료의 경우에는 Montgomery와 Runger(1993a)의 <표 4>에서 보는 바와 같이 $\hat{\sigma}^2_{repeatability} = 0.88$, $\hat{\sigma}^2_{reproducibility} = 0.011$ 로 나온다. 이 결과는 GR&R 분석에서 구한 값들과 조금 차이는 있지만 그렇게 심한 것 같지는 않다. 그러

나 작업자와 제품간 교호작용이 크다면 두 분석결과는 Montgomery와 Runger(1993a)의 또다른 예에서와 같이 다르게 나타날 수도 있다.

실험계획법을 이용한 분석은 분산분석이므로 각각 분산요소에 대한 통계적 해석이 용이하며, 각 추정량에 대한 신뢰구간의 추정 또한 Montgomery와 Runger(1993b)에서와 같이 가능하다. 따라서 GR&R 분석을 할 때에는 분산분석을 함께 실시하는 것이 좋다.

다음으로 실험계획법적인 접근방법과 GR&R 분석방법을 비교하기 위하여 GR&R 분석에서 이용되는 $\hat{\sigma}_{reproducibility}$ 에 대한 기대값을 구해본다. 우선 이인자 요인배치 모형으로부터 i 번째 작업자의 평균관측값은

$$\bar{X}_i = \mu + O_i + \frac{\sum P_j}{p} + \frac{\sum(OP)_{ij}}{p} + \frac{\sum \Sigma R_{k(ij)}}{pn}$$

로 추정되며, 계측기의 재현성은 \bar{X}_i 들로부터 범위를 구하여 수정항 d_2 로 나누어 구한다. 즉, $\hat{\sigma}_{reproducibility} = R \bar{X}_i / d_2$ 이다. 이제 계측기 재현성의 추정량 $\hat{\sigma}_{reproducibility}$

$= R \bar{X}_i / d_2$ 에 대한 기대값을 구해보면 $(\sigma_0^2 + \frac{\sigma_p^2}{p} + \frac{\sigma_{op}^2}{p} + \frac{\sigma_R^2}{pn})^{1/2}$ 이 되어 (Montgomery와 Runger(1993a)에 있는 결과는 잘못된 것임) 실험계획법에서 상정하는 계측기의 재현성 $(\sigma_0^2 + \sigma_{op}^2)^{1/2}$ 과 큰 차이가 있음을 알 수 있다. 이것을 봐서도 두 분석결과가 항상 똑같게 나오지 않을 수 있다는 것을 알 수 있다.

마지막으로 Montgomery와 Runger(1993a)의 133쪽에서 제안하는 계측기 재현성의 또다른 추정량에 대해 살펴본다. 거기서 제안하는 추정방법은 각각의 제품과 작업자에 대한 평균측정값

$$\bar{X}_{ij} = \mu + O_i + P_j + (OP)_{ij} + \frac{\Sigma R_{k(ij)}}{n}$$

를 우선 구한 다음, 각 제품별로 작업자들의 계측결과를 가지고 범위 R_j 를 구하여, 마지막으로 이 범위들의 평균 \bar{R} 를 작업자의 수를 고려한 수정계수 d_2 로 나눈다. 이를

$\hat{\sigma}_{reproducibility}^* = \bar{R} / d_2$ 라고 두고 그 기대값을 구해보면 $(\sigma_0^2 + \sigma_p^2 + \sigma_{op}^2 + \sigma_R^2/n)^{1/2}$ 으로 나와 (Montgomery와 Runger(1993a)에 있는 결과는 잘못된 것임), 이것 또한 실험계획법에서 상정하는 계측기의 재현성 $(\sigma_0^2 + \sigma_{op}^2)^{1/2}$ 과 큰 차이가 있음을 알 수 있다.

5. 결론

계측기 관리는 품질관리의 기본이므로 상황에 맞는 계측기 관리를 전개할 필요가 있다. 그러나 현실적으로 계측기 관리는 일년에 한두번 정도 교정을 하는 일로 대신하고 있을 뿐이다. 이러한 경우 계측결과를 있는 그대로 받아들여 공정의 상태를 질 못 판단할 수 있다. 따라서 계측기의 능력을 수시로 점검하여 계측기 때문에 생기는 문제를 없애야 한다.

계측기 능력분석의 목적은 계측기 능력을 제대로 파악하여 문제가 되는 부분은 고쳐서 측정관리를 원활히 하며, 궁극적으로 전체적인 공정관리를 효율적으로 하는데 그 목적이 있다. 본 논문에서는 계측기 능력분석에서 많이 사용되는 measurement error study, GR&R 및 실험계획법적 접근방법에 대하여 간단히 살펴보았다.

첫 번째 방법은 분석이 간단하다는 장점은 있으나 계측기의 능력이 부족한 경우 어느 요인에 문제가 있는지 전혀 알 수가 없다. 아울러 분석결과가 너무 단순하므로 이에 통계적 해석을 부가하여 잘못된 결론에 이르는 것을 미연에 방지해야 할 필요가 있다.

두 번째 방법 또한 분석이 간단하다. 그러면서도 계측기 능력에 문제가 있다면 이것이 계측기 자체의 문제 때문인지 또는 계측자 때문에 생기는 문제인지 쉽게 파악할 수 있다. 그러나 이 방법은 각각의 요인에 의한 계측자료의 변동 등을 점추정치 하나로만 나타낸다는 단점이 있다. 두 번째 방법에서는 주어진 자료를 관리도와 같은 그림으로 나타내면 계측현상을 한눈에 파악할 수 있어서 분석이 훨씬 더 용이하다.

세 번째 방법으로, 처음 두 방법에 비해 보다 많은 이론이 정립된 실험계획법적 분석방법을 이용하여 계측자료를 분석할 수 있다. 본 논문에서는 특히 이인자 요인배치법을 들고 있는데, 이 분석방법을 이용하면 작업자 및 제품의 주효과, 작업자와 제품간 교호작용에 대한 효과를 쉽게 검정할 수 있으며, 작업자와 제품간 교호작용이 큰 경우에는 계측기의 재현성을 기존의 GR&R 분석방법보다 편의가 없게 더 잘 추정할 수 있다.

한편 본문에서는 반복이 있는 이원 요인배치 변량모형을 이용하여 계측기의 정밀도 및 재현성에 대한 점추정에만 관심을 두었는데, 사실 이들의 구간추정은 물론 계측기의 정밀도나 재현성이 일정한 범위안에 있게 하기 위한 표본의 크기 등에 대해서도 알아볼 필요가 있다. 아울러 우리는 지금까지 변량모형에 대해서만 살펴보았는데, 품질특성치를 측정하는 작업자가 항상 같아서 작업자가 변량인자가 아닌 모수인자로 취급되어야 할 필요가 있을 수 있다. 이 경우에는 변량인자와 모수인자를 함께 고려한 혼합모형을 이용하여 자료를 분석해야 한다. 또한 본문에서는 모든 작업자가 똑같은 제품을 측정한다고 가정하므로 요인배치법을 사용했는데, 각각의 작업자가 지역적으로 떨어져 있기 때문에 서로 다른 제품들을 측정하게 되는 경우도 있다. 그러나 이 경우에도 적절한 실험계획법의 모형이 이용될 수 있다. 다음으로 작업자나 제품 요인 이외에 또다른 요인이 측정결과를 변동시킬 수 있다. 예를 들어, 계측장비가 여러 대인

경우 이것 또한 변동요인으로 넣어야 하므로, 이때에는 삼원배치법을 이용해야 한다. 이와 같은 복잡한 상황하에서도 실험계획법의 모형은 쉽게 응용되어 분석을 용이하게 해준다. 그러나 이원배치나 삼원배치에서 상정하는 가정들이 제대로 맞아야 분석결과가 의미가 있다는 것을 일러둔다. 마지막으로, 지금까지는 계량형의 자료의 경우 어떻게 분석을 실시하는지 생각해보았는데, 계수형의 자료에 대해서도 똑같은 분석을 해야 할 것이다.

참고문헌

- [1] Deming, W. E.(1982), *Quality, Productivity and Competitive Position*, MIT Center for Advanced Engineering Study, Cambridge, MA.
- [2] Deutler, T.(1991). "Grubbs-Type Estimators for Reproducibility Variances in an Interlaboratory Test Study," *Journal of Quality Technology*, Vol. 23, No. 4, pp. 324-335.
- [3] Floyd, D. A. and Laurent, C. J.(1995), "Gauging: An underestimated consideration in the application of statistical process control," *Quality Engineering*, Vol. 8, No. 1, pp. 13-29.
- [4] Hemmerle, W. J. and Hartley, H. O.(1973), "Computing Maximum Likelihood Estimates for the Mixed A.O.V Model Using the W Transformation," *Technometrics*, Vol. 15, pp. 819-831.
- [5] Milliken, G. A. and Johnson, D. E.(1984), *Analysis of Messy Data*, Volume 1: *Designed Experiments*, Van Nostrand Reinhold, New York.
- [6] Montgomery, D. C.(1991), *Introduction to Statistical Quality Control*, 2nd ed., John Wiley & Sons, New York.
- [7] Montgomery, D. C. and Runger, G. C.(1993a), "Gauge Capability Analysis and Designed Experiments. Part I," *Quality Engineering*, Vol. 6, No. 1, pp. 115-135.
- [8] Montgomery, D. C. and Runger, G. C.(1993b), "Gauge Capability Analysis and Designed Experiments. Part II," *Quality Engineering*, Vol. 6, No. 2, pp. 289-305.
- [9] Searle, S. R.(1971), *Linear Models*, John Wiley & Sons, Inc. New York.
- [10] Tsai, P.(1988), "Variable Gage Repeatability and Reproducibility Study Using the Analysis of Variance Method," *Quality Engineering*, Vol. 1, No. 1, pp. 107-115.