품질실험계획 모형의 특징 및 적용방안
Implementation and Features for Design of Experiment

최 성 운*
Sung-Woon Choi*

Abstract

The research presents implementation strategy and features for experimental design of quality. The MINITAB models of experimental design classfies by the kinds of factors, orthogonality, blocking, confounding and rotationability. The results discussed in this paper can be easily used for practictioners.

Keywords : Implementation, Features, MINITAB Model, Experimental Design, Factor, Orthougonality, Blocking, Confounding, Rotationability.

1. 서론

고객이 요구하는 품질 스펙(Specification)에 대해 기업에서 생산한 실제 품질 데이터가 만족하는 가를 확인, 평가하는 통계적 품질 및 공정관리(Statistical Quality and Process Control)와는 달리 품질실험설계(Design of Experiment for Quality)는 신제품 개발을 위한 제품기술조건 또는 개량제품개발을 위한 생산기술조건을 새롭게 설정하기 위한 용도로 사용된다.

* 경원대학교 산업공학과

품질실험설계는 일본형 타구치 실험설계 방법과 서구형 실험설계 방법으로 구분된다. 타구치 실험설계에서는 특성값을 능동적, 수동적 계량연속형 데이터의 동특성과 망소, 망대, 망목 계량연속형 데이터와 망소, 망대 계수이산형 데이터인 정특성으로 구분한다. 인자는 특성값에 일정한 영향을 주는 상수인자, 수준을 컨트롤할 수 있는 제어인자, 실험가능하나 제어 불가능하고 특성값에 영향을 주는 잡음인자, 능동적 동특성과 비례관계를 갖는 신호인자, 실험불가능하고 특성값에 영향을 주지 않는 오차 인자 등으로 분류된다. 타구치 실험설계에서는 SN비(Signal-to-Noise Ratio)와 생산 스펙에 대한 고객의 기능한계(Functional Limit)를 고려한 품질손실함수(Quality Loss Function) 개념을 고려한다. 미국형 실험설계에서는 고객이 요구하는 스펙을 창조, 개량하기 위해서 계량연속형 데이터인 특성값(Characteristic Value, Dependent Variable, Response)이 상한스펙(USL), 하한스펙(LSL), 양쪽스펙의 실험목표를 달성할 수 있도록 설계하는 통계적인 기법이다. 실험설계의 모형은 특성값에 영향을 주는 인자(Factor, Independent Variable)의 종류와 수준(Level, Treatment)의 실험 랜덤화 순서(Randomization Order)에 따라 달라진다. 기술적인 재현성 여부에 따라 모수고정인자(Fixed Factor)와 랜덤변량인자(Random Factor)로 분류되는데 모수고정인자는 제어가능하고 콘트롤이 가능하여 재현성이 있어 모평균의 추정에 의미가 있으며, 랜덤변량인자는 집단, 날짜, 블록(Block)과 같이 재현성이 없는 경우 모분산의 추정에 의미가 있다. 실험 대상인 부품 또는 제품의 재사용 여부에 따라 재사용이 가능한 경우 교차인자(Crossed Factor), 파괴검사와 같이 재사용이 불가능한 경우 지분인자(Nested Factor)로 분류되며 기술적 관점에서 모수고정인자가 완전 랜덤화 설계(CRD: Completely Randomized Design)가 곤란한 경우나 랜덤변량인자인 샘플링 오차를 설계할 경우에 적용된다. 반복(Replication)은 CRD를 원칙으로 하나 실험의 경제성과 효율성을 위해 되풀이(Repetition)를 취하기도 한다.

이와 같이 품질실험설계는 특성값의 차수[1,4-6], 인자의 종류[2-3,8],반복의 형태에 따라 실험순서와 설계방법이 다른데도 불구하고 전문가와 기업실무자의 이해부족으로 실험설계보다는 통계패키지에 의한 분석에 치중하여 그릇된 의사결정을 내리는 일이 빈번하다.

따라서 본 연구에서는 식스시그마운동과 품질분임조활동에서 가장 많이 사용되는 품질실험계획 모형[7,9]의 특징과 적용방안을 MINITAB 사용자가 이해하기 쉽도록 유형화하여 제시한다.

2. ANOVA 모형의 특징 및 적용방안

MINITAB에서 지원되는 ANOVA(Analysis of Variance)방법으로는 일원, 이원ANOVA, 균형 ANOVA, GLM(Generalized Linear Model), 완전내포분석, ANOM(Analysis of Means), MANOVA(Multivariate ANOVA)등이 있으나 본 연구에서는 특성값에 영향을 주는 인자(Factor)의 종류에 따른 모형의 특징 및 적용방안을 논의하기로 한다.

ANOVA는 3개이상의 모평균을 비교하기 위해 사용되는 분산분석의 F검정방법이다.

인자가 1개일 경우 일원 ANOVA, 2개일 경우 이원 ANOVA라고 하며 수준간의 평균차와 수준내의 오차와의 비교를 통한 유의성을 판정하는 방법이다. 일원배치법의 ANOVA에 대한 모수적 표현방법으로 데이터구조 모형 $y_{ij} = \mu + a_i + e_{ij}$를 사용하는데 $(y_{ij} - \mu) = e_{ij} + a_i$은 통계량 $(y_{ij} - \overline{\overline{y}}) = (y_{ij} - \overline{y_{i.}}) + (\overline{y_{i.}} - \overline{\overline{y}})$으로 표현되며 좌변과 우변은 $\overline{y_{i.}}$를 소거하면 같아진다. 이 식으로 편차제곱합(Sum of Square, Variation), 자유도(Degree of Freedom), 평균제곱(Mean Square, Unbiased Variance)을 구하여 a_i와 e_{ij}에 대한 평균제곱의 비율인 F 검정을 실시한다. 평균제곱은 2개 이상의 데이터의 산포를 비교하기 위한 불편분산으로 기하학적으로 수평선을 기준으로 한 높낮이의 차이를 알아보기 위한 방법이다. 회귀분석은 직선 또는 적합함수에 따른 적극적인 분석방법이다. ANOVA는 e_{ij} 잔차에 대한 정규성, 불편성, 등분산성, 독립성의 가정을 잔차히스토그램, 잔차 정규확률도, 잔차대 적합치, 잔차대 순서의 개별그림으로 판독하는데 정규성 검정, 등분산 검정, ARIMA시계열 분석을 보조적으로 활용하는 것이 바람직하다.

일원, 이원 ANOVA는 a_i가 모평균이 의미가 있는 모수고정인자인 경우 적용되며 모분산이 의미가 있는 랜덤변량인자는 균형 ANOVA와 GLM을 활용해야 한다. 랜덤변량인자만으로 구성된 샘플링, 측정오차를 설계할 경우 완전내포분석을 사용한다. 교호작용이 의미가 없는 라틴, 그레코, 초그레코 라틴방격법은 균형 ANOVA, GLM, 타구치 설계등을 이용하며, 모수고정인자와 랜덤변량인자의 혼합모형(Mixed Model)인 난괴법(Randomized Block Design)은 교호작용 $(ab)_{ij}$의 모수고정인자의 합을 영(Zero)으로 제약하는(Restricted) 경우와 제약하지 않는(Unrestricted) 경우 모두를 알고 싶은 경우 균형 ANOVA를 사용한다. 제약혼합모형은 독립성을 상실하는 대신에 분석의 용이함이 있고 비제약 혼합모형은 독립성을 유지하는 대신에 Satterthwaite 자유도를 사용한 복잡한 근사 F_o검정을 실시해야 한다. MINITAB에서는 비제약 혼합모형을 Default로 설정하고 있으므로 실무자는 사용수준에 따른 선택을 고려해야 한다. 통상 생산, 측정현장의 ANOVA는 제약혼합모형을, 고도의 정밀한 측정정밀도를 평가하는 Gage R&R ANOVA인 경우 비제약 혼합모형을 사용하나 전문가들의 두가지 사용에 대한 장단점의 논란은 아직도 진행중이다.

실험대상이 되는 부품, 제품의 재사용(Recycle)이 가능한 경우 완전랜덤화 설계(CRD: Completley Randomized Design)가 가능한 교차인자와 파괴검사와 같이 재사용이 불가능한 지분인자로 구분되는데 교차인자인 경우 CRD의 반복(Replication)이 가능한 반면 지분인자인 경우 특정조건에서의 되풀이(Repetition)만 가능하다. 모수고정인자가 시간, 비용, 설비특성관점에서 완전랜덤화가 곤란한 경우 분할법(Split-Plot Design)을 사용하는데 1차 단위가 일원,이원배치인 단일 분할법, 2차분할법(Split-Split-Plot Design), 이방분할법(Multi-Cell Design)이 있으며 균형ANOVA를 활용한다. MINITAB의 MSA(Measurement System Analysis)에서 지원하는 Gage R&R은 교차인자모형과 지분인자인 내포모형이 있는데 두 모형 모두 부품과 계측자를 랜덤 샘플링하는 변량인자로 가정한다. 부품, 계측자 이외에 다양한 계측조건을 고려하는 Gage R&R 교차인자와 지분인자의 결합모형(Combined Model)을 사용하는데 이

경우 Sattethwaite 자유도와 근사F 검정을 실시해야 한다.

QC Story 15단계중 대책실시 단계에서 가장 많이 사용되는 ANOVA는 일원, 이원, 반복있는 이원, 삼원배치법이나 오차항의 가정검토가 누락되어 있어 이에 대한 보완이 요구되며 교호작용의 유무에 따른 반복의 선택도 실무적 관점에서 선행적으로 고려되어야 한다. DMAIC 5단계중 Measure 단계에서 측정시스템의 Gage R&R 사용시 교차인자 모형과 지분인자의 내포모형의 선택이 실험대상의 재사용 여부에 따라 이루어져야 한다. 예를들어 공기업에서 지역마다의 시설물에 대한 Gage R&R을 적용할 경우 Gage R&R(내포)모형을 사용해야 한다.

3. DOE 모형의 특징 및 적용방안

MINITAB에서 지원되는 DOE방법으로는 k^n 완전요인 배치법(CFD: Complete Factorial Design), k^{n-p}(FFP: Fractional Factorial Design), Plackette-Burman 설계, RSD(Response Surface Design), 혼합물 설계(Mixture Design), 타구치 설계(Taguchi Design)등이 있다.MINITAB에서 ANOVA모형보다는 DOE(Design of Experiment)모형으로 분류하는 이유는 이 방법들이 직교성, 중심점, 블록화, 교락화, 회전성 등의 인자수준들을 설계하는 특징에 연유한다.

k^n CFD는 수준 k를 고정하는 대신 많은 인자 n을 배치하는 선별형(Screening)실험계획모형이다. $2^n CFD$인 경우 효과(Effect)=(합의 대비(Contrast))/(총실험횟수/2)의 두 수준간의 평균차를 파악할 수 있으며 편차 제곱합 S=(합의 대비)^2/총실험횟수, 자유도 $\nu=1$로 ANOVA를 계산한다.

고차의 교호작용을 주효과 또는 저차의 교호작용과 교락하여 실험횟수를 절약하는 경우 $k^{n-p}FFP$의 실험계획모형을 사용한다, 교락되는 요인은 n인자수에 따른 해상도(Resolution)의 별명(Alias)으로 파악하여 실험을 설계한다. 주효과만을 검토하고자 할 경우 해상도 Ⅲ 또는 Plackette-Burman 설계를, 주효과와 2차 교호작용을 검출하고자 할 경우 해상도 Ⅳ, 해상도 Ⅴ를 사용한다. 해상도Ⅲ, Ⅳ,Ⅴ는 (1+2), (1+3, 2+2), (1+4, 2+3)의 표현으로 Ⅲ인 경우 주효과와 2차 교호작용이, Ⅴ인 경우 주효과와 4차 교호작용 또는 2차 교호작용과 3차 교호작용이 별명관계로 교락되어 불필요한 고차의 교호작용은 검출할 수 없다는 의미이다.

$2^n CFD$는 2 수준으로 직선효과만을 검출할 수밖에 없기 때문에 중심점(CP: Center Point)에 의한 곡률(Curvature)효과를 파악하기 위한 실험을 계획한다. $2^2 CFD$+중심점 실험계획은 인자 I, x_1, x_2, x_1x_2에 의한 직교행렬과 특성치 y에 의한 $y=x\beta+\epsilon$의 1차 반응표면식을 $\hat{\beta}=(x'x)^{-1}x'y$로 추정하고 $\hat{y}=\hat{\beta_0}+\hat{\beta_1}x_1+\hat{\beta_2}x_2+\hat{\beta_{12}}x_1x_2$의 계수 $\hat{\beta}$에 대한 효과를 $2\times\hat{\beta}$ 또는 두 수준의 평균차인 효과 공식으로 구한다. 계수와 중심점의 P-Value로 유의성을 판정하여 특히 중심점의 P-Value가 α보다 클 경우 곡률이 존재하지 않는다는 것을 암시한다. ANOVA에서 편차 제곱합은 $S_{계수}$=((1)수준평균--(-1)수준평균)^2,

$S_{잔차오차} = \sum(y-\hat{y})^2$, $S_{곡률} = n_f n_c(\overline{y_c}-\overline{y_f})/(n_f+n_c)$, $S_{순수오차} = \sum(y-\overline{y_c})^2$ 으로 계산되며 첨자 f는 요인(Factorial)점, 첨자 c는 중심(Center)점이다. 곡률효과가 유의적인 경우 2차 회귀방정식의 RSD를 적용해 보는 것이 바람직하다.[8]

비반복 $2^k CFD$인 경우 인자의 유의성을 판정하는 경우 효과에 대한 정규확률도, 파레토 차트와 함께 사용되는 방법이 Length 방법이다. 각 인자의 효과를 C_j라고 했을 때 $S_0 = 1.5 \times median(|C_j|)$, PSE(Pseudo Standard Error)$=1.5 \times median(|C_j| < 2.5S_0)$, ME(Margin of Error)$=t_{1-\frac{\alpha}{2}}((2^k-1)/3) \times PSE$, SME(Simultaneous ME)$=t_p((2^{k-1})/3)$ PSE, 여기서 $p = 1-(1+0.95^{1/(2^k-1)})/2$이다. 각 인자의 효과가 계산된 ME와 SME보다 클 경우 정규확률도의 직선에서 멀어지고 파레토차트는 소수중점항목으로 선정된다.[8]

2차 회귀모형인 RSD는 2^k요인점, 2^k축점, n_0중심점의 직교블럭화에 의한 CCD(Central Composite Design)모형이다. 이론 회전성 $\alpha = (2^k)^{1/4}$에 대해 실제 회전성 $\alpha = (n_{1f}(n_{2f}+n_{2c})/2(n_{1f}+n_{1c})^{\frac{1}{2}}$를 비교평가한다. 요인꼭지점과 축점의 극단적인 수준을 피할 경우 Box-Behnken 설계 방법을 사용하며 축점을 면의 중심으로 하는 경우 α값에 내접설계를 하는 면중심(Face-Centered RSD) RSD를 사용한다. x의 관심영역(Region of Interest)에 의한 변수변환에 따라 Coded, Uncoded Variable로 $\hat{\beta} = (x'x)^{-1}x'y$를 추정하는데 $(\hat{x}x)^{-1}$의 행렬식, 평균 Trace, 고유값에 따라 D, A, E-Optimal의 최적실험이 설계된다. 최적화를 위해 회귀식은 고유값과 고유벡터에 의한 스펙트랄 분해(Spectral Decomposition)로 표현되며 정준분석(Canonical Analysis)의 도함수에 의해 최대, 최소, 안장점(Saddle Point)의 최적해가 구해진다. 고유값이 모두 양수인 경우 최대값, 음수인 경우 최소값, 섞인 경우 안장점이 나오며 시각화를 위해 2차원의 등고선도와 3차원의 표면도, 만족함수에 의한 반응 최적화 방법을 사용한다. 다중(Multiple) RSM인 경우 스펙 특성값간의 상관관계가 고려되지 않은 종합 만족도 함수를 사용한다는 점에 유의해야 하며 특히 개별 만족도의 가중치와 중요도의 실무적 이해를 전제로 값을 부여해서 사용해야 한다.

혼합물 설계는 비율(Proportion)의 합이 1이 되어야 한다는 조건 $x_1 + x + \cdots + x_k = 1$로 회귀식의 상수항 β_0가 생략된 실험계획모형이다. 상한, 하한 제약이 있는 경우 꼭지점(Extreme Vertices)을 사용하여 꼭지점, 2중, 3중 혼합, 중앙점, 축점의 설계 방법에 따라 심플렉스 중심, 격자 설계 방법을 사용한다. 분석모형으로는 1차, 2차, 특수 3차, 완전 3차 모형이 있다.

타구치 설계는 잡음(Noise)인자의 설정이 중요하며 고객의 신뢰성 사용조건인 온도, 습도, 진동과 제어할 수 없는 생산조건을 선정한다. 동특성의 신호인자와 망목, 망대, 망소의 정특성에 대한 설계분석 방법이 MINITAB에서 지원되고 있으나 국가품질경영대회에 출전한 식스시그마 분임조에서 그다지 많이 사용되고 있지 않다.

특정 인자를 선정하고 나머지 인자를 Pooling하거나 고려하지 않고 실험을 설계할 경우 선형, 제곱, 교호작용의 회귀분석, 잔차오차(RE:Residual Error), 적합성 결여

(LOFE: Lack of Fit Error), 순수오차(PE : Pure Error), 전체(Total)에 대한 분산분석을 실시한다. 이 경우 RE= PE+LOFE인데 $(y_{ij}-\widehat{\overline{y_{i.}}})=(y_{ij}-\overline{y_{i.}})+(\overline{y_{i.}}-\widehat{\overline{y_{i.}}})$의 관계식에 의해 편차제곱합, 자유도, 평균제곱으로 계산되어 유의성을 판정한다. 여기서 LOFE의 P-Value가 α보다 큰 경우 모형이 적합하다는 귀무가설을 채택한다.

DMAIC의 5단계중 잠재인자에서 치명인자를 선정하는 Analyze 단계에서는 Plackett-Burman 설계, $2^{k-p}FFP$등으로 주효과 중심의 인자를 선정하여 개선을 실시하고 Improve 단계에서 교호작용, 곡률효과, 2차 효과를 검출하기 위해 $2^{k}CFD$, RSM등을 사용하는 것이 바람직하다.

4. 결 론

본 연구에서는 QC Story 15단계, DMAIC 5단계에서 측정시스템 분석, 치명인자 선정, 개선대책을 실시할 경우 적용되는 MINITAB지원 품질실험설계 모형의 특징과 활용방안을 제시하였다. 특히 인자재현성에 따른 모수고정인자, 랜덤변량인자와 실험대상의 반복사용에 따른 교차인자, 지분인자에 의한 ANOVA 모형의 활용방안을 논의하였으고 직교성, 블록화, 교락화, 중심점에 의한 실험설계 모형의 특징과 활용방안도 제안하였다.

5. 참 고 문 헌

[1] 최성운, "안전 및 환경적용을 위한 최소 실험계획", 대한안전경영과학회지, 7(5)(2005) : 69-84.
[2] 최성운, "GLM에서 제약과 비제약 혼합모형의 고찰 및 확장", 2009년 대한안전경영과학회 춘계학술대회 발표논문집,(2009): 185-192.
[3] 최성운, "일반화 선형모형의 디자인 행렬을 이용한 품질실험설계", 2009년 대한안전경영과학회 춘계학술대회 발표논문집, (2009) : 423-427.
[4] Box G.E.P, Draper N.R., Empirical Model-Building and Response Surfaces, Wiley, 1987.
[5] Box G.E.P, Draper N.R, Response Surfaces, Mixtures, and Ridge Analysis, 2nd Edition, Wiley, 2007.
[6] Cornell J., Experiments with Mixtures : Designs, Models, and the Analysis of Mixture Data, 3rd Edition, Wiley, 2002.
[7] Dey A., Murerjee R., Fractional Factorial Plans, Wiley, 1999.
[8] Montgomery D.C., Design and Analysis of Experiments, 6th Edition, Wiley,2005.
[9] Mukerjee R., Wu C.F.J., A Modern Theory of Factorial Design, Springer, 2006.