블록 반복측정을 이용한 품질통계 모형의 유형화

최 성 운*
*경원대학교 산업공학과

Model Classification of Quality Statistics Using Block Repeated Measures

Sung Woon Choi*
*Department of Industrial Engineering, Kyungwon University

Abstract

Dependent models in quality statistics are classified as serially autocorrelated model, multivariate model and dependent sample model. Dependent sample model is most efficient in time and cost to obtain samples among the above models.

This paper proposes to implement parametric and nonparametric models into production system depended on demand pattern. Nonparametric models have distribution free and asymptotic distribution free techniques. Quality statistical models are classified into two categories ; the number of dependent sample and the type of data. The type of data consists of nominal, ordinal, interval and ratio data. The number of dependent sample divides into 2 samples and more than 3 samples.

Keywords : Dependent Sample, Parametric, Nonparametric, Distribution Free, Asymptotic Distribution Free, Demand Pattern, Type of Data, Number of Sample

1. 서 론

품질통계에서는 모집단(로트, 공정)에서 랜덤 샘플링한 표본(샘플, 시료)을 동등하고 독립인 분포(iid : Identically and Independent Distributed)의 데이터 즉 랜덤 샘플(Random Sample)의 독립표본 모형으로 가정한다.

그러나 계측과 측정기술의 발달로 샘플링 빈도(Sam-pling Frequency)가 높아짐에 따라 데이터간의 간격이 좁아지고 자동상관(Autocorrelation)이 되는 Box Jenk-ins의 ARMA(Autoregressive and Moving Average), ARIMA(AR and Intergrated MA), ARMAX(ARMA with Exogeneous Variable) 등의 시계열(Serial) 종속 모형으로 가정된다.

또한 고속컴퓨터 통계 소프트웨어패키지의 출현으로 측정변수간의 공분산(Covariance)과 상관계수가 존재하는 경우 복잡한 계산을 효율적으로 분석, 평가할 수 있다. 이러한 다변량(Multivariate) 종속모형은 개체(Obs-ervation, Individual, Object, Case, Experimental Unit)간의 관계를 규명하는 판별, 분류, 클러스터링, 다차원 스케일링, 대응(Correspondence) 분석과 변수(Variable, Characteristic, Attribute) 간의 관계를 규명하는 주성분, 인자, 행렬도(Biplot), 정준, 경로(Path) 분석 등이 있다.

그리고 품질통계에서 실험계획을 실시할 경우 파일롯 플랜트의 구조상 한번에 몇 개의 샘플만을 특정 인자처리조건에서 실험이 가능할 경우 CRD(Completely Randomized Design)의 반복(Replication)을 실시하게 되면 시간과 비용의 관점에서 많은 샘플이 필요로 하게 되므로 비효율적인 실험이 된다. 이 경우 특정 인자처리 조건을 블록으로 하여 반복 측정(Repeated Measures)을 실시하는 실험계획을 이용할 경우 적은 샘플로 효율적인 실험이 가능해진다.

2007년 5월 접수; 2007년 6월 수정본 접수; 2007년 6월 게재확정

특히 의료, 건강, 안전통계(Medical, Health and Safety Statistics)와 같이 실험대상이 되는 환자, 안전 근로자를 선정하기 어렵거나 시간에 따른 추이를 관찰하고자 할 경우 블록에 의한 종속(Dependent)표본 모형을 사용하면 좋다.

이와 같이 품질통계의 종속형 모형은 시계열 종속모형, 다변량 종속모형, 종속표본 모형 등의 세가지 모형으로 분류된다.

본 연구에서는 종속표본 모형을 대상으로 생산형태에 따른 모수 통계모형과 비모수 통계모형을 유형화하여 명목(Nominal, Categorical), 순위(Ordinal, Ranking), 구간(Interval), 비율(Ratio)자료 등의 적용방안을 제시한다.

명목 자료는 제품을 검사할 때 상품, 중품, 하품과 같이 2개 이상의 범주(Category)로 분할표(ContingencyTable)를 이용할 수 있을 경우에 적용되며, 2개 범주인 경우는 불량(부적합품), 결점(부적합) 등의 계수치 자료에 적용된다. 순위형 자료는 고객 만족 조사시 여러 제품에 대한 순위로 제품을 평가할 경우 사용된다.

구간자료는 섭씨, 화씨 온도와 같이 절대적인 크기를 비교할 수 없는 경우에 적용되며, 비율 자료는 품질 통계에서 계측기에 의해 구해지는 모든 계량치 자료에 적용된다. 명목, 순위자료는 이산형(Discrete) 자료로서 비모수통계 모형의 적용대상이 되며 구간, 비율자료는 연속형(Continuous) 자료로서 분포를 가정하는 대표본(L-arge Sample)의 모수통계 모형을 적용하는 것이 바람직하나 모집단의 분포나 모수를 가정하지 않고 소표본(Small Sample) 혹은 중표본(Medium Sample)의 비모수통계모형을 활용할 수 있다.

소품종 대량생산에서는 품종이 제약되므로 모집단의 누적된 정보를 분포로 활용할 수 있어 대표본의 모수통계 모형을 적용할 수 있다. 그러나 다품종 소량생산에서는 품종의 다양성으로 인하여 특정 제품의 모집단에 관련된 분포를 알 수 없으며 비반복적 생산인 경우 소표본의 비모수통계 모형을 사용해야만 한다. 한편 반복적인 다품종 소량생산에서는 누적된 중표본을 대상으로 정규근사 분포를 이용하는 비모수통계 모형을 적용하는 것이 바람직하다.

2장의 2개 종속표본 모형에서는 구간, 비율자료에 대한 t_0검추정의 분포를 가정하는 대표본의 모수통계 모형을 제안하고 범주자료의 McNemar 모형, 순위자료의 부호, Wilcoxon 모형, 구간비율 자료의 Walsh, 무작위 모형 등의 분포를 가정하지 않는 소표본, 중표본의 비모수통계 모형을 제시한다. 3장의 3개 이상 종속표본 모형에서는 구간, 비율자료의 난괴법, 라틴 방격법, 단일 분할법, 다단계 분할법, 2단 분할법, 범주자료의 2방 분할법 등의 모수통계 모형을 제안하고 범주자료의 Cochran 비모수통계 모형, 순위자료의 Friedman, Page, 다중비교 모형 등의 분포무관(Distribution Free) 모형과 순위자료의 Doksum, Hollander 모형 등의 점근분포무관(Asymptotic Distribution Free) 모형을 제시하고 4장에서 결론을 맺는다.

2. 2개 종속표본 모형

2.1 모수통계 모형 : 구간, 비율자료

모수통계 모형은 구간, 비율자료를 대상으로 모집단의 분포를 알고 있는 경우 정규성, 독립성, 불편성, 분산성, 비공선형성 등의 가정을 만족하는 모수를 이용하는 방법이다. 품질통계에서 평균은 z, t분포, 분산은 χ^2, F분포를 이용하고 부적합품(불량)은 간략화 관점에서 초기하, 이항, 포아송 분포를 조건에 따라 선택하고 부적합(결점)은 포아송분포를 사용한다.

신뢰성 통계에서는 수명을 확률변수로, 초기고장기간은 와이블, 감마분포, 우발고장기간은 지수, 포아송분포, 와이블, 감마 분포 마모고장기간은 정규, Rayleigh, 대수정규, 와이블, 감마분포 등으로 가정하여 모수를 이용한다.

소품종 대량생산인 경우 대량의 데이터를 쉽게 구할 수 있고 한정된 품종의 누적된 정보로 모집단의 분포나 모수를 가정할 수 있다. 2개 종속표본 모형인 경우 종속표본의 차이(Difference) d는 정규성과 독립성을 만족하는 분포로 가정되어야 t분포를 사용할 수 있다. 2개 종속표본의 모수통계 모형은 t_0차이 검정으로 검정전에 모집단의 정규분포 혹은 독립성의 가정여부를 알기 위해 누적된 표본의 정보 즉 대표본을 사용해야만 한다. t_0검정은 차이 데이터 d의 평균($\overline{d}$)과 표준편차(s_d)를 이용하여 검정하는 방법이며 t분포를 이용하여 판정한다. 검정 통계량 $t_0 = \dfrac{\overline{d}-\delta}{s_d/\sqrt{n}}$ 이며 δ는 모집단 차이 모수값이다.

분석결과를 판정하는 방법은 Neyman-Pearson의 가설 검정(Testing Hypothesis)과 Fisher의 유의성 검정(Test of Significance) 등이 있다.[7] 전자는 유의수준(Significance Level) α를 5% 또는 1%로 미리 정해놓고 검정 통계량 값에 의한 유의확률(Significance Probability) P-Value가 α보다 작게 나오면 유의적인 의사결정(Decision Making)을 하는 방법이다. 후자는

전자의 방법에서 유의수준 α의 합리적인 설정에 의문을 제기하며 검정 통계량 값에 대한 유의확률 P-Value로 검정사안에 따라 검정이용자가 유의성의 강약을 판단하는 것이 바람직하다는 의견을 제시하고 있다.

분석결과의 판정방법은 모수, 비모수 통계 모두 동일하고 사람이 분포 표를 이용하여 검정을 실시하는 경우 유의수준 α의 설정에 따라 의사결정에 다른 오류가 있을 수 있다는 사실을 염두에 두어야 하며, 컴퓨터 통계패키지를 이용할 경우 P-Value를 해석하는 객관적이고 경험적인 기준을 확보하여야 한다.

2.2 비모수 통계

2.2.1 McNemar모형 : 범주, 순위자료

McNemar 검정은 범주, 순위자료를 대상으로 2×2 분할 도수표(Contingency Freguency Tables)를 사용하는 방법이다. 검사자에게 검사반응의 첫번째 2개의 집합과 두번째 2개의 집합의 반응변화를 관찰하고자 할 경우 사용된다. 이 방법은 사전(Before) 관련 2행과 사후(After) 관련 2열로 이루어진 2×2 분할도수표로 데이터를 집계할 수 있으며 대각행렬 도수 f_{11}, f_{22}에 적합도검정을 적용한다.

대각행렬의 기대도수 $\dfrac{f_{11}+f_{22}}{2} > 5$일 경우 McNemar 검정을 사용하고 5보다 작으면 2.2.2절의 이항부호 검정을 적용한다. McNemar 검정통계량은 연속성수정(Correction of Continuity)이 포함되면,

$\chi_0^2 = \dfrac{(f_{11}+f_{22}-1)^2}{f_{11}+f_{22}}$ 이며 자유도와 유의수준에 의한 χ^2임계값 표[3] P-Value로 2.1절에서 언급한 바와 같이 Neyman-Pearson의 가설검정 또는 Fisher의 유의성검정방식을 선택하여 검정의 분석결과를 판단한다.

McNemar 검정을 3개 이상 $\times$ 3개 이상의 분할도수표로 확장할 경우 주변동질성 검정모형을 적용할 수 있다.[5]

2.2.2 부호 모형 : 순위자료

이항부호 검정은 연속분포를 갖는 순위자료를 대상으로 맞추어진 쌍(Matched or Mating Pair)의 차이에 대한 부호를 사용하는 방법이다.

동점(Tie)일 경우 즉 제로(Zero)를 제외하고 플러스와 마이너스의 부호를 갖는 차이를 보여 주는 쌍의 수(n)와 플러스와 마이너스 부호 중 더 적은 부호의 수(x)의 한쪽 P-Value 표[3]에 의하여 Fisher 방식의 검정에 대한 분석결과를 파악한다. 이 방법은 n<25인 소표본 비반복 다품종 소량생산에 적용될 수 있으며 n>25인 중표본 반복 다품종 소량생산에서는 $\mu = \dfrac{n}{2}$, $\sigma = \dfrac{\sqrt{n}}{2}$ 를 $Z_0 = \dfrac{x-\mu}{\sigma}$에 대입하여 정규근사 공식을 사용하며 연속성 수정이 포함되는 경우 $Z_0 = \dfrac{(x\pm0.5)-\mu}{\sigma}$ 를 적용하면 된다.

2.2.3 Wilcoxon 모형 : 순위자료

Wilcoxon 검정은 2.2.2절의 부호검정에서의 방향(Direction) 이외에 크기(Magnitude)까지도 고려하는 방법이다. 부호가 있는 차이에 부호에 관계없이 순위를 매기는 데 동점일 경우 평균순위를 사용한다. 부호붙은 순위들 중 더 작은 순위합계(T)를 계산하고 부호를 갖는 차이들의 개수(n)를 고려하여 한쪽, 양쪽 유의수준에 따른 임계값 표[3] P-Value로 Neyman-Pearson 또는 Fisher방식을 이용하여 검정의 분석결과를 판단한다.

이 방법은 n<25인 소표본 비반복 다품종 소량생산에 적용될 수 있으며 n>25인 중표본 반복 다품종 소량생산에서는 $\mu = \dfrac{n(n+1)}{4}$, $\sigma = \sqrt{\dfrac{n(n+1)(2n+1)}{24}}$ 를 $Z_0 = \dfrac{x-\mu}{\sigma}$ 에 대입하여 정규근사 공식을 사용한다.

Wilcoxon 검정 통계량에서 Walsh 평균을 이용하는 방법[6]은 2개 독립표본 모형에서만 적용되는 방법으로 2개 종속표본 모형에 적용할 수 없다.

2.2.4 Walsh 모형 : 구간, 비율자료

Walsh 검정은 대칭 모집단의 구간, 비율자료를 대상으로 부호붙은 차이점수(d_i)의 배열순서 즉 오름차순의 최대값과 최소값에 의한 방법이다.

쌍들의 수(n)와 한쪽, 양쪽 유의수준에 따른 임계값 표[3] P-Value로 Neyman-Pearson 또는 Fisher 방법을 이용하여 검정의 분석결과를 판단한다.

2.2.5 무작위 모형 : 구간, 비율자료

무작위 검정은 구간, 비율자료를 대상으로 쌍의 수 n<12일 경우 소표본 비반복 다품종 소량생산에 적용되는 방식이다. 부호붙은 차이점수를 d_i라 하고 쌍들의 수를 n이라고 할 경우 가능한 결과들의 수 2^n의 기각역에 있는 극단적인 결과들의 수를 유의수준 α에 2^n을 곱한 $\alpha \cdot 2^n$으로 결정한다. 관찰된 $\sum_{i=1}^{n} d_i$ 중 최대

값이 $\alpha \cdot 2^{n}$ 기각역에 있는 $\sum_{i=1}^{n} d_i$값보다 작을 경우 유의적인 판단을 하는데 관찰된 $\sum_{i=1}^{n} d_i$가 P-Value 역할을 한다.[3] 13<n<25일 경우 소표본 비반복 다품종 소량생산에 적용되는 방식으로 2.2.3의 Wilcoxon모형을 사용한다. n>25일 경우 중표본 반복 다품종 소량생산에 적용되며 $\mu=0$, $\sigma=\sqrt{\sum_{i=1}^{n} d_i^2}$을 $Z_0=\frac{\sum_{i=1}^{n} d_i-\mu}{\sigma}$에 대입하여 정규근사 공식을 사용한다.

3. 3개 이상 종속표본 모형

3.1 모수통계 모형

3개 이상 종속표본 모형의 모수통계 모형은 품질통계에서 실험계획법으로, 오차항이 정규성, 독립성, 불편성, 등분산성의 가정을 만족해야 하며 의학 및 보건통계에서는 블록계획법과 반복 측정자료의 분산분석법으로 이용된다. 따라서 2.1절과 같이 모수통계 모형은 모집단의 가정을 확인하기 위해 실험계획 전에 대표본이 요구되므로 소품종 대량생산의 품질통계에 적용한다.

3.1.1 난괴법 : 구간, 비율자료

난괴법(RBD : Randomized Block Design)은 구간, 비율자료의 특성값을 대상으로 2원 배치법 실험계획을 실시한다. 하나의 인자는 블록인자로 1원 배치법 CRD(Completely Randomized Design)의 반복(Replication)과 달리 블록내에서 반복측정(Repeated Measures)을 실시한다. 블록 수준 간 차이가 있어 다른 인자처리 수준의 값을 반복(Repetition)하거나 시간에 따른 효과를 관찰하고자 할 경우 적용된다.

예를 들어 검사자나 계측기의 종류가 서로 다르거나 검사자에 의한 재현 정밀도(Reproducibility)와 반복 정밀도 (Repeatability)의 특성값이 시간의 인자에 의해 영향을 받는가를 관찰할 경우 검사자나 계측기는 블록인자로 하는 난괴법을 실시한다. 난괴법 데이터의 구조식 $y_{ij}=\mu+a_i+b_j+e_{ij}$이며 e_{ij}는 오차항으로 모수통계 모형의 정규성, 독립성, 불편성, 등분산성 가정[2]을 만족해야 하며 b_j가 블록인자의 주효과(Main Effect)가 된다.

3.1.2 라틴 방격법 : 구간, 비율자료

2개의 블록인자에 라틴 방격을 배열한 실험계획법으로 검사자와 계측기 종류에 의한 재현 정밀도와 반복 정밀도의 시간요인에 따른 변화를 관찰할 경우 사용된다. 데이터 구조식 $y_{ijk}=\mu+a_i+b_j+c_l+e_{ijl}$로 a_i는 검사자의 주효과, b_j는 계측기의 주효과로 블록인자에 해당한다. 2개의 블록인자의 직교하는 라틴 방격을 2개, 3개 배열했을 경우 그레코라틴, 초그레코라틴 방격이라 하며 블록화와 직교화를 이용하여 각 인자의 교호작용을 무시하고 주효과만을 검출한다.

4×4 라틴, 그레코, 초그레코 라틴 방격법은 각각 CRD의 4원, 5원, 6원 배치법과 데이터 배열이 같으나 블록인자의 직교성을 이용하여 실험횟수를 줄이는 실험계획이다.

3.1.3 단일 분할법 : 1차 단위 1원 배치 : 구간, 비율자료

실험장치의 구조상 특정 인자 처리조건에서 몇 개의 샘플을 반복(Repetition)측정하거나 CRD의 완전 랜덤화 반복(Replication)시 특정 인자 처리조건을 변경하는 데 시간과 비용이 많이 소요되는 고온 열처리 실험에서 랜덤화의 곤란정도에 따라 블록내에서 반복측정(Repeated Measures)하는 분할법(Split-Plot Design)을 사용한다. 1차단위가 1원 배치인 단일 분할법의 데이터 구조식 $y_{ijk}=\mu+a_i+r_k+e_{(1)ik}+b_j+(ab)_{ij}+e_{(2)ijk}$로 r_k는 반복(Repetition), $e_{(1)ik}$, $e_{(2)ijk}$는 각각 1차, 2차 단위오차를 의미한다.

반복 측정자료의 분산분석법은 분할법같은 일변량(Univariate)방법이 있고 Mauchly의 구형성(Sphericity)이 만족되지 않을 경우에 자유도 손실을 줄이는 Greenhouse-Geisser, Huynh-Feldt방법을 사용하거나 Pillai의 Trace, Wilks의 Lamda, Hotelling의 Trace, Roy의 Maximum Root등의 다변량(Multivariate) 방법을 활용한다. [1]

3.1.4 단일 분할법 : 1차 단위 2원 배치 : 구간, 비율자료

단일 분할법의 1차 단위가 1원, 2원 배치일 경우, CRD의 반복있는 2원 배치법 데이터 배열과 같지만 실험의 랜덤화 순서와 데이터 구조식이 달라진다. 1차 단위가 2원 배치 단일 분할법은 두 인자처리 조합이 블록으로 분할되어 데이터 구조식,

$y_{ijk}=\mu+a_i+b_j+e_{(1)ij}+c_k+(ac)_{ik}+(bc)_{jk}+e_{(2)ijk}$가 되며 반복(Repetition)일 경우 $y_{ijk}=\mu+a_i+b_j+r_p+(ab)_{ij}+e_{(1)ijp}+c_k+(ac)_{ik}+(bc)_{jk}+(abc)_{ijk}+e_{(2)ijkp}$가 되어 각각 반복 없는 또는 반복 있는 3원 배치법의 데이터 배열은 같으나 블록 분할구(Split-Plot)로 인해 분산분석법에 차이가 있다.

3.1.5 다단계 분할법 : 구간, 비율자료

다단계 분할법은 지분 실험법(Nested Design)이라고 하며 모든 인자가 블록인자로 이루어져 내포되어 있는 실험계획법이다. 품질통계에서는 샘플링오차, 측정오차 등을 검토하기 위한 방법으로 사용되며 데이터의 구조식 $y_{ijkp}=\mu+a_i+b_{j(i)}+c_{k(ij)}+e_{p(ijk)}$로 CRD의 3원배치법과 데이터 배열은 같으나 내포된 블록인자로 인해 교호작용(Interaction)이 존재하지 않는다.

3.1.6 2단 분할법 : 구간, 비율자료

2단 분할법(Split-Split-Plot Design)은 랜덤화가 어려운 순서로 블록분할구로 인자를 선정하여 3차 단위까지 분할하는 실험계획이다. CRD의 반복있는 3원 배치법의 데이터 배열과 같으나 데이터구조식 $y_{ijkp}=\mu+a_i+r_p+e_{(1)ip}+b_j+(ab)_{ij}+e_{(2)ijp}+c_k+(ac)_{ik}+(bc)_{jk}+(abc)_{ijk}+e_{(3)ijkp}$ 가 되어 분산분석법이 다르다.

3.1.7 2방 분할법 : 범주자료

앞절의 난괴법, 라틴 방격법, 분할법등은 구간, 비율자료의 블록 반복측정된 방식이라면 2방 분할법은 이항범주자료의 블록실험계획법이다. 데이터 배열은 반복있는 2원 배치법과 같으나 데이터의 구조식 $y_{ijk}=\mu+a_i+b_j+e_{(1)ij}+e_{(2)ijk}$ 이며, 두 인자 처리 조건의 블록내에서 반복 측정되는 실험계획이다.

3.2 비모수통계 모형

3.2.1 Cochran모형 : 범주자료

Cochran검정은 범주자료 혹은 이원화된(Dichotomous) 순위자료를 대상으로 2.2.1절의 2×2 McNemar 모형을 $l\times m$ 분할도수표로 확장한 방식이다. 성공(Success)을 1, 실패(Failure)를 0으로 했을 경우, i번째 행의 성공의 총수는 R_i가 되고 j번째 열의 성공의 총수는 c_j이며 이 열들의 평균은 $\overline{c}$가 된다. 검정통계량은 이론식

$$\chi_0^2=Q_0=\frac{m(m-1)\sum_{j=1}^{m}(c_j-\overline{c})^2}{m\sum_{i=1}^{l}R_i-\sum_{i=1}^{l}R_i^2}$$ 과 계산식

$$\chi_0^2=Q_0=\frac{m(m-1)[m\sum_{j=1}^{m}c_j^2-(\sum_{j=1}^{m}c_j)^2]}{m\sum_{i=1}^{l}R_i-\sum_{i=1}^{l}R_i^2}$$ 이 되며

[3] 기각치는 $\chi^2(l-1:\alpha)$로 Neyman-Pearson과 Fisher 방식에 의해 검정의 분석결과를 판단한다.

3.2.2 Friedman 모형 : 순위자료

Friedman검정은 l개의 블록과 m개의 처리수준의 2원배치법 데이터 배열표에 의한 분포무관 방식이다. 각 블록행내의 순위를 구하고 이를 다시 각 처리열의 순위합 $R_{.j}$, 열평균 $\overline{R}_{.j}$, 총평균순위 $\overline{R}$를 구한다. 검정통계량의 이론식과 계산식은,

$$S_0=\frac{12l}{m(m+1)}\sum_{j=1}^{m}(\overline{R}_{.j}-\overline{\overline{R}})=\frac{12}{lm(m+1)}\sum_{j=1}^{m}R_{.j}^2-3l(m+1)$$ 이 되며 기각치는 $s(l,m;\alpha)$[4]로 Neyman-Pearson과 Fisher 방식에 의해 검정의 분석결과를 판단한다.

이 방법은 $l<13$인 경우 소표본 비반복 다품종 소량생산에 사용되는 방법으로 $l>13$인 경우는 중표본 반복 다품종 소량생산에 적용되며 기각치 $s(l,m;\alpha)$ 대신 $\chi^2(m-1;\alpha)$를 사용한다. 동점인 경우 평균순위를 사용하나 블록 i에서 k번째 동점 그룹의 크기를 t_k라 할 경우 $T_i=(\sum^k t_k^3-\sum^k t_k)$이며 검정통계량

$$S_0'=\frac{S_0}{1-\sum_{i=1}^{l}T_i/[lm(m^2-1)]}$$ 을 이용한다.

3.2.3 Page 모형 : 순위자료

Page 검정은 각 l개의 블록 행들과 m개 열의 처리조건이 순위를 갖는 경우 적용되며, 처리효과에 대한 순서 대립가설을 검정하기 위한 분포무관 방식이다.

Page 통계량 $L_0'=\sum_{j=1}^{m}jR_{.j}$로 j는 처리수준의 순서에 의한 가중치로 간주될 수 있으며 기각치 $L(l,m;\alpha)$[4]로 Neyman-Pearson과 Fisher 방식에 의해 검정의 분석결과를 판단한다. 이 방법은 $l<20$인 경우 소표본 비반복 다품종 소량생산에 사용되는 방법이며 $l>20$인 경우는 중표본 반복 다품종 소량생산에 적용되며

$$\mu=\frac{lm(m+1)^2}{4},\quad \sigma=\sqrt{\frac{l(m^3-m)^2}{144(m-1)}}$$ 를

$$Z_0=\frac{L_0-\mu}{\sigma}$$ 에 대입하여 정규근사 공식을 사용한다.

3.2.4 Doksum 모형 : 순위자료

Doksum 검정은 2.2.3절의 Wilcoxon 부호순위 개념을 확장한 것으로 3.2.2절의 Friedman 검정이 블록내의 순위만을 고려했으나 Doksum 검정은 블록간, 블록내의 순위를 모두 고려하는 점근 분포무관 방식이다. Doksum 검정통계량,

$$\chi_0^2 = A_0 = \frac{\sum_{j=1}^{m} \overline{Hj.} - [(m-1)/2m]^2}{(m-1)V/2m}$$ 이며

$V = \frac{2l-1+(m-2)[7(l-2)+13-6n]}{3ml(l-1)}$ 이고 $\overline{Hj.}$ 은 블록간, 블록내 순위가 조정된 값이다. 기각역은 $\chi^2(m-1:\alpha)$로 Neyman-Pearson과 Fisher방식에 위해 검정의 분석결과를 판단한다.

Doksum 검정은 점근 분포무관 방법으로 대표본 소품종 대량생산에 적용된다.

3.2.5 Hollander 모형 : 순위자료

Hollander 검정은 3.2.4절의 Doksum 검정과 다르게 3.2.3절과 같이 m개 열의 처리조건이 순위를 갖는 경우 블록내 순위뿐 아니고 블록간의 정보도 고려하는 점근 분포무관 방식이다. Hollander 검정통계량은 T는 블록간, 블록 내 조정된 순위합, $\mu = \frac{m(m-1)l(l+1)}{8}$,

$a = \sqrt{\frac{1}{144} l(l+1)(2l+1)m(m-1)3+2(m-2)^2\hat{\rho}}$

를 $Z_0 = \frac{T-\mu}{\sigma}$ 에 대입하면 정규근사 공식을 사용할 수 있다. 여기서 $\hat{\rho} = 12\lambda - 3$이고 $\frac{1}{4} \le \lambda \le \frac{4}{27}$ 이며, λ는 상한인 $\frac{4}{27}$ 를 사용한다. Hollander 검정은 점근 분포무관 방법으로 대표본 소품종 대량생산에 적용될 수 있다.

3.2.6 다중 비교모형 : 순위자료

3.2.2절의 Friedman 순위합에 기초한 전체 처리비교 공식이 $|R_{.s} - R_{.t}| \ge r(l, m:\alpha)[4]$이면 두 처리수준 s와 t는 차이가 있다는 판단을 내린다. 이 방법은 $l < 15$인 소표본 비반복 다품종 소량생산에 적용될 수 있으며 $l > 15$인 경우 중표본 반복 다품종 소량생산에 사용되며 전체비교 공식은 $|R_{.s} - R_{.t}| \ge q(l=\infty, m:\alpha)\sqrt{\frac{lm(m+1)}{12}}$ 이 된다. 처리, 대조비교 공식은 $(R_{.s} - R_{.1}) \ge r^*(l, m-1:\alpha)$[4] 이며 s수준이 1수준보다 크다는 판단을 내린다. 이 방법은 $l < 18$인 소표본 비반복 다품종 소량생산에 사용되는 방법이며 $l > 18$인 중표본 반복 다품종 소량생산인 경우에는 $(R_{.s} - R_{.1}) \ge M(\frac{1}{2}, m-1:\alpha)\sqrt{\frac{lm(m+1)}{6}}$ 를 이용한다.

4. 결 론

품질통계의 종속형 모형은 시계열 종속모형, 다변량 종속모형, 종속표본 모형 등의 세가지 모형으로 분류된다. 본 연구에서는 종속표본 모형을 대상으로 생산형태에 따른 모수통계 모형과 비모수통계 모형을 유형화하여 명목, 순위, 구간, 비율자료의 적용방안을 제시하였다.

1. 2개 종속표본 모수통계 모형에서 구간, 비율자료에 대한 t_0차이 검정은 대표본 소품종 대량생산에 적용된다.

2. 2개 종속표본 비모수통계 모형에서 McNemar 모형은 범주, 순위자료에 적용되며 부호모형, Wilcoxn 모형은 소, 중표본 반복, 비반복 다품종 소량생산의 순위자료에 적용된다. Walcoxon모형은 소, 중표본 반복, 비반복 다품종 소량생산의 순위자료에 적용된다. Walsh 모형은 구간, 비율자료에 적용되며 무작위 모형은 소, 중표본 반복, 비반복 다품종 소량생산의 구간비율 자료에 적용된다.

3. 3개 종속표본 모수통계 모형의 구간, 비율자료에 대해서는 블록 반복측정된 난괴법, 라틴 방격법, 단일 분할법, 다단계 분할법, 2단 분할법과 범주자료에 대한 이방 분할법의 모수통계 모형은 대표본 소품종 대량생산에 적용된다.

4. 3개 종속표본 비모수통계 모형에서 Cochran 모형은 범주자료에 적용하고 Friedman 모형, Page 모형, 다중 비교모형은 소, 중표본 반복, 비반복 다품종 소량생산의 순위자료에 적용되는 분포무관 방법이며, Doksum 모형, Hollander 모형은 대표본 소품종 대량생산의 순위자료에 적용하는 점근 분포무관 방식이다.

5. 참 고 문 헌

[1] 박동권, SPSS시리즈 : 분산분석과 반복측정자료, 민영사, 2002.
[2] 박성현, 현대실험계획법, 민영사, 2005.
[3] 소영일 외, SPSS를 활용한 비모수통계학, 법문사, 1987.
[4] 송문섭 외, 비모수통계학개론, 자유아카데미, 1989.
[5] 안윤기, SPSS시리즈 : 가설검정, 민영사, 2004.
[6] 이승훈, Minitab을 이용한 공학통계 자료분석, 이레테크, 2006.
[7] 최종후 외, 학술논문과 통계적기법, 자유아카데미, 1990.

저 자 소 개

최 성 운

현 경원대학교 산업공학과 교수. 한양 대학교 산업공학과에서 공학사, 공학석사, 공학박사 학위를 취득하고, 1994년 한국과학재단 지원으로 University of Minnesota에서 1년간 Post-Doc을 수행했으며, 2002년부터 1년 반동안 University of Washington에서 Visiting Professor를 역임하였음. 주요 관심분야는 자동화 생산 및 장치 산업에서의 품질관리이며, 컴퓨터 · 정보통신시스템의 신뢰성 설계 및 분석, 서비스 사이언스, RFID시스템에도 관심을 가지고 있음.

주소: 경기도 성남시 수정구 복정동 산65번지 경원대학교 산업공학과 ☏ 031) 750-5366