信賴性應用硏究
제3권, 제1호, pp. 59-72, 2003

다구찌 누적법의 특징과 활용

이우선, 송일성

성신여자대학교 통계학과

Discussion of the Taguchi's Accumulation Analysis

Woo Sun Lee, Il seong Song

Department of Statistics, Sungshin Women's Univ.

Abstract

The Accumulation Analysis is a main method of Taguchi's data analysis strategy. However, Many statistician(specially Hamada, Wu and Nair etc.,) pointed that this method is too much complicated while not being useful. In this article, We demon -strate the method of this Accumulation Analysis by examples and introduce the argument points and the alternative solutions discussed among many statisticians.

1. 서론

순서(順序)지어진 범주형 분할표를 분석하는 경우 피어슨의 χ^2-검정법의 문제점을 해결하기 위하여 다구찌(田口玄一, G. Taguchi)가 새롭게 제안하여 발전시켜 온 것이 누적법(累積法, Accumulation Analysis)이다. 다구찌는 그의 實驗計劃法 新版(1962)의 제 15 장과 16 장에서 순서형 분류치에 관한 분할표의 분석법으로서 누적법을 소개 하고 있다. 그러나 이러한 방법이 널리 알려지게 된 것은 다구찌의 논문(1974,1987 Saishin Igaku)에 자세히 소개된 이후이다. 다음은 다구찌의 1974년 논문에 소개된 내용이다.

<예1.1, 1974>두 가지의 의약품 A와 B를 사용해서 각각 80명씩에 표준 투약을 하고 다른 80명씩에 대해서는 어떤 처치 투약을 하여 다음과 같은 가상적인 관측 도수를 얻었다. 여기에서는 실험 결과를 4개의 크기 순위의 계급 범주로 나누어서 관찰하여 기록한 것이다. 순서 지어진 계급 범주라는 것은, '전연 효과가 없다(-)', '약간 효과가 있다(+)', '효과가 있다(++)', '현저하게 효과가 있다(+++)'의 4가지이다.

<표 1.1> 약품 A, B에 대한 효과 실험

약품 A

	-	+	++	+++	합계
표준	40	24	10	6	80
처치	24	40	10	6	80
차이	-16	16	0	0	
c_i	64/160	64/160	20/160	12/160	

* 차이=처치-표준, c_i=(표준+처치)/160

χ^2=8.00, α=0.047

약품 B

	-	+	++	+++	합계
표준	40	24	10	6	80
처치	24	29	16	11	80
차이	-16	5	6	5	
c_i	64/160	53/160	26/160	17/160	

χ^2=7.33, α=0.06

이와 같은 경우, 피어슨의 χ^2-검정에서의 귀무가설은 각 효과 범주 내에서 표준과 처치 사이에서 차이가 없다는 것으로, 이 예에서는 다같이 80명을 대상으로 하고 있으므로 모든 범주 내에서의 차이는 0이라는 것이다. 그런데 약품 A에서는 이 차이가 들쑥날쑥 하거나 전혀 차이가 없는 것으로 보이나, 약품 B의 경우는 두 번째 계급 범주부터는 분명히 처치 효과가 있는 것으로 나타나 있다. 이와 같이 가설과 어긋나 있는 것이 분명한데도 불구하고 유의 수준 5%하에서 가설이 기각되지 않고 도리어 약품 A의 경우에서 가설이 기각되는 불합리한 결과를 확인 할 수 있다. 이와 같이 분할 표를 관찰해서 얻는 당연한 상식과 검정결과가 어긋나는 이 사실을 다구찌는 다음과 같이 논술하고 있다. "피어슨의 χ^2-검정은 이와 같은 순서 지어진 범주 형 분할표의 경우에 있어서 효과의 크기를 검정하는데 적당하지 못하다. 위의 예에서 약품 A보다는 약품 B의 효과가 크다는 것이 분명하다. 그런데도 피어슨

의 χ^2-검정결과는 엉뚱한 결과를 가져올 수 있는 문제점을 가지고 있다. 이를 해결하기 위하여 이른바 누적법을 사용하면 이러한 문제점을 개선 할 수 있다."

2. 다구찌의 누적법

<예1.1, 1974>에서 약품 A, B에 관한 분할표인 <표 1.1>에서 다음과 같은 방법으로 도수를 누적시킨 <표 2.1>을 작성한다. 우선 순서 지어진 4개 계급 범주에 대해서 (-)열은 계급 제 I조로 열의 명칭만 바꾸고, (-)열과 (+)열의 대응 도수를 합한 것을 계급 제 II조로 하고, 이 제 II조와 (++)열의 서로 대응하는 도수를 합한 것으로 계급 제 III조를 구성한다. 마찬가지로 계급 제 IV조를 구성할 수 있으나 이것은 처리별 도수 합계로 되어 분석 대상에서는 제외된다.

<표 2.1> <표 1.1>에서 작성한 누적표

약품 A					약품 B				
A	I	II	III	IV	B	I	II	III	IV
표준	40	64	74	80	표준	40	64	74	80
처리	24	64	74	80	처리	24	53	69	80
합계	64	128	148	160	합계	64	117	143	160
비율	d_1	d_2	d_3	1	비율	d_1	d_2	d_3	1

이 표의 각 칸의 수치는 기본 관측치 0과 1로 이루어진 것들의 합계라는 것으로 이해하여야 한다. 예를 들어 A에서 계급 I조에서 표준 처리의 40이라는 것은 표준 처리를 받은 80명 중 제 I조에 속하면 1 그렇지 않으면 0으로 한 관측치 80개의 합계가 40이라는 것이다. 이와 같이 생각하여 각 조 별로 분산분석을 하고 이것을 각 요인별로 합한다.

예를 들어 i번 계급조에 대해서 전체 제곱합 $SS_T(i)$, 처리 제곱합 $SS_t(i)$, 그리고 오차 제곱합 $SS_E(i)$를 계산한다. 그런데 각 계급조 간에 분산이 같지 않다는 것이 문제이다. 이것을 조정하기 위하여 2항분포에서의 분산 공식 $d_i(1-d_i)$를 이용한다. 이와 같은 분산으로 i조에서의 제곱 합을 나누어 표준화를 하는 것이다. 여기에서 i조에 대해서 $W_i = 1/[d_i(1-d_1)]$와 같이 정의한 가중치(weight)를 i 조의 각 제곱 합에 곱하는 것에 의해서 표준화된다. 이러한 것을 구체적으로 설명해 본다. 다음은 위의 예에 대해서 필산으로 분석을 하는 경우 그 과정을 설명하고 있는 것이다.

위의 예에서 약품 B의 경우, χ^2의 검정에서 5% 수준 아래 이에 미달되지만 누적법으로

분석하는 경우, F값이 어떻게 되어 유의성이 있는 것으로 판단되는가를 살펴보자.
약품 B의 누적표에서,

$d_1 = 64/160$, $d_2 = 117/160$, $d_3 = 143/160$이므로 $W_1 = 160^2/[64(160-64)] = 4.1667$, $W_2 = 160^2/[117(160-117)] = 5.0885$, $W_3 = 160^2/[143(160-143)] = 10.5306$, 그리고

$CF(1) = 64^2/160 \cdot W_1$, $CF(2) = 117^2/160 \cdot W_2$, $CF(3) = 143^2/160 \cdot W_3$, 와 같이 보정항을 구한다.

$SS_T(1) = (1 \times 64 + 0 \times 96)W_1 - CF(1) = (64 - 64^2/160)W_1 = 160$

을 얻는다. 여기에서 주목할 것은 이 값은 총 측정 도수인 160, 즉 하나의 상수(constant)라는 것이다. 또 이것은 계급 조에 따라 달라지지도 않는다. 그러므로 전체 제곱 합은 $SS_T = 160 \times 3 = 480$이다.

각 조의 처리 제곱 합은 2 수준간의 것이므로 다음과 같이 계산된다.

$SS_t(1) = (40-24)^2/160 \cdot W_1 = 6.6667$, 자유도=1

$SS_t(2) = (64-53)^2/160 \cdot W_2 = 3.8482$, 자유도=1

$SS_t(3) = (74-69)^2/160 \cdot W_3 = 1.6454$, 자유도=1,

따라서, $SS_t = 6.6667 + 3.8482 + 1.6454 = 12.1603$, 자유도=3이다.

여기에서 이 값을 약품 B의 통계량이라는 뜻으로 T_B로 표시하자. 오차 제곱 합은 $SS_E = 480 - 12.1603 = 467.8397$와 같이 구해진다. 이상의 것을 정리한 것이 다음 분산분석표이다.

<표 2.2> 약품 B에 관한 누적표에 대한 분산분석표

변인	자유도(df)	제곱 합(AS)	평균 제곱(MS)	F
처리	3	T_B=12.16	4.05	4.11
오차	474	467.84	0.99	
계	477	480		

여기에서 F-검정을 하면 5% 수준하에서 뿐만 아니라, 1% 수준하에서도 유의성이 나타나 우리의 상식과 잘 부합된다는 것을 알 수 있다. 약품 A의 경우는 $T_A = 6.6667$로 되어 χ^2-검정 때와는 달리 5% 유의 수준 아래 유의성이 나타나지 않는다.

이러한 방법은 다음과 같은 다요인(Multifactor) 실험의 경우에도 확대 응용된다.

<예1.2, 1974> 어떤 주조(molding)공정에서 최적 주조 조건을 구하기 위하여 다음과 같은 실험 계획 아래 실험을 실시하였다. 여기에서 고려된 인자와 수준은 다음과 같다.

A=자료(資料), 2 수준 (현재의 것과 새로운 것)

B=주조(鑄造)의 형식, 2 수준 (현재의 형식과 새로운 형식)

C=경화(硬化) 방법, 2 수준 (현재의 방법과 새로운 방법)

D=용해(鎔解) 온도, 2 수준 (현재의 온도와 현재 온도+30℃)

관측 도수는 각 인자 조합(실험 조건)별 실험 단위에서 20개씩의 제품을 조사하여 수(no flow-holes), 우(a few flow-holes), 양(several flow-holes), 그리고 가(many flow-holes)로 분류하였다. 이것을 정리한 것이 <표 2.4>이다.

<표 2.3> 주조 실험의 계획과 데이터

실험 번호	A ①	B ②	A B ③	C ④	E R ⑤	D ⑥	A D ⑦	관측 도수 수	우	양	가	누적 도수 I	II	III	IV
1	1	1	1	1	1	1	1	15	3	1	1	15	18	19	20
2	1	1	1	2	2	2	2	14	3	2	1	14	17	19	20
3	1	2	2	1	1	2	2	16	2	2	0	16	18	20	20
4	1	2	2	2	2	1	1	19	1	0	0	19	20	20	20
5	2	1	2	1	2	1	2	17	1	0	2	17	18	18	20
6	2	1	2	2	1	2	1	17	2	1	0	17	19	20	20
7	2	2	1	1	2	2	1	10	2	4	4	10	12	16	20
8	2	2	1	2	1	1	2	16	3	1	0	16	19	20	20
							합계	124	17	11	8	124	141	152	160

이것을 분석하기 위하여 각 인자별로 관측 도수를 정리하여야 한다. 예를 들면,

인자 A의 경우:

<표 2.4> 분할표 및 누적표

인자 A의 분할표	수	우	양	가	계	누적표 I	II	III
A1	64	9	5	2	80	64	73	78
A2	60	8	6	6	80	60	68	74

이것은 2행 4열의 분할표이다. 따라서 위의 약품 B의 경우와 마찬가지로 분석할 수 있다.

W_1=5.7348　　W_2=9.5558　　W_3=21.0526

등을 계산하여 다음과 같이 AA-통계량이 구해진다.

T=4.1718

이와 같은 분석을 각 인자에 대하여 실시하여 정리한 것이 다음 <표 2.5> 이다. 여기에서 TA, TB등으로 표시한 것은 인자 A의 통계량 T, 인자 B의 통계량 T등을 나타내는 것이다.

<표 2.5> 인자별 데이터의 정리와 AA-통계량*

요인	수준	수	우	양	가	I	II	III	AA-통계량
A_1	1	64	9	5	2	64	73	78	
A_2	2	60	8	6	6	60	68	74	TA=4.1718
B_1	1	63	9	4	4	63	72	76	
B_2	2	61	8	7	4	61	69	76	TB=0.6809
AB_1	1	55	11	8	6	55	66	74	
AB_2	2	69	6	3	2	69	75	78	TAB=13.9680
C_1	1	58	8	7	7	58	66	73	
C_2	2	66	9	4	1	66	65	69	TC=11.8684
D_1	1	67	8	2	3	67	75	77	
D_2	2	57	9	9	5	57	66	75	TD=8.9482
AD_1	1	61	8	6	5	61	69	75	
AD_2	2	63	9	5	3	63	72	77	TAD=1.2072
ER_1	1	64	10	5	1	64	74	79	
ER_2	2	60	7	6	7	60	67	73	TER=8.2368

* 처리제곱합 SS_t를 Accumulation Analysis Statistic라 하여 AA-통계량이라고 표시한 것이다.

<표 2.6> <표 2.5>에서의 AA-통계량을 분산분석표로 종합한 것

변인	자유도	제곱합	평균제곱	F
A	3	4.1718	1.3906	1.47
B	3	0.6809	0.2270	0.24
$A \times B$	3	13.9680	4.6560	4.93**
C	3	11.8684	3.9561	4.19**
D	3	8.9482	2.9827	3.16**
$A \times D$	3	1.2072	0.4024	0.43
ER	3	8.2368	2.7456	2.91*
$Error$	456	430.9188	0.9450	
T	477	480		

이 예에서 처리 조합(인자 수준 조합) $A_iB_jC_kD_l$의 실현치를 Y_{ijkl}로 표시하고 이 변수의 기대치를 $\mu(A_iB_jC_kD_l)$로 표시하기로 한다. 그러면 위 분산분석표에서 다음과 같은 모형을 생각할 수 있다. 이 모형에서 ab_{ij}는 a_i와 b_j간의 교호작용을 나타낸다.

$$E[Y_{ijkl}] = \mu(A_iB_jC_kD_l) = m + a_i + b_j + ab_{ij} + c_k + d_l \quad (2.1)$$

$$Var(Y_{ijkl}) = \sigma^2 \quad (2.2)$$

처리 조합의 추정치는 다음과 같다.

$$\begin{aligned}\widehat{\mu}(A_iB_jC_kD_l) &= \widehat{m} + \widehat{a}_i + \widehat{b}_j + \widehat{ab}_{ij} + \widehat{c}_k + \widehat{d}_l \\ &= \overline{M} + (\overline{A}_i - \overline{M}) + (\overline{B}_j - \overline{M}) + (\overline{A_iB_j} - \overline{A}_i - \overline{B}_j + \overline{M}) \\ &\quad + (\overline{C}_k - \overline{M}) + (\overline{D}_l - \overline{M}) \\ &= \overline{A_iB_j} + \overline{C}_k + \overline{D}_l - 2\overline{M}\end{aligned} \quad (2.3)$$

여기에서 $\overline{M}$, $\overline{A}_i$ 등은 전체 평균치, 인자 A의 i 수준 평균치 등을 나타낸다.

<표 2.7> <표 2.4>에서 A_iB_j 수준 조합의 도수와 백분율(%),
그리고 C_k, D_l의 수준별 도수와 백분율(%)

요인 수준	도수와 백분율(%) 수	우	양	가	합계	누적 백분율(%) I	II	III
A_1B_1	29	6	3	2	40			
	(72.5)	(15.0)	(7.5)	(5.0)		72.5	87.5	95.0
A_1B_2	35	3	2	0	40			
	(87.5)	(7.5)	(5.0)	(0.0)		**87.5**	**95.0**	**100.0**
A_2B_1	34	3	1	2	40			
	(85.0)	(7.5)	(2.5)	(5.0)		85.0	92.5	95.0
A_2B_2	26	5	5	4	40			
	(65.0)	(12.5)	(12.5)	(19.0)		65.0	77.5	90.0
C_1	58	8	7	7	80			
	(72.5)	(10.0)	(8.75)	(8.75)		72.5	82.5	91.25
C_2	66	9	4	1	80			
	(82.5)	(11.25)	(5.0)	(1.25)		**82.5**	**93.75**	**98.75**
D_1	67	8	2	3	80			
	(83.75)	(10.0)	(2.5)	(2.5)		**83.75**	**93.75**	**96.25**
D_2	57	9	9	5	80			
	(71.25)	(11.25)	(11.25)	(6.25)		71.25	82.5	93.75

여기에서 최적 인자 조합과 인자 수준은 A_1B_2와 C_2, D_1 이라는 것을 알 수 있다. 따라서 백분율을 비율로 표시하여 추정되는 누적 비율은 (2.3)에서 다음과 같다.

$$\widehat{\mu_I}(A_1B_2C_2D_1) = 0.875 + 0.825 + 0.8375 - 2(124/160) = 0.9875 \quad (2.4)$$

$$\widehat{\mu_{II}}(A_1B_2C_2D_1) = 0.950 + 0.9375 + 0.9375 - 2(141/160) = 1.0625 \quad (2.5)$$

그런데 비율의 추정치는 [0,1]의 범위 내에 있어야 한다. 그러나 이러한 범위를 벗어나는 일은 비율을 가감해서 얻는 수에서는 생길 수 있는 일이다. 이러한 문제점을 극복하고 비율의 추정치가 [0,1]의 구간 내에 있도록 하는 방법으로 오메가(Ω)법이라는 것을 사용하고 있다. 이것은 로그 변환의 일종인 로짓 변환과 같은 것이다. 다구찌는 다음과 같이 비율 p에 대한 변환을 사용하고 있다. 이것을 p의 데시벨(dB)이라고 한다.

$$dB = -10\log_{10}\left(\frac{1}{p} - 1\right) \tag{2.6}$$

이와 같은 데시벨(decibel)을 이용할 수 있는 수치 표가 마련되어 있다. 이 표에 의하면

$$\begin{aligned} dB(\widehat{\mu_I}) &= dB(0.875) + dB(0.825) + dB(0.8375) - 2dB(0.775) \\ &= 8.451dB + 6.734dB + 7.137dB - 2\times 5.371dB \\ &= 11.580dB \end{aligned} \tag{2.7}$$

데시벨 표를 역으로 이용하여 $\widehat{\mu_I} = 0.935$를 얻는다. 마찬가지로 하여 $\widehat{\mu_{II}} = 0.987$ 이다.

한편, 현재의 조건하에서는 $\widehat{\mu_I} = 0.751$ $\widehat{\mu_{II}} = 0.977$과 같이 된다.

<예 2.3> 다음 예는 Taguchi(1987, 189쪽)에 수록되어 있고, Taguchi and Wu(1990)에서 아크-용접(Arc-Welding) 실험으로 소개되고 비판의 대상으로 논의되고 있는 것이다.

두 개의 강철 판 사이의 아크-용접을 위한 최적 조건을 알아내기 위한 실험이다. 여기에 9개의 제어 인자가 취해졌다. 이것을 A, B, C, D, E, F, G, H, 그리고 I 로 표시한다. 그리고, 교호작용 $A\times G$, $A\times H$, $G\times H$, $A\times C$를 알아보기로 하였다.

실험 계획은 $L_{16}(2^{15})$직교배열표를 이용하였고, 이들 각 조건에서의 반응 관측치로는 반응 관측치 겉모양(appearance), X선 검사 결과, 용접 부분의 공학적 강도, 그리고 작업성(workability)이 측정 기록되었다. 그러나, 여기에는 작업성에 관한 3개 범주, 즉 쉽다(easy), 보통이다(normal), 어렵다(difficult)의 어느 범주에 속하는가를 판단하여 기록한 것을 분석한다. 작업성외의 다른 반응 관측치에 대한 것은 그 종류와 관측치만을 소개하고, 다구찌에 의해서 분석하여 얻은 종합적인 결론을 소개한다.

<표 2.8> $L_{16}(2^{15})$를 이용한 실험계획과 작업성에 관한 데이터

	직교배열표															관찰치 작업성					
	A	G	A G	H	A H	G H	B	D	E	F	I	e	e	A C	C						
열번호 / 실험번호	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	E	N	D	I	II	III
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	0	1	0	0	1	1
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	0	1	0	0	1	1
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	0	0	1	1	1
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	0	0	1	1	1
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1	0	0	1	0	0	1
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1	1	0	0	1	1	1
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2	0	0	1	0	0	1
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	0	1	0	0	1	1
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1	0	1	0	0	1	1
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1	0	1	0	0	1	1
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2	0	1	0	0	1	1
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	0	0	1	1	1
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2	0	0	1	0	0	1
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2	0	1	0	0	1	1
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1	0	1	0	0	1	1
	a	b	a b	c	a c	b c	a b c	d	a d	b d	a b d	c d	a c d	b c d	a b c d						

각 실험 조건(인자 수준 조합)에 대한 작업성이 3개 중 하나로 판정되었다는 특징이 있다. 따라서 데이터가 극단적으로 제한적이고 또 이것은 특징적으로 결핍 자료(the paucity of data)라는 것을 알 수 있다. 그러나 다구찌는 여기에서도 다른 경우와 마찬가지로 누적법을 적용하여 분석하고 있다. <예 2.2>의 경우와 마찬가지로 인자별로 정리한 2행 3열의 분할표를 작성하고, 이것에서 계산한 AA통계량 T, 그리고 χ^2값을 표시한 것이 다음 표이다.

<표 2.9> 인자별 데이터의 정리와 AA 통계량 T, 그리고 χ^2

인자	E	N	D	I	II	AA-T	χ^2
A1	3	3	2	2	6		
A2	1	6	1	1	7	1.7436	2.3333
B1	2	4	2	2	6		
B2	2	5	1	2	7	0.4103	0.4444
C1	3	4	1	3	7		
C2	1	5	2	1	6	1.7436	1.4444
D1	3	5	0	3	8		
D2	1	4	3	1	5	5.0256	4.1111
E1	2	5	1	2	7		
E2	2	4	2	2	6	0.4103	0.4444
F1	0	5	3	0	5		
F2	4	4	0	4	8	9.0256	7.1111
G1	1	7	0	1	8		
G2	3	2	3	3	5	5.0256	6.7778
H1	2	4	2	2	6		
H2	2	5	1	2	7	0.4103	0.4444
I1	1	5	2	1	7		
I2	3	4	1	3	7	1.7436	1.4444
AC1	2	4	2	2	6		
AC2	2	5	1	2	7	0.4103	0.4444
AG1	2	5	1	2	7		
AG2	2	4	2	2	6	0.4103	0.4444
AH1	1	6	1	1	7		
AH2	3	3	2	3	6	1.7436	2.3333
GH1	1	6	1	1	7		
GH2	3	3	2	3	6	1.7436	2.3333

다음은 이것을 분산분석표로 정리한 것이다.

<표 2.10> 작업성에 관한 분산분석표

변인	자유도	SS	V=MS	F
A	2	1.74	0.87	2.90
B	2	0.41		
C	2	1.74	0.87	2.90
D	2	5.03	2.52	8.40
E	2	0.41		
F	2	9.03	4.52	15.07
G	2	5.03	2.52	8.40
H	2	0.41		
I	2	1.74	0.87	2.90
AG	2	0.41		
AH	2	1.74	0.87	2.90
AC	2	0.41		
GH	2	1.74	0.87	2.90
RESIDUAL	4	2.16	0.54	
Total	30	32.00		

작업성의 분석에서는 인자 D, F, G가 중요인자로 간주된다.

참고로 다른 반응 관측치인 겉모양(Appearance), X선 검사 결과, 용접 부분의 공학적 강도에 관한 관측 자료는 다음과 같다.

<표 2.11> 겉모양, X-선 검사 등에 관한 관측치

실험 번호	겉모양 앞면 상	겉모양 앞면 중	겉모양 앞면 하	겉모양 뒤면 상	겉모양 뒤면 중	겉모양 뒤면 하	X-선 상	X-선 중	X-선 하	장력 강도 kg/mm^2	신장력 %
1	1	0	0	1	0	0	3	1	0	43.7	33.6
2	0	1	0	0	0	1	3	1	0	40.2	40.2
3	0	1	0	0	1	0	4	0	0	42.4	30.5
4	1	0	0	0	1	0	2	2	0	44.7	23.7
5	0	1	0	0	0	1	2	2	0	42.4	34.7
6	0	0	1	0	0	1	4	0	0	45.9	21.8
7	1	0	0	0	1	0	2	2	0	42.2	24.8
8	0	1	0	0	0	1	3	1	0	40.6	29.8
9	0	1	0	1	0	0	2	1	1	42.4	33.7
10	0	1	0	0	1	0	3	1	0	45.5	25.5
11	1	0	0	0	1	0	3	1	0	43.6	36.9
12	0	1	0	1	0	0	4	0	0	40.6	29.0
13	1	0	0	0	0	1	0	3	1	44.0	30.3
14	0	0	1	0	0	1	3	0	1	40.2	39.0
15	1	0	0	0	1	0	1	3	0	42.5	27.9
16	1	0	0	1	0	0	4	0	0	46.5	40.8

다음은 작업성을 포함해서 모든 관측치에 대한 분석결과를 종합한 것이다.

<표 2.12> 요인효과에 대한 종합정리표

		작업성(%)			표정(%)			X-선			장력강도	伸張度	판정
인자	수준	E	N	D	T	M	B	T	M	B			
A	1										42.76	29.9	
A	2										43.16	32.9	*
B	1										41.89	33.0	
B	2										44.04	29.8	*
C	1				50	31	19				44.51	29.7	*
C	2				19	50	31				41.41	33.1	
AC	11										44.42	26.0	
AC	12										41.40	33.8	
AC	21										44.90	33.4	(*)
AC	22										41.42	32.4	
D	1	38	62	0	44	44	12	53	41	6			표시인자
D	2	12	50	38	25	38	37	84	16	3			
E	1												*
E	2												*
F	1	0	62	38				84	13	6	43.16	33.3	
F	2	50	50	0	38	56	6	53	44	3	42.76	29.5	*
G	1	12	88	0	34	26	43						*
G	2	38	25	37									
AG	11											32.0	
AG	12											27.8	
AG	21											31.3	(*)
AG	22											34.5	
H	1				25	32	43						
H	2				44	50	6						*
AH	11										43.05	32.6	
AH	12										42.48	27.2	
AH	21										43.02	32.1	
AH	22										43.30	33.6	(*)
GH	11				38	50	12						
GH	12				38	62	0						(*)
GH	21				12	13	75						
GH	22				50	38	12						
I	1										43.15	28.6	
I	2										42.77	34.2	*

최적조건(최적 인자 수준 조합)은 $A_2B_2C_1D_1E_1F_2G_1H_2I_2$($E_1$대신에 E_2도 무방)이다.

3. 결론 및 토의

앞 절에서 소개한 다구찌의 누적법을 살펴보면 전통적인 통계적 방법의 관점에서 납득하기 어려운 점들을 내포하고 있다. Nair(1986b)는 다구찌의 누적법을 본질적으로 따지고 대

안을 제시하는 논문을 쓰게 된 동기를 다음과 같이 말하고 있다. "미국에서 다구찌 방법이 널리 받아들여져 사용되고 있는 것이 사실이다. 따라서 이 방법의 장단점을 따지고, 이것이 부적절하다면 이를 대신 할 수 있는 단순하고도 효율적인 방법을 제시할 필요가 있다"는 것이다. Nair(1986a,b) 그리고 Hamada and Wu(1990)등에 나타난 다구찌의 누적법에 의한 문제점으로 지적되고 있는 내용을 다음과 같이 요약하여 정리할 수 있다.

(1) 분산분석표에서 처리 제곱합 SS_t만이 빈도 관측치 에서 계산된 것이라는 것을 알 수 있다. 이와 같은 통계량 SS_t를 Nair(1986)는 Taguchi's statistic라 하고 Box and Jones(1986)는 Accumulation statistic T라 하였고, Hamada and Wu(1990)에서는 AA statistic T라고 하여 이에 대한 성질을 논의하고 있다.

(2) $AA-$통계량 T는 분산분석표에 제시된 자유도를 가진 χ^2-분포에 따르지도 않고 분산분석표에 계산된 분산비(分散比) F도 F-분포를 한다고 생각할 수 없다.

(3) 전체 제곱합 SS_T는 미리 계획된 표본의 크기에 의해서 정해지는 하나의 상수이다. 그리고 오차 제곱합 SS_E는 이 상수에서 통계량 T를 뺀 것으로, 따라서 통계량 T와 서로 독립이 아니다. 또, 오차에 관한 평균제곱 $MS_E = SS_E/$자유도 는 일반적으로 1에 가까운 값을 갖게 된다.

이러한 문제점을 알고 순서 지어진 범주형 분할표를 올바르게 분석하기 위하여 우리는 다음과 같은 사항을 검토 할 수 있다.

(1) 다꾸찌의 누적법에 따라 $AA-$통계량 T를 계산하되 이를 직교성분으로 분할하여 검토 할 수 있다.
(2) 비-누적법의 적용이 합당한지를 검토하여 본다.
(3) 점수 수열을 가중치로 한 SN비를 계산하는 방법을 고려해 볼 수 있다.
(4) 로짓모형(Logit Model)을 이용한 분석을 검토 할 수 있다.

참고문헌

[1] 박성현(1990). 응용실험계획법, 영지문화사.
[2] 백운봉(1997). 품질을 좋게 하는 방법, 자유아카데미.
[3] 백운봉(1989). 累積法, 응용통계 4, 1-19, 고대 통계연구소.
[4] 이우선(1997). 최신실험설계, 영풍문고.
[5] Cochran, W.G.(1950). The Comparison of Percentages in Matched Samples, *Biometrika,* 37, 256-266.
[6] Hamada, M. and Wu, C.F.J.(1990). A Critical Look at Accumulation Analysis and

Related Methods, *Technometrics,* 32, 119-162.

[7] McCullagh, P.(1980). Regression Models for Ordinal Data (with discussion) *J. of Royal Statistical Society, Ser. B,* 109-142.

[8] Nair, V.N.(1986a). Components of Cumulative Chi-squared-Type Tests for Ordered Alternatives in Contingency Tables. as Statistical Research Report 18, Murray Hill, NJ: AT&T Bell Laboratories

[9] Nair, V.N.(1986b). Testing in Industrial Experiments With Ordered Categorical Data, *Technometrics,* 28, 283-311.

[10] Nair, V.N.(1987). Chi-squared-Type Tests for Ordered Alternative in Con-tingency Tables, *JASA* 82, 283-291.

[11] Phadke, M.S., Kacker, R.N., Speeny, D.V., and Grieco, M.J.(1983), Off-Line Quality Control in Integrated Circuit Fabrication Using Experimental Design, *The Bell System Technical Journal,* 62, 1273-1310.

[12] Park, S.H.(1996). *Robust Design and Analysis for Quality Engineering,* Chap-man & Hall

[13] SAS Institute Inc. (1990). *SAS/STAT User Guide,* Version 6, Fourth Ed.

[14] Satterthwaite, F.E.(1946). An Approximate Distribution of Estimates of Vari-ance Components, *Biometrics Bulletin,* 2, 110-114.

[15] Snedecor, G.W. and Cochran, W.G.(1989). *Statistical Methods;* Eight Edition, 206-209, The Iowa State University Press.

[16] Taguchi, G.(1974). A New Statistical Analysis for Clinical Data, The Accumu-lating Analysis, in Contrast With the Chi-Square Test, *Saishin Igaku (最新醫學),* 29, 806-813.

[17] Taguchi, G.(1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Costs,* White Plains, NY: UNIPUB/Kraus.

[18] Takeuchi, K. and Hirotsu, C. (1982). The Cumulative Chi-square Method Against Ordered Alternative in Two-Way Contingency Tables, *Reports of Statistical Application Research, Union of Japanese Scientist and Engineers,* 29, 1-13.

[19] Tukey, J.W. (1977). *Exploratory Data Analysis,* Addison-Wesley Pub. Co. Inc.