

순차 범주형 데이터 분석을 위한 최적 모수 설계에 관한 연구

A Study on the optimal parameter design by analyzing
the ordered categorical data

全 泰 俊 : 전남대학교 공과대학 산업공학과
洪 南 杓 : 전남대학교 대학원 산업공학과
朴 鎬 一 : 전남대학교 대학원 산업공학과

<< 요약 >>

제품 개발에 관한 응용 연구 혹은 개발 연구의 실험 결과가 품질 특성의 본질적인 성격이나 측정시의 편의 때문에 순차 범주형 자료(ordered categorical data)로 분류되는 경우가 있다.

본 논문에서는 망목 특성 문제(nominal-the-best type problem)를 분석하는데 있어서 기존의 다구찌 누적법이 순차 범주형 자료 분석법이 안고 있는 문제점들을 고찰하고, 이를 개선하기 위해 품질 손실에 근거한 목표 누적법을 제시한다.

본 논문에서 제시한 기법을 post-etch contact window 데이터에 적용해 본 결과 인자의 최적 수준을 결정하는데 용이하였다.

I. 서 론

제품개발에 관한 응용연구 혹은 기초연구의 실험의 결과로 얻어지는 데이터는 종종 순차 범주형 자료(ordered categorical data)로 분류되는 경우가 많다. 순차 범주형 자료는 제품 또는 공정의 품질 특성이 주관적인 성격을 가지는 경우와 품질 특성을 측정할 수 없거나 어려운 경우에 많이 발생한다. 그러나, 순차 범주형으로 정리된 데이터는 이산적(discrete)이기 때문에, 분석하는데 있어서 연속형 데이터에 비해 많은 어려움이 따르게 된다. 따라서 실험자는 전형적으로 상세한 체계적 모형화(systematic modeling)와 데이터 분석에는 거의 관심을 기울이지 않고 있는것이 보통이다.

실험자의 주요 목적은 공정 또는 제품의 질을 향상시키기 위해 필요한 중요한 인자를 인하고, 그 최적수준을 결정하는 것이다. 따라서 본 논문에서는 이와 같은 실험의 목적을 합리적으로 달성 할 수 있는 data 분석의 한 방법으로 목표치가 있는 순차 범주형 data 분석의 새로운 방법을 제시한다.

본 논문은 모두 4장으로 구성되어 있으며, 2장에서는 다구찌 누적법에서의 범주형 데이터 분석 기법의 특징 및 문제점들을 살펴보고, 3장에서는 망목 특성인 경우의 각 인자의 최적 수준을 결정하는 새로운 분석 기법을 제시한다.

II. 순차 범주형 데이터를 분석하는 기존의 누적법

1. 다구찌의 누적법

Y_{ik} 가 수준 i 에서 범주 k 에 있는 관찰 빈도이면($k=1, \dots, K; i=1, \dots, I$), 다구찌의 누적법은 다음과 같은 세단계를 걸쳐 분석한다.

단계 1. 얻어진 데이터로부터 누적 빈도수(Cumulative Frequencies)를 계산한다. 즉, i 째 수준에서의 누적 빈도는,

$$C_{ik} = \sum_{j=1}^k Y_{ij}$$

가 된다.

단계 2. 각 누적 범주에 가중치(weights)를 할당한다.

$$W_k = \frac{1}{\frac{CUM_k}{CUM_T} \times \left[1 - \frac{CUM_k}{CUM_T} \right]} = \frac{CUM_T^2}{CUM_k(CUM_T - CUM_k)}$$

여기서, CUM_T 는 총 관찰 빈도, CUM_k 는 누적 범주 k 에서의 총 빈도.

단계 3. 가중치를 곱하여 제곱합을 계산한다.

$$SSA = n \sum_{k=1}^{K-1} \sum_{i=1}^I (C_{ik} - C_{.k})^2 / [C_{.k}(n - C_{.k})]$$

여기서, $C_{.k} = \sum_{i=1}^I C_{ik} / I$ 는 수준에 대한 평균을 나타낸다.

<표 2.1>에 주어진 바와 같은 특별한 가상 데이터를 고려하여 보자.³⁾ <표 2.1>의 데이터는 약품의 효과별로 순차적으로 되어 있기 때문에 순차 범주형 데이터이다.

<표 2.1> 처방과 무처방 환자에 대한 가상(hypothetical) 데이터

	약 품 A 효과				약 품 B 효과			
	없음	적음	중간	큼	없음	적음	중간	큼
무처방	40	24	10	6	40	24	10	6
처방	24	40	10	6	24	29	16	11

<표 2.1>에 제시된 약품 데이터를 누적법으로 분석해보자. 누적 빈도수를 계산하여 분산 분석을 실시한 결과가 <표 2.2>에 나타나 있다.

<표 2.2> 분산 분석표

요 인	자유도	제곱합	제곱 평균	F ₀
약품 A	3	6.67	2.22	2.24
약품 B	3	12.17	4.06	4.11**
잔 차	474	467.83	0.99	
총 합	477	480.00		

NOTE: $F_{3, 474}(0.01) = 3.78$

누적법은 범주의 순서를 고려하여 계산하기 때문에 약품 B가 유의하게 분석되고 있다. 그러나, 다구찌의 누적법에서는 누적 범주(Cumulative Category)를 정의할때, 어떤 범주가 목표 범주이든지간에 단순히 첫번째 범주에서 부터 순차적으로 누적하여 분석한다. 따라서 누적분석은 목표 범주에서 품질 손실이 가장 적고, 목표치에서 멀어짐에 따라 더 많은 손실이 발생한다는 손실함수⁽⁵⁾의 개념을 고려하지 못한 분석 방법이다. 이러한 방법으로 목표치가 중간 범주가 되는 망목 특성 문제인 경우에도 같은 방법으로 분석하는 것은 적절한 방법이 되지 않는다.

Ⅲ. 最適 母數 設計을 위한 目標 累積法

3.1 目標 累積法의 概要

다구찌의 누적법에서는 누적 범주를 정의할때, 어떤 범주가 목표 범주(target category)인가에 상관없이 단순히 첫번째 범주에서부터 순차적으로 누적해 나간다. 따라서, 누적법에 의한 분석시 첫번째 범주가 목표 범주인 경우를 제외하고는 목표 범주에서의 득수는 누적 득수이므로 목표 범주에서 많은 누적 득수를 갖는 수준을 최적 수준으로 선정하면 잘못된 결정을 유발할 수 있다. 결과적으로 목표 범주를 고려하지 않고 분석하는 누적법은 손실 함수의 개념을 활용하지 못하는 셈이 된다. 이러한 문제를 해결하기 위해 본 논문에서 제시하는 목표 누적법에서는 누적 범주를 구성할때, 다구찌의 누적법에서 처럼 단순히 첫번째 범주에서부터 누적해 나가지 않고 목표 범주를 기준으로 하여 누적하게 된다. 만일, 목표 범주가 첫번째 범주인 경우의 누적 범주는 다구찌의 누적법에서 정의하는 누적 범주와 동일하게 될 것이다.

다구찌의 누적법과 비교하였을때, 가장 큰 차이점은 누적 범주를 정의하는 방법과 각 인자의 최적 수준을 결정하는 방법이다. 목표 누적법의 간단히 절차를 살펴보면 다음과 같다.

- 단계 1. 누적 범주의 정의 : 목표 범주를 기준으로 목표 누적 범주를 정의하고, 각 누적 범주에서 누적 득수를 계산한다. 예를 들면, 범주가 5개 있고 목표 범주가 Ⅲ범주라면, 각 누적 범주는 다음과 같이 구성되고

누적 범주	범주들
(I)	Ⅲ + Ⅱ + I
(Ⅱ)	Ⅲ + Ⅱ
(Ⅲ)	Ⅲ
(Ⅳ)	Ⅲ + Ⅳ
(Ⅴ)	Ⅲ + Ⅳ + Ⅴ

각 누적 범주에서의 득수(누적 득수)는 누적 범주를 구성하는 각 범주의 득수를 합한 것이 된다.

- 단계 2. 인자별 수준별 누적 득수의 계산 : 각 인자가 득수 분포에 미치는 영향을 파악하기 위해 단계 1에서 구한 누적 득수와 인자 배치표를 이용하여 각 누적 범주에서의 인자별 수준별 누적 득수와 누적 득수 백분율을 계산하고, 선도표(line plot)를 작성한다.

- 단계 3. 분산 분석 : 단계 2에서 계산된 누적 득수에 대해 다구찌의 누적법과 동일한 방법으로 분산 분석을 실시하여 각 인자의 유의성 여부를 파악한다.

단계 4. 최적 수준의 결정 : 단계 3에서 유의하게 판명된 인자의 최적 수준을 결정한다. 최적 수준을 결정할때는 실험자의 판단에 의해 그 손실 정도가 같다고 고려되는 범주들을 하나의 범주로 묶어서 처리하거나, 손실 정도가 가장 크다고 고려되는 누적 범주를 제외시키고 나머지 누적 범주만을 고려하여 최적 수준을 결정하게 되는데 목표 누적 범주와 고려되고 있는 누적 범주에서 항상 높은 곡선을 갖는 수준을 최적 수준으로 선정한다. 만일, 상충이 발생하는 경우는 상충되는 각각의 수준들을 모두 최적 수준으로 선정한다. 단계 3에서 유의하지 않다고 판명된 인자의 최적 수준은 작업성이나 경제성을 고려하여 현용 수준으로 한다.

3.2 適用事例

본 사례는 AT & T Bell Lab.에서 CMOS회로에 필요한 3.5- μm 의 콘택트 윈도우(contact window)형성 공정을 최적화하기 위해 Phadke등이 행한 실험이다.

범주	설명(μm)	범주	설명(μm)
I	윈도우가 열리지 않거나 인쇄되지 않음	III	[2.25, 2.75]
II	(0, 2.25)	IV	[2.75, 3.25]
		V	(3.25, ∞)

Phadke등은 post-etch window 크기를 위와 같이 5 가지 범주로 나누고 실험 데이터를 각 범주별로 분류하여 <표 3.1>의 결과를 얻었다[9]. 윈도우 형성과정은 모두 11가지 단계로 이루어지는데 여섯번째 단계인 플라즈마 에칭후에 목표로 하는 window 크기는 3 μm 이었다.(본 논문에서 언급되는 윈도우는 콘택트 윈도우를 의미한다.) 따라서, 범주 IV가 가장 바람직한 범주가 된다. 이러한 범주들은 윈도우 크기에 대해 순서화되어 있음에 주목하자.

<표 3.1> post-etch window 크기 자료 - 득수

실험 번호	I	II	III	IV	V
1	10	0	0	0	0
2	0	3	3	2	2
3	1	0	0	9	0
4	10	0	0	0	0
5	10	0	0	0	0
6	5	3	2	0	0
7	10	0	0	0	0
8	5	0	0	5	0
9	0	1	4	5	0
10	2	5	3	0	0
11	1	1	2	6	0
12	1	0	1	3	5
13	5	0	3	2	0
14	6	3	1	0	0
15	10	0	0	0	0
16	10	0	0	0	0
17	0	0	4	3	3
18	0	0	0	0	10
전체 평균	.48	.09	.13	.19	.11

3.3 누적 범주의 정의

최적 모수 설계를 위해 목표 범주가 IV범주임을 고려하여 목표 범주 IV를 기준으로 하여 다음과 같이 목표 누적 범주를 정의하고, 각 누적 범주에서의 득수를 계산한다. 그

결과가 <표 3.2>의 우변에 나타나 있다.

$$(I) = IV + III + II + I$$

$$(II) = IV + III + II$$

$$(III) = IV + III$$

$$(IV) = IV$$

$$(V) = IV + V$$

<표 3.2> post-etch window 크기 데이터의 누적 득수

실험번호	I	II	III	IV	V	(I)	(II)	(III)	(IV)	(V)
1	10	0	0	0	0	10	0	0	0	0
2	0	3	3	2	2	8	8	5	2	4
3	1	0	0	9	0	10	9	9	9	9
4	10	0	0	0	0	10	0	0	0	0
5	10	0	0	0	0	10	0	0	0	0
6	5	3	2	0	0	10	5	2	0	0
7	10	0	0	0	0	10	0	0	0	0
8	5	0	0	5	0	10	5	5	5	5
9	0	1	4	5	0	10	10	9	5	5
10	2	5	3	0	0	10	8	3	0	0
11	1	1	2	6	0	10	9	8	6	6
12	1	0	1	3	5	5	4	4	3	8
13	5	0	3	2	0	10	5	5	2	2
14	6	3	1	0	0	10	4	1	0	0
15	10	0	0	0	0	10	0	0	0	0
16	10	0	0	0	0	10	0	0	0	0
17	0	0	4	3	3	7	7	7	3	6
18	0	0	0	0	10	0	0	0	0	10
합 계	86	16	23	35	20	160	74	58	35	55

3.2.2 인자 수준별 누적 득수의 계산

다음 단계는 각각의 인자가 득수 분포에 미치는 효과를 파악하기 위하여, 먼저 <표 3.2>의 우변에 있는 누적 득수를 이용하여 각 누적 범주에서의 인자별 수준별 누적 득수를 계산한다. 예를들면, 인자 A의 1 수준에서의 득수는 L₁₈ 직교 배열(orthogonal array)의 인자 배치[9]를 참조하여 다음과 같이 계산된다.

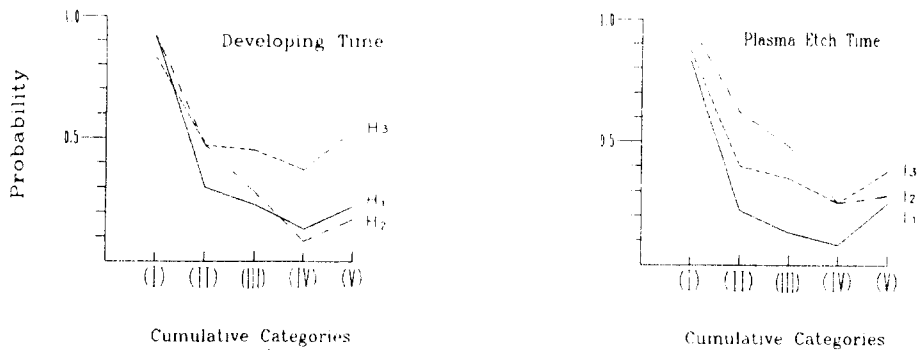
	(I)	(II)	(III)	(IV)	(V)
실험 1	10	0	0	0	0
실험 2	8	8	5	2	4
실험 3	10	9	9	9	9
실험 4	10	0	0	0	0
실험 5	10	0	0	0	0
실험 6	10	5	2	0	0
실험 7	10	0	0	0	0
실험 8	10	5	5	5	5
실험 9	10	10	9	5	5
합 계	88	37	30	21	23

나머지도 비슷한 방법으로 계산할 수 있으며, <표 3.3>는 이를 정리한 것이다. 표에서 밑줄친 부분은 현용 수준이며, 누적 범주에 대한 확률들은 각 범주에서의 누적 득수를 각 수준의 총 관찰 득수로 나눔으로써 얻어진다. 예를 들면, 누적 범주 (I)에서 인자 A의 1 수준에서의 누적 득수 88을 총 관찰 득수 90으로 나누면, 누적 득수 백분율 0.98을 얻는다. 이러한 누적 득수 백분율을 꺾은선 그래프로 나타내면 [그림 3.1]과 같

다. 유의한 인자는 꺾은선 그래프에서 보면, 수준에 따라 확률의 편차가 크게 나타나는데, [그림 3.1]로부터 점성·굽는 온도(BD), 스핀속도(C), 구멍(F), 현상 시간(H), 플라즈마 에칭시간(I)이 수준에 따라 큰 확률 편차를 갖기 때문에 유의한 인자로 고려할 수 있겠다.

<표 3.3> post-etch window 데이터에 대한 인자들의 효과

인자 수준	누적 득수					누적 득수 백분율(%)				
	(I)	(II)	(III)	(IV)	(V)	(I)	(II)	(III)	(IV)	(V)
표면 면적										
A ₁	88	37	30	21	23	0.98	0.41	0.33	0.23	0.26
A ₂	72	37	28	14	32	0.80	0.41	0.31	0.16	0.36
점성, 굽는 온도										
B ₁ D ₁	53	38	29	20	27	0.88	0.63	0.48	0.33	0.45
B ₂ D ₁	60	14	8	2	2	1.00	0.23	0.13	0.03	0.03
B ₁ D ₂	47	22	21	13	26	0.78	0.37	0.35	0.22	0.43
스핀 속도										
C ₁	60	13	8	2	2	1.00	0.22	0.13	0.03	0.03
C ₂	55	33	26	16	21	0.92	0.55	0.43	0.27	0.35
C ₃	45	28	24	17	32	0.75	0.47	0.40	0.28	0.53
굽는 시간										
E ₁	57	26	24	14	17	0.95	0.43	0.40	0.23	0.28
E ₂	43	17	14	7	24	0.72	0.28	0.23	0.12	0.40
E ₃	60	31	20	14	14	1.00	0.52	0.33	0.23	0.23
구멍										
F ₁	50	18	11	6	16	0.83	0.30	0.18	0.10	0.27
F ₂	53	17	14	10	17	0.88	0.28	0.23	0.17	0.28
F ₃	57	39	33	19	22	0.95	0.65	0.55	0.32	0.37
노출 시간										
G ₁	52	26	23	13	21	0.87	0.43	0.38	0.22	0.35
G ₂	48	30	18	7	19	0.80	0.50	0.30	0.12	0.32
G ₃	60	18	17	15	15	1.00	0.30	0.28	0.25	0.25
현상 시간										
H ₁	55	18	14	8	13	0.92	0.30	0.23	0.13	0.22
H ₂	55	28	17	5	10	0.92	0.47	0.28	0.08	0.17
H ₃	50	28	27	22	32	0.83	0.47	0.45	0.37	0.53
에칭 시간										
I ₁	50	13	8	5	15	0.83	0.22	0.13	0.08	0.25
I ₂	58	37	29	15	17	0.97	0.62	0.48	0.25	0.28
I ₃	52	24	21	15	23	0.87	0.40	0.35	0.25	0.38
계	160	74	58	35	55	0.89	0.41	0.32	0.19	0.31



[그림 3.1] 윈도우 크기 데이터에 대한 인자 효과의 선도표

3.2.3 분산 분석

세번째 단계는 꺾은선 그래프로부터 유의하다고 추정하였던 인자의 유의성을보다 정확히 파악하기 위해 분산 분석을 실시한다. 재분석한 결과가 <표 3.4>에 나타나 있다.

<표 3.4> 윈도우 크기 데이터의 재분석

요 인	자유도	제공합	제공평균	F ₀	기여율(%)
A	4	18.36	4.59	8.02	1.79
BD	8	100.98	12.62	22.14**	10.71
C	8	100.61	12.58	22.07**	10.67
E	8	46.20	5.78	10.10	4.63
F	8	58.20	7.28	12.72**	5.96
G	8	24.47	3.06	5.37	2.21
H	8	54.35	6.79	11.92**	5.53
I	8	52.68	6.59	11.56**	5.35
간차	835	477.72	0.57		
총변동	895	900.00			

NOTE: $F_{4,835}(0.01)=3.32$, $F_{8,835}(0.01)=2.51$

<표 3.4>으로부터 각 인자에 대한 F₀값은 유의 수준 5%에서 모두 유의하지만, 기여율이 5% 이하인 A, E, G인자는 그다지 중요하지 않다고 말할 수 있다. 이러한 결과는 꺾은선 그래프로부터 추정한 결과와 일치한다.

3.2.4 최적 수준의 결정

네번째 단계는 각 제어인자의 최적 수준을 결정한다. 손실면에서 보면, 목표 범주 IV에서는 손실이 가장 적고, 반면에 범주 I에서 가장 많은 손실이 발생한다. 만일, 범주 I과 V에서 발생하는 손실은 그 손실 정도가 같다고 가정한다면, 한개의 누적 범주로 간주하여 분석할 수 있다. 본 논문에서는 누적 범주 (I)로 묶었다. 또한, 총 관찰 dot수를 알고 있기 때문에 한개의 범주는 고려하지 않아도 되므로(자유도가 4이다), 누적 범주 (IV), (III), (II)만을 고려 대상으로 하여 최적 수준을 결정한다. 먼저 목표 범주 (IV)부터 각 수준별 누적 dot수를 비교하여 누적 범주 (IV), (III), (II)에 대해서 항상 많은 dot수를 갖는 수준이 존재하면 이를 최적 수준으로 한다. 이러한 수준은 [그림 3.1]의 선도표에서 보면 누적 범주 (IV), (III), (II)에 대해서 항상 높은 곡선을 갖게 된다. 만일, 상충이 발생하는 경우에는 상충되는 각각의 수준들을 모두 최적 수준으로 선정한다. 이러한 방법으로 유의한 인자 BD, C, F, H, I에 대한 최적 수준의 선정 결과가 <표 3.5>에 나타나 있다. <표 3.5>에서 유의한 인자로 판명된 점성·굽는 온도(BD), 구멍(F), 현상 시간(H), 에칭시간(I)은 B₁D₁, F₃, H₃, I₂가 각각 최적 수준임을 쉽게 알 수 있다. 스피ن속도(C)의 최적 수준은 C₂와 C₃를 모두 최적 수준으로 선정하였으며, 확인 실험을 실시하여 한개의 수준을 선택하게 된다. 유의하지 않은 인자 A, E, G에 대해서는 작업성이나 경제성을 고려하여 현용 수준인 A₁, E₂, G₂를 최적 수준으로 선정한다. 따라서, 분석 결과 제안되는 최적 수준 조합은 A₁B₁C₂D₁E₂F₃G₂H₃I₂(최적 수준 조합 1)과 A₁B₁C₃D₁E₂F₃G₂H₃I₂(최적 수준 조합 2)의 2개이다.

<표 3.5> 유의한 인자의 최적 수준

요 인	누적-dot수			최 적	요 인	누적-dot수			최 적
	(IV)	(III)	(II)			(IV)	(III)	(II)	
B ₁ D ₁	20	29	38	●	H ₁	8	14	18	
B ₂ D ₁	2	8	14		H ₂	5	17	28	
B ₁ D ₂	13	21	22		H ₃	22	27	28	●
C ₁	2	8	13		I ₁	5	8	13	
C ₂	16	26	33	●	I ₂	15	29	37	●
C ₃	17	24	28	●	I ₃	15	21	24	
F ₁	6	11	18						
F ₂	10	14	17						
F ₃	19	33	39	●					

3.3 累積法과 目標 累積法의 比較

3.3.1 분산 분석표의 비교

다구찌의 누적법에 의해 분석한 분산 분석표[9]에 의하면, E, F, I 인자가 유의하지 않은 인자로 판명되고 있음을 알 수 있다. 그러나 기여율을 계산하여 기여율이 3%이하인 인자는 유의하지 않다고 보면, A, E, F, I 인자는 유의하지 않게 된다. 반면에 목표 누적법에 의한 <표 3.4>의 결과를 보면, A, E, G 인자가 유의하지 않은 인자임을 알 수 있다.

누적법과 목표 누적법에 의한 분산 분석 결과 얻어진 인자의 유의성 여부를 <표 3.6>에 정리하였다.

<표 3.6> 인자의 유의성 검출 비교

인 자	누적법	목표 누적법
A	●	×
BD	●	●
C	●	●
E	×	×
F	×	●
G	●	×
H	●	●
I	×	●

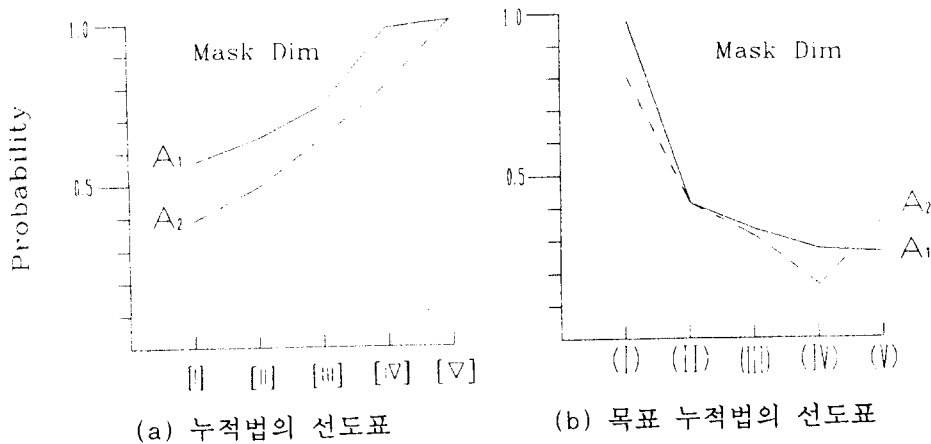
NOTE) ●: 유의함 ×: 유의하지 않음

Phadke 등의 post-etch line width의 연속형 데이터의 분석 결과[9]에 의하면, 모든 인자는 SN 비에 유의하지 않았으며 A, B, C, G, H 인자가 평균에 유의한 인자로 분석되었다. Nair[7]는 post-etch window 크기 데이터를 분석하기 위해 Ridit 분석[2]에서의 점수를 이용하여 평균 효과와 분산 효과를 갖는 인자를 분리하여 찾아낼 수 있도록 두 가지 점수법(scoring method)을 제안하였는데, 그의 분석 결과에 의하면 인자 I는 분산 효과에 유의한 인자로 확인된 바 있다. 그러나, 누적법에서의 분산 분석 결과에 의하면 인자 I는 유의하지 않은 인자로 판명되고 있다.

3.3.2 선도표의 비교

3.1절에서 언급하였던 바와 같이 누적법과 목표 누적법의 가장 큰 차이점은 누적 법

주의 정의에 있다. [그림 3.2]의 선도표의 비교를 보면, 다구찌의 누적법과 목표 누적법과의 차이가 매우 잘 나타나 있는데, 그림에서 좌측의 선도표는 범주 I부터 누적하여 다구찌가 정의한 누적 범주에 대한 선도표이다. 선도표에서 보면, 누적법의 선도표는 A 인자가 수준간에 확률 편차가 크므로 유의한 인자로 예상될 수 있지만, 목표 누적법에 의한 선도표는 수준간의 확률 편차가 작으므로 A 인자는 유의하지 않다고 예상할 수 있다.



[그림 3.2] 인자 A의 선도표 비교

누적법에서 최적 수준을 결정하는 과정을 보면, [그림 3.2]의 (a)에 있는 누적법의 선도표로부터 목표 범주 IV에서 높은 확률 곡선을 갖는 1 수준을 최적 수준으로 선정하지 않고 모든 윈도우가 열리는 것이 중요하기 때문에 범주 I에서 더 적은 득수를 갖는 2 수준을 최적 수준으로 선정하였다. 이는 목표 범주 IV에서 높은 확률 곡선을 갖는 1 수준에 대한 再考를 하게 한다. 따라서, 득수 분포표를 다시 찾게 되고 2 수준 인자가 목표 범주 IV에서 정말로 많은 득수를 갖고 있는지를 확인하게 된다. <표 3.3>에서 보면 목표 범주 IV에서 1 수준의 득수가 2 수준의 득수보다 많음을 알 수 있다. 그러나, 목표 누적법의 선도표를 보면 범주 IV, III, II에서 높은 확률 곡선을 갖는 수준이 1 수준이 되며, 2 수준에 대해서는 생각할 필요가 없게 된다. 나머지 인자들에 대해서도 이와 비슷한 결론을 얻을 수 있다. 이와 같이 목표 누적법에 의하면 각 인자의 최적 수준을 선도표로부터 용이하게 결정할 수 있게 된다.

3.3.3 최적 수준의 비교

Phadke 등은 누적법을 적용하여 유의한 인자의 최적 수준을 결정할 때, 가장 바람직하지 않은 범주 I에서 最少의 득수를 갖고 목표 범주 IV에서 最多의 득수를 갖는 수준을 최적 수준으로 선정하고, 만일 상충이 발생한 경우는 모든 윈도우가 열리는 것이 중요하기 때문에 범주 I에서 더 적은 득수를 갖는 수준을 최적 수준으로 하였다[9].

반면에 목표 누적법에서는 모든 누적 범주를 고려하지 않고, 누적 범주 (IV), (III), (II)만을 고려하여 각 제어 인자의 최적 수준을 결정하였다. <표 3.7>는 각 방법에 의한 인자의 최적 수준의 비교를 나타내고 있다.

표에서 Phadke 방법은 post-etch window 크기의 연속형 데이터를 이용한 SN 비 분석법[12]을 의미하며, 밀출친 인자는 유의한 인자들이고 분산 분석시 유의 수준 5%에 대해서 유의한 인자는 *를 유의 수준 1%에 대해서 유의한 인자는 **로 나타내었다. 목표 누적법에 의해 선정된 최적 수준은 인자 A, BD와 F를 제외하고는 Phadke 방법과 일치하고 있음을 알 수 있다.

<표 3.7> 각 분석 방법의 최적 수준 비교

인 자	방 법		
	Phadke 방법	누적법	목표 누적법
A	A_2^{**}	A_2^{**}	A_1
BD	$B_1 D_2^{**}$	$B_1 D_1^{**}$	$B_1 D_1^{**}$
C	C_3^{**}	C_3^{**}	C_3^{**}
E	E_2	E_2	E_2
F	F_2	F_2	F_3^{**}
G	G_2^{**}	G_2^{**}	G_2
H	H_3^{**}	H_3^{**}	H_3^{**}
I	I_2	I_2	I_2^{**}

4. 結 論

실험의 주요 목적은 공정의 질을 향상시키기 위해 유의한 인자를 올바르게 찾아내고, 그 최적 수준 조합을 결정하는 데 있다. 이러한 이유로 다구찌의 누적법은 실험자의 관심을 끌기에 충분하다. 그러나, 누적법은 품질 손실 함수에서 유도된 SN 비의 개념을 고려하지 않고 분석하며, 목표 범주가 정해져 있는 망목 특성 문제인 경우 품질 특성치의 산포를 줄이는 것이 무엇보다도 중요한 일인데, 누적 통계량은 중요한 분산 효과를 알아내는데 미흡하다. 본 논문에서 제시한 목표 누적법은 손실 함수를 고려한 모수 설계에 그 근거를 두고 있으며, 특히 다구찌의 누적법과는 달리 목표 범주를 고려한 누적 범주를 정의하여 데이터를 분석하였다.

목표 누적법을 실제 데이터 분석에 적용한 사례를 통하여 선도표만으로도 각 인자의 최적 수준을 쉽게 확인할 수 있었으며, 유의한 인자의 검출성이 높았다. 본 논문에서 제시한 목표 누적법은 인자의 최적 수준을 선정하는 과정에서 실험자의 판단에 의해 그 손실 정도가 같다고 고려되는 범주들을 하나의 범주로 묶어서 처리하거나, 손실 정도가 가장 크다고 고려되는 누적 범주를 제외시키고 나머지 범주만을 고려하여 최적 수준을 결정하였는데, 이는 매우 주관적이라고 할 수 있다. 그러므로 본 논문의 이러한 취약성을 보완하여 손실 정도를 정량화시켜 누적 범주를 고려하는 객관적인 방법이 모색되어야 할 것으로 생각된다.

참 고 문 헌

1. Nair, V. N. (November 1986) Testing in Industrial Experiments with Ordered Categorical Data, Technometrics, Vol.28, No.4, 283-291.
2. Phadke, M. S, Kackar, R. N, Speeney, D. V, and Grieco, M. J. (1983) Off-line Quality Control in Integrated Circuit Fabrication Using Experimental Design, The Bell System Technical Journal, Vol.62, 1273-1310.
3. Taguchi, G. (1974) A New Statistical Analysis for Clinical Data. the Accumulating Analysis in Contrast with the Chi-square Test, Saishin Igaku, Vol.29, 806-813.
4. Taguchi, G. (1987) System of Experimental Design, UNIPUB/Kraus International Publications, Vol.1, 73-78.
5. Taguchi, K. and Wu, Y. (1980) Introduction to off-line Quality Control, Central Japan Quality Control Association, Nagoya, Japan