1. 서론

통계학자들은 동질적인 집단(homogeneous groups)을 정의하는 데 종종 관심을 가진다. 점 추정의 정밀도(precision)는 부분적으로는 표본이 추출된 층(strata) 내의 동질성에 따라 달라진다. 유의미한 차이를 검정하는 데 사용되는 통계적 검정들은, 층 내 동질성을 포함하는 비교뿐만 아니라 층 간 차이를 포함하는 비교에 근거한다. 표본 추출이나 추론 문제 외에도, 한 모집단을 서로 뚜렷하게 대비되는 하위 그룹들로 어떻게 나눌 수 있을지, 즉 같은 그룹의 구성원들이 서로 비교적 유사한지를 아는 것도 중요하다.

본 논문은 다음과 같은 통계적 문제 기술에 초점을 맞춘다: K개의 원소 집합이 주어졌고, 각 원소는 가중치 w_i와 수치적 값 a_i를 가진다고 하자. 그리고 K보다 작은 양의 정수 G가 주어졌을 때, K개의 원소를 G개의 상호 배타적이면서도 전체를 포괄하는 하위 집합들로 나누는 체계적이고 실용적인 절차를 찾는 것이다. 이때 다음의 가중 제곱합

$$D = \sum_{i=1}^{K} w_i (a_i - \hat{a}_i)^2$$

이 최소화되도록 한다. 여기서 $\hat{a}_i$는 원소 i가 속한 하위 집합 내 a값들의 가중 산술 평균(weighted arithmetic mean)을 의미한다. 이 문제는 **그룹화 문제(grouping problem)** 라고 불린다. 이때의 D값은 분산 분석(ANOVA)의 문맥에서 **제곱 거리(squared distance)** 라고 부르며, 널리 알려진 제곱합(sums of squares)으로 인식된다. 그룹화 체계는 일반적으로 **분할(partition)** 이라고 하며, D가 최소가 되는 분할은 **최소제곱 분할(least squares partition)** 이라고 한다.

그룹화 문제는 두 가지 하위 유형으로 구분된다: (1) **비제약 문제(unrestricted problem)** — 분할에 대한 제약이나 부가 조건이 없는 경우, 그리고 (2) **제약 문제(restricted problem)** — 사전 지식, 이론, 또는 편의상의 이유로 특정 조건이 미리 주어지는 경우. 각 유형의 예시는 이어지는 장에서 설명될 것이다. 문헌에서 소개된 바 있으며, 단순한 유형의 문제를 다루는 계산 방법들도 함께 제시될 것이다.

연속적인 빈도 분포에 대한 유사한 문제는 Dalenius [3], [5] 및 Dalenius와 Gurney [4]에 의해 연구된 바 있다. 정규분포의 특수한 경우는 Cox [2]의 최근 메모에서도 다뤄졌다. 이들이 제안한 방법은 다음의 조건에서 본 논문의 문제를 푸는 데 유용하다: 개체 수가 매우 많고, 이들의 분포가 비교적 단순한 연속 곡선으로 근사될 수 있으며, 그룹 수 G가 비교적 작을 때—예를 들어 5 이하—그리고 허용 가능한 분할에 어떤 제약도 없는 경우이다. 그렇지 않은 경우에는 본 논문의 접근 방식이 더 적절하다고 여겨진다.

본 논문에서는 w_i와 a_i가 완전한 확실성 하에 주어진 것으로 가정하며, 이는 기술적이고 비확률적인 접근을 의미한다. 그러나 이러한 접근은 표본추출 및 다른 확률적 응용에도 이어진다. 또한 동질성 측도로서 제곱편차(squared deviations)를 더하여 얻은 D를 별다른 정당화 없이 사용하지만, 이 값은 실용적인 문제에 있어 유용하고 적절한 척도로 간주된다.[2]

2. 비제약 문제 (The Unrestricted Problem)

비제약 문제는 친숙한 상황을 고려함으로써 쉽게 이해될 수 있다. 예를 들어, 모집단 내의 관련 변수에 대한 정보가 있을 때, 비례 층화 표본 추출(proportional-stratified sampling)에 적합한 층(strata)의 수를 선택하는 최적 방법을 찾고자 한다고 가정하자.[3]

Hansen, Hurwitz, 그리고 Madow [8]는 Atlanta 가족들의 소득 수준에 대한 데이터를 다룬 연구에서, Mensershausen [9]의 선행 연구를 바탕으로 이러한 층화 표본 추출의 논의를 제시하였다. 이들 가족은 소득 계층별로 10개 계급으로 나뉘며, 이들의 빈도 분포는 그림 791에 제시되어 있다. 문제는 이 10개 계급을 3개의 큰 계층으로 합쳐서 전체 가구의 평균 소득 추정량이 분산이 작은 상태로 나오도록 하는 것이다. 즉, 층화된 표본을 기반으로 한 평균 추정량이 분산이 작도록 만드는 것이 목표이다.[4]

다양한 층화 방법과 표본 추출 방법이 제안된다. 여기서는 원래의 10개 계급을 3개의 층으로 결합하는 다양한 가능한 조합에 초점을 맞출 것이다. 이때 각 층 간의 표본 수는 비례적으로 할당되며, 각 층 내에서는 무작위로 표본을 추출한다고 가정한다. 또한 모집단 평균을 추정하는 표본 평균이 전체 평균과 같다고 가정한다.

이러한 조건 하에서, w_i가 모집단에서 소득 계급 i의 가중치이고, a_i가 계급 i의 평균 소득이며, $\hat{a}_i$가 계급 i가 속한 층의 평균 소득일 때, 전체 추정치의 분산은 방정식 (1)에 나오는 D에 비례한다는 것이 잘 알려져 있다. 즉, 계급들의 평균 소득에서 벗어난 정도를 최소화하는 방식으로 계급을 층으로 나누는 것이 전체 추정치의 분산을 줄이는 데 효과적이다.

해당 문제의 해답은—그림 791을 육안으로 보아서는 명확하지 않지만—다음과 같다. 원래의 10개 소득 계층을 낮은 소득에서 높은 소득으로 (그림 791의 왼쪽에서 오른쪽으로) 번호를 매기면, 클래스 16~~을 하나의 계층 (stratum)으로,~~ 79를 또 다른 계층으로, 그리고 10을 단독으로 세 번째 계층으로 둔다. 이 계층화 방법은 Hansen, Hurwitz, Madow나 Dalenius, Gurney가 동일한 예시를 논의하면서도 언급하지 않은 방식이다.[5]

이 소규모 문제에서는 손으로 계산하여 해를 구하는 것도 충분히 가능하다. 직관적으로도 명확하며, 다음과 같이 증명할 수 있다: 10개의 원래 계층이 소득에 따라 정렬되어 있을 때 ($i < j$ 이면 $a_i < a_j$), 고려해야 할 분할은 연속 분할(contiguous partitions)뿐이다. 연속 분할이란, 순서가 정해진 원소 집합을 연속적인 하위 집합들로 나누는 분할로, 아래 조건을 만족해야 한다:

만약 원소 i, j, k가 순서 $i < j < k$ 를 가지고 있고, i와 k가 같은 하위 집합에 속한다면, j 역시 같은 하위 집합에 속해야 한다.

따라서 최적의 그룹화를 찾기 위해서는 단지 가능한 36가지 연속 분할 각각에 대해 D 값을 계산하고, 그중 최소 D 값을 가지는 분할을 선택하면 충분하다.[7]

일반적인 비제약 문제에서 K개의 원소를 G개의 그룹으로 나누는 경우도, 동일한 논리에 따라 $\binom{K-1}{G-1}$개의 연속 분할만을 고려하면 충분하다.[8]

이해를 돕기 위해 요약하자면:

- 정렬된 원소들을 연속적인 그룹으로만 나누면 최적해를 놓치지 않는다.
- 따라서 전체 분할 경우의 수가 아니라 연속 분할 경우의 수 $\binom{K-1}{G-1}$만 계산하면 된다.
- 이 가정은 원소들이 수치적으로 정렬되어 있다는 전제에서 유효하다.

아래는 "3. The Restricted Problem" 섹션의 한국어 번역입니다:

3. 제약 문제 (The Restricted Problem)

앞선 예제에서 연구자는 그룹화 문제를 해결하는 과정에서, 10개의 원래 계층(소득 수준 순)을 우선 정렬한 뒤, 이들을 연속적인 그룹으로 나누어 제곱합 D를 최소화하는 방법을 자유롭게 선택할 수 있었다. 하지만 이제는 이 원소들을 **사전 지식에 기반한 특정 순서(a priori ordering)** 에 따라 정렬해야 한다고 가정해보자. 이 순서는 소득 순서와는 다를 수 있다. 다시 말해, 제곱합을 최소화하려는 목적은 같지만, 그 해가 유효하기 위해서는 원소들이 **사전 순서**에 따라 연속적으로 분할되어야 한다.

예를 들어, 원래의 10개 원소가 같은 거리 내에 위치한 가족들이고, 이들의 **위치 순서**가 알려져 있다고 하자. 이 경우에도 최대한 동질적인(소득 기준) 3개의 그룹을 만들고 싶지만, 동시에 이 그룹들이 **위치상으로도 연속적인** 집단이 되도록 해야 한다. 이와 같은 조건은 D를 최소화하는 목적 외에 부가적인 **제약 조건**으로 작용하며, 이로 인해 해가 앞선 경우보다 더 큰 D 값을 가질 수도 있다.

이 외에도 여러 형태의 부가 조건이 있을 수 있다. 예를 들어:

- 원소들이 완전한 순서가 아닌, **부분 순서**만 만족해야 할 수도 있다.

- 원소들이 다차원 공간의 점들과 연결되어 있을 수도 있으며, 이 경우는 a_i 값 외에 해당 점들의 좌표에 대한 수학적 제약이 적용될 수 있다.[9]
- 특정 분할은 공간적 위치나 정렬 조건과 무관하게, **명시적으로 금지**될 수도 있다.

이 논문 두 번째 단락에서 정의한 그룹화 문제에 대하여, 만약 어떤 형태로든 사전 제약이 존재한다면 이 문제는 **제약 문제(restricted problem)** 라고 부른다. 이때 제약이란 K개의 원소를 G개의 하위 집합으로 분할하는 경우의 수에 대해, **서로 배타적이고 전체를 포괄하는 분할이라는 조건 외에 추가적인 제약 조건이 존재하는 경우**를 말한다.

대부분의 실용적인 문제는 이 제약 문제 유형에 속한다. 왜냐하면 실제 연구자들은 **항상 사전 지식이나 현실적인 제약, 편의성** 등을 고려하여 그룹화 조건에 영향을 주기 때문이다. 실제로 제약 문제의 범위가 너무 넓기 때문에 일반적인 해법을 제시하기 어려우며, 경우에 따라 **불가능에 가깝다**.

실용적으로 의미 있는 주요 제약 조건들을 정의하는 것조차 이 논문의 범위를 넘는다. 대신, 다음 장에서는 단순한 형태의 제약 문제 하나를 예로 들며 계산 예제를 통해 설명할 것이다. 그 예제는 다음과 같다: **a_i의 수치적 순서를 따르는 완전한 사전 순서(a priori ordering)가 주어진 경우**, 이 순서를 따르며 분할하는 문제이다.

Wilks와 Roberts [11]의 교재에도 이와 유사한 문제가 소개된다.

최대 동질성을 위한 그룹화 (Grouping for Maximum Homogeneity)

96년에 걸친 호수 수위 변화에 대한 예제가 소개된다. 그림 793에 재현된 그래프는 수위가 높았던 시기와 낮았던 시기가 있었음을 보여주지만, 뚜렷한 규칙성이나 주기성은 나타나지 않는다. 이 현상의 분석은 주로 구간별 이동평균(runs and moving averages)에 기반하고 있다. 이제 G개의 시기를 정의하고자 한다고 가정하자. 이때, 각 시기 내에서의 수위 변동이 제곱 거리 D로 정의되어 최소화되도록 하는 것이 목표다.

물론 각 시기는 시간 순서상 연속된 연도로만 구성되어야 하며, 이것이 바로 **a priori 순서 제약** 이다. 이 제약 조건이 주어질 때는, 앞서 제시된 형식의 **제약 문제(restricted problem)** 가 된다. 이 경우에는 모든 가중치 w_i가 1로 동일하며, 해는 시간 순서를 따른 연속된 분할이어야 한다. 단, 이 순서가 반드시 수위 순서일 필요는 없다.

1부터 10까지의 G 값에 대해 문제를 푼 결과는 표 794에 제시되어 있다. 이 해답은 다음 섹션에서 설명될 자동 컴퓨터 프로그램을 통해 제공되었다. 표에서:

- 첫 번째 열은 다양한 G 값 (군집 수),
- 두 번째 열은 최소화된 D 값,
- 그 다음의 삼각형 형태 배열은 "P"라는 헤드로 시작하며, 각 G에 대해 최적의 분할을 식별한다.

각 행의 최적 분할은 해당 행의 G값에 대해 최소의 D 값을 제공하며, 각 하위 집합의 가장 큰(항상 96) 원소 번호로 식별된다.

예를 들어, $G = 3$ 일 경우 세 시기는:

- 연도 1~30
- 연도 31~61
- 연도 62~96

이 된다.

$G = 1$ 인 경우는 자명하다. 전체 연도가 하나의 집합이 되며, 이때 D 값은 전체 평균에 대한 제곱 편차들의 합이다.

덧붙이자면, 이 프로그램은 **복수의 해가 존재하더라도 하나만 출력**하며, 이 프로그램에 대해서는 뒤에서 더 자세히 논의된다.

다음은 이미지 속 **"4. METHODS OF SOLUTION (해결 방법)"** 섹션의 전문 한국어 번역입니다:

4. 해결 방법 (Methods of Solution)

작은 규모의 비제약 문제—예를 들어 $K \leq 20$, $G \leq 5$ 정도의 경우—에서는 가능한 모든 연속 분할을 나열하고, 각 분할에 대해 D 값을 계산한 뒤, 그 중에서 최소 D 값을 갖는 분할을 선택하는 방식으로 해를 구하는 것이 현실적으로 가능하다. 동일한 크기의 일부 제약 문제들도 이러한 방식으로 해결될 수 있다. 제약 문제의 경우, 허용 가능한 분할 수는 비제약 문제보다 적을 수 있지만, 해당 제약 조건을 충족하는 분할만을 고려하기 위한 작업에는 추가 시간이 소요된다.

K는 크지만 G는 여전히 작을 경우, 일부 비제약 문제는 a_i의 크기순으로 정렬한 빈도 분포를 시각적으로 관찰하여 해 또는 근사 해를 얻는 것이 가능하다. 그룹 간 경계는 데이터가 희소하거나 가중치가 작게 나타나는 부분에 설정할 수 있다. 하지만 이러한 희소 영역의 수가 분할해야 하는 개수와 일치하지 않는 경우, 이 방법은 쉽게 적용되지 않는다.

a_i의 분포가 비교적 단순한 연속 함수(예: 단봉형 unimodal 분포)로 근사 가능할 경우, Dalenius의 방법 [3], [5]을 적용할 수 있다.[10]

보다 일반적인 문제—예를 들어 a_i가 임의의 분포를 가지거나 K 또는 G가 더 클 경우—에는 조합적 접근법(combinatorial approach)이 요구된다.[11]

원칙적으로는, K개의 원소를 G개의 하위 집합으로 나누는 모든 분할의 수는 유한하기 때문에, 각 분할을 하나씩 고려하여 최소 D 값을 갖는 분할(또는 분할들)을 선택하는 방식으로 해를 찾는 것이 가능하다. 그러나 이러한 접근을 수정하지 않는 한, 조합 수가 지나치게 커져서(10번 주석에 제시된 수식이 보여주듯이) 모든 가능성을 고려하는 것이 사실상 불가능해진다. 이는 최신 디지털 컴퓨터로도 마찬가지이다.

예를 들어, Lake Michigan-Huron 호수 수위 문제에서 $K = 96$, $G = 10$일 때 가능한 연속 분할 수는 1조(1 trillion)를 약간 넘는다. 프린스턴 유형의 고속 컴퓨터가 D 값을 10밀리초에 하나씩 계산한다고 가정하고, 하루 24시간 풀가동하더라도, 이 모든 분할을 하나씩 비교하는 데 **280년 이상**이 소요된다.

제곱 거리 함수의 가산적 성질 덕분에, 부분 최적화(suboptimization)를 통해 계산량을 크게 줄일 수 있다. 이러한 최적화 절차는 다음 보조정리(lemma)에 의해 암시된다:

부분최적화 보조정리(Suboptimization Lemma):
집합 A를 서로소인 두 부분집합 A_1과 A_2로 분할했다고 하자. 만약 P_1'이 A_1을 G_1개의 부분집합으로 최소제곱 분할(least squares partition)한 것이고, P_2'가 A_2를 G_2개의 부분집합으로 최소제곱 분할한 것이라면, $A_1 : A_2$를 $G_1 + G_2$개의 하위집합으로 나누는 부분분할 중 최소제곱 분할은 $P_1' : P_2'$가 된다.

즉, 일단 집합 A_1에 대해 최소제곱 분할이 계산되었다면, A_2의 여러 분할에 대해 다시 계산하지 않아도 된다. 적절한 중간 결과를 기록해두면, 이 보조정리를 통해 전체 원소 집합에 대한 다양한 분할을 따로 계산하지 않고도 해결할 수 있다. 실제로 이 정리를 적용했을 때, Lake Michigan-Huron 문제에서 $G = 1$부터 $G = 10$까지 해를 구하는 데 **단 3분**이 소요되었다.

이 보조정리는 허용 가능한 분할 조건 하에서도 유효하므로, **제약 문제**에도 적용 가능하다.

일리노이 대학교에서 **Illiac 자동 디지털 컴퓨터**를 위한 프로그램이 작성되었으며, 작성자 본인이 비제약 문제와 사전 순서가 주어진 제약 문제 모두에 대해 그 정확성을 검증하였다. 이 프로그램은 $K \leq 200$, $G \leq 10$ 범위 내에서 다음을 해결할 수 있다:

- $G = 1, 2, \ldots, \hat{G}$ (단, $\hat{G} \leq 10$)에 대해 각각,
- 최소 제곱 거리 D를 가지는 최적의 분할을 식별

가장 큰 문제(즉, 가장 큰 K와 G)에 대한 실행 시간은 **약 14분**이다 (입출력 포함).

데이터는 표준 전신 타이프(teletype tape)를 통해 입력된다.
비제약 문제라면 w_i와 a_i의 수치 값을 해당 순서대로 컴퓨터에 입력해야 한다.
Illiac은 이들에 대해 연속성과 위 보조정리를 고려하여 관련된 모든 분할에 대해 D 값을 체계적으로 계산한다.
G개의 집합으로 나누는 문제에서는, 먼저 $G-1$개의 분할 문제를 풀면서 얻어진 결과들을 체계적으로 재사용한다.

표 794에 있는 해답은 이 프로그램의 출력 그대로이며, 총 $\hat{G}$개의 행으로 구성되어 있다. 일부 의미 없는 소수점은 생략되었으나, 나머지는 실제 출력과 동일하다.

이 계산 프로그램에는 세 가지 주요 제한사항이 있다:

1. K와 G의 크기가 여전히 제한적이다.
2. 크기를 더 키우기 위한 수정이 가능하더라도 실행 시간은 빠르게 늘어난다.
3. 다중 해(multiple solutions)를 식별하지 않으며, 근사 해(near solution)도 제공하지 않는다.

또한, 완전한 일차원 순서를 제외한 **제약 문제**는 처리할 수 없다. 앞으로 더 나은 프로그램의 발전이 기대된다.

5. 일반화 (Generalization)

이 논문에서 공식화된 그룹화 문제를 일반화하는 몇 가지 방법을 간략히 제시한다. 이 문제에 대한 확률적 접근법은 문헌 [6]에 제시되어 있으며, 이는 고정된 G를 전제로 하는 대신, G의 선택 자체를 의사결정의 일부로 포함시키려는 논리적 근거도 제공한다. 즉, G의 선택은 더 상세한 정보의 가치와 그에 따른 **세부 정보 비용(cost of detail)**을 비교하는 문제로 해석된다.

명시적인 비용 이론이 없더라도, G가 변화함에 따라 D 값이 어떻게 변하는지를 아는 것은 (예: 표 794의 두 번째 열 참조) 연구자가 처음에 G 값을 확신하지 못할 때 적절한 G를 선택하는 데 도움이 될 수 있다. 이를 위해 “Illiac” 프로그램이 설계된 것이다.

3장에서 언급한 바와 같이, 한 차원을 넘는 다차원 속성 공간에서 요소들을 정의하는 것도 가능하나, 본 논문에서는 제곱 거리 D를 하나의 차원에서 측정하는 방식으로 제한하였다. 물론 이론적으로나 계산적으로 보았을 때, 다차원 공간에서 정의된 거리 기준(distance criterion)을 사용하는 것이 가장 바람직할 것이다. 이런 접근의 필요성은 예를 들어, **Hagood와 Bernert [7]의 다변량 층화 문제(multivariate stratification problem)**에서 확인된다.

이와 같은 접근에서는, 거리 판단에서 각 차원의 상대적 중요성을 반영하기 위한 **가중치 체계(weighting system)**가 필수적이다.

이 논문의 **일차원적 접근 방식**은 **다차원 그룹화 문제**에 대한 근사 해법으로 사용될 수도 있다. 예비 단계로, 각 요소에 대해 단일 값으로 축소시킬 수 있으며, 이는 예를 들어:

- 첫 번째 주성분(first principal component)을 추출하거나,
- 다른 요인분석(factor analysis) 방법을 통해 수행할 수 있다.

이러한 절차는 현재는 잘 알려져 있으며, 많은 컴퓨터 시스템에서 정형화되어 있다. 이후에는 본 논문에서 제시한 계산 기법을 활용하여 원하는 그룹 수로 요소들을 분할할 수 있다. 물론 이 근사화의 적절성은 **다차원**

분포에서 첫 번째 주성분이 얼마나 지배적인가에 따라 달라진다.

부록: 최소제곱 분할은 연속 분할임의 증명

텍스트에서 정의된 **비연속(non-contiguous)** 분할을 생각해보자.

원소 i 와 k 는 평균이 $\hat{a}_{ik}$ 인 한 부분집합에 속하고

원소 j 는 평균이 $\hat{a}_j$ 인 다른 부분집합에 속하며

$a_i < a_j < a_k$ 라 하자.

이때, $\hat{a}_{ik}, \hat{a}_j$ 의 값이 어떠하든 다음 중 하나는 항상 참이다

1. $|a_j - \hat{a}_j| \geq |a_j - \hat{a}_{ik}| > 0$
2. $|a_i - \hat{a}_{ik}| \geq |a_i - \hat{a}_j| > 0$
3. $|a_k - \hat{a}_{ik}| \geq |a_k - \hat{a}_j| > 0$

즉, 세 개의 서로 다른 점 a_i, a_j, a_k 중 하나는
자신이 속한 부분집합의 평균과의 거리보다,
다른 부분집합 평균과의 거리가 더 가깝거나 같으면서 두 거리 모두 양수이다.

그런 점을 a 라 하고,
그가 속한 집합을 A, 평균을 $\hat{a}$,
"외부" 집합을 B, 평균을 $\hat{b}$ 라 하면,

$$|a - \hat{a}| \geq |a - \hat{b}| > 0 \tag{4}$$

본문의 식 (1)에 따라, 주어진 분할의 제곱 거리 D 는 다음과 같이 쓸 수 있다:

$$D = \sum_{i=1}^{K} w_i a_i^2 - W_A \hat{a}^2 - W_B \hat{b}^2 - R \tag{5}$$

여기서

- $W_A = \sum_{i \in A} w_i$,
- $W_B = \sum_{i \in B} w_i$,
- R은 A, B 이외의 다른 집합들에 대한 가중 평균 제곱합.

이제 a를 집합 A에서 B로 옮긴 새로운 분할을 생각해보자.
새로운 평균을 각각 $\hat{a}'$, $\hat{b}'$ 라 하면:

$$(W_A - w)\hat{a}' = W_A \hat{a} - wa \tag{6}$$

$$(W_B + w)\hat{b}' = W_B \hat{b} + wa \tag{7}$$

여기서 w는 점 a의 가중치.

새 분할의 제곱 거리 D'는:

$$D' = \sum_{i=1}^{K} w_i a_i^2 - (W_A - w)\hat{a}'^2 - (W_B + w)\hat{b}'^2 - R \tag{8}$$

이제 식 (8)에서 식 (5)를 빼고, (6), (7)을 대입하고 정리하면 다음이 된다:

$$D - D' = w\left[\frac{W_A}{W_A - w}(a - \hat{a})^2 - \frac{W_B}{W_B + w}(a - \hat{b})^2\right] \tag{9}$$

식 (4)와 모든 가중치가 양수이며 $W_A > w$ 인 조건에서, 위 식의 우변은 **항상 양수**이다. 따라서 $D > D'$.

즉, 어떤 비연속 분할이라도 항상 더 작은 D 값을 갖는 **다른 분할로 대체 가능**하며, 그 분할은 동일한 집합 수를 유지한다.
결국, **최소 제곱 분할은 항상 연속 분할이어야 한다.**