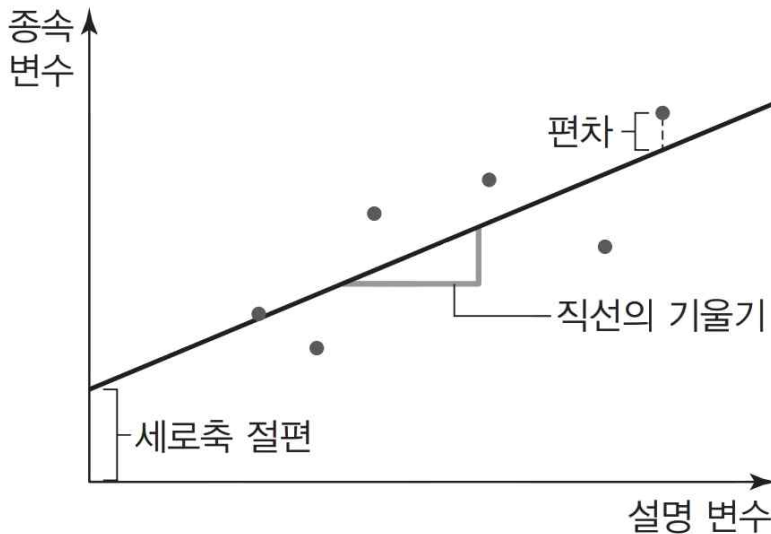


사회 과학 분야에서 수치 자료를 분석할 때 기본적으로 쓰이는 방법 중 하나가 최소 제곱법이다. 최소 제곱법은 이론적으로 영향을 주는 요소인 ①**설명 변수**가 영향을 받는 대상인 ②**종속 변수**에 비례적인 영향을 준다고 가정하고, 그 영향 정도를 실제 수치 자료로부터 추정하는 데 쓰는 보편적인 방법이다. 설명 변수가 변화할 때 종속 변수가 변화하는 평균치를 ③**회귀 계수**라 부르는데, 바로 이 회귀 계수가 ④**최소 제곱법을 통해 구하고자 하는 대상**이다. 구체적으로 보면, 최소 제곱법은 종속 변수의 실제 수치에서 종속 변수의 예측치를 뺀 수치를 편차라고 지칭하고, 각 실제 수치에 대한 편차를 ⑤**각각 제공하여** 모두 더한 것을 최소화한다. 종속 변수의 예측치는 설명 변수에 미지의 회귀 계수를 곱하여 구함으로써 편차를 미지의 상태인 회귀 계수의 함수로 표현하고, 함수인 ⑥**편차 제공의 합을 최소화하는** 구체적인 회귀 계수 수치를 구하는 것이다. 현실에서는 설명 변수 외에도 다양한 요소들이 종속 변수에 영향을 미치기 때문에, 설명 변수에 의해 이론적으로 예측되는 종속 변수와 실제 관찰되는 수치 사이의 ⑦**과리**를 편차로 볼 수 있다.



최소 제곱법을 시각적으로 다음과 같이 설명할 수도 있다. ⑧**2차원 평면**에 설명 변수를 가로축에 두고, 종속 변수를 세로축에 둔다. 예를 들어 6개 관찰 대상에 대해 설명 변수와 종속 변수의 실제 수치 자료가 수집되었다면, 각 관찰 대상의 설명 변수와 종속 변수 조합을 평면에 한 점으로 표시할 수 있다. 설명 변수와 종속 변수의 관계를 보여 주는 6개의 점을 모두 표시하고, 점들 사이를 지나가는 임의의 직선을 그린다. ⑨**각 점으로부터 직선까지의 수직 거리를 편차의 절댓값으로 보고**, 6개의 수직 거리 각각을 제공한 수를 모두 더한다. 최소 제곱법은 이 편차 제공의 합을 직선의 기울기와 세로축 절편을 약간씩 바꾸어 구한 편차 제공의 합과 비교하여, 편차 제공의 합이 가장 작은 직선을 택하는 과정으로 이해할 수 있다. ⑩**편차 제공의 합이 가장 작은 직선의 기울기가 설명 변수와 종속 변수의 관계를 나타내는 회귀 계수**이다. 결과적으로 ⑪**직선은 점들의 아래쪽이나 위쪽이 아니라 사이에 위치하게 된다**. 또한 ⑫**각 편차를 제공하기 때문에 직선에서 상대적으로 멀리 떨어져 있는 점들이 기울기에 영향을 더 많이 미친다**.

최소 제공법을 가계의 소득과 식료품 소비의 연관 정도를 알아보는 데 적용한다고 가정하자. ⑬**소득에 따라 ⑭식료품 지출액이 어떻게 달라지는지는 식료품 지출액이 소득의 영향을 받는다는 인과 관계를 전제로 최소 제공법을 이용하여 계산할 수 있다.** 만약 ⑮**소득과 식료품 지출액이 각각 월간 100만 원 단위 수치이고 ⑯최소 제공법으로 구한 직선의 기울기가 0.1이라면, 월 소득이 100만 원 증가할 때 식료품 지출액이 평균적으로 10만 원 증가하는 것으로 볼 수 있다.** 한편, 세로축 절편이 0.3이라면 ⑰**소득이 0이더라도 식료품 지출액이 평균적으로 30만 원이라는 의미이다.** 소득 자료에 공적 및 사적 보조금이 포함되어 있지 않을 경우 보조금 수입으로 식료품 소비를 할 수 있게 되므로 소득이 0이더라도 ⑱**세로축 절편이 0보다 클 수 있다.** 또는 실제로는 소득이 0인 경우가 없더라도 수치 자료에 최소 제공법을 적용한 결과 소득이 0이라는 가상적인 상황에 대한 식료품 지출액이 0보다 크게 도출될 수 있다.

위와 같이 최소 제공법을 사용하면 잘 알려진 엥겔의 법칙도 쉽게 해석이 된다. 엥겔의 법칙은 ⑲**소득이 적을수록 소득에서 식료품 지출액이 차지하는 비율이 높다**는 것이다. 설명 변수를 가계의 소득, 종속 변수를 식료품 지출액으로 설정했을 때, ⑳**세로축의 절편이 0보다 크고 직선의 기울기는 0보다 크고 1보다 크지 않다면 특정 소득에서 식료품 지출액이 차지하는 비율은 소득이 적을수록 높게 나타난다.** 이러한 현상이 나타나는 이유는 무엇일까? 모든 소비는 소득의 영향을 받지만, 식료품은 누구나 필수적으로 소비하기 때문에 ㉑**식료품 지출액의 규모는 예컨대 문화 예술 분야 등에 대한 지출에 비해서 소득 수준에 따라 크게 다르지 않은 편이기 때문이다.** 소득의 큰 부분을 식료품 소비에 쓰면 최소한의 생계유지 외에 다른 지출을 할 여유는 적게 마련이므로 소득 대비 식료품 지출액의 비율은 삶의 질을 ㉒**판단하는 근거 자료**가 될 수 있다. 소득 대비 식료품 지출액의 비율은 한 국가 내의 소득 계층별로 파악되기도 하고, 국가 전체의 평균치로서 일정 시점 간격으로 집계되기도 한다. 또한 국가별 비교 자료로 쓰이기도 한다.

㉓**복지 정책 시행에 중요한 기준** 중 하나인 ㉔**최저 생계비**를 산정할 때, 미국에서는 ㉕**최소한으로 필요한 식료품 지출액에 3을 곱하는 방식**을 이용한다. 저소득층 가계의 소득에서 식료품 지출액이 차지하는 비율이 1/3 정도 된다는 조사 결과를 바탕으로 한 것이다. 이렇게 소득과 식료품 지출액의 관계가 정책에 직접 이용되는 사례도 있지만, ㉖**소득의 충분성**이나 ㉗**식료품 가격 수준의 적정성** 등을 판단해야 하는 정책들에서도 ㉘**소득과 식료품 지출액의 관계**에 대한 구체적인 정보는 필수적이다. 이러한 사례 외에도 최소 제공법은 어떤 현상들 간에 인과 관계가 있다고 제시하는 이론을 수치 자료로 검증하거나, 이론에서 예측하는 변수 간 관련 정도를 구체적 수치로 나타내고자 할 때 보편적으로 사용된다.

1. ①과 ②의 관계를 서술하세요.

인과 관계(① : 원인, ② : 결과)

2. ①을 나타내는 다른 말을 서술하세요.

독립 변수

3. ③이 '무엇'의 평균치인지 식으로 나타내세요.

종속 변수의 변화량/설명 변수의 변화량

4. 다음은 ④의 의미를 나타낸 것입니다. 빈 칸을 채우세요.

최소 제곱법에서 구하고자 하는 (목표)가 되는 (수치)

5. ⑤의 이유를 추론하여 서술하세요.

편차는 종속 변수의 실제 수치에서 종속 변수의 예측치를 뺀 수치이다. 따라서 종속 변수의 실제 수치가 예측치보다 큰 경우 편차가 0보다 크고, 종속 변수의 실제 수치가 예측치보다 작은 경우 편차가 0보다 작다. 이러한 상황에서 편차들을 그냥 다 더하는 경우 편차들의 부호가 다르기 때문에, 종속 변수의 실제 수치와 예측치의 차이가 크더라도 정작 편차들을 다 더한 값은 0에 가깝게 되어 편차들을 더한 수치가 실제 수치와 예측치의 차이를 반영하지 못하는 상황이 발생할 수 있다. 편차를 제곱하면 그 값은 항상 0 이상이 되고, 편차가 클수록 편차의 제곱도 크기 때문에 편차들의 제곱을 더하는 경우 이러한 상황을 방지할 수 있다.

6. 다음은 ⑥의 의미를 나타낸 것입니다. 빈 칸에 들어갈 알맞은 말을 고르거나 빈 칸을 채우세요.

- (1) (종속) 변수에 대한 (예측치)가 실제 (종속) 변수의 수치와 잘 (들어맞게) 하는
- (2) (종속) 변수에 대한 (예측)의 정확성을 (높이는)

7. ⑦ 대신 쓸 수 있는 말을 2가지 이상 서술하세요.

불일치, 차이

8. ⑧을 나타내는 다른 말을 서술하세요.

좌표평면

9. ⑨를 보고 '2차원 평면'에서 다음 수치들에 해당하는 것을 서술하세요.

종속 변수의 실제 수치	점의 세로축 값
종속 변수의 예측치	점의 가로축 값에서의 직선의 세로축 값

(※ 본문에서 가로축 값, 세로축 값이라는 말을 사용해서 예시답안도 그를 활용해 작성했으나, x값, y값이라는 단어를 사용하셔도 상관없습니다.)

10. 다음은 ⑩의 이유를 서술한 것입니다. 빈 칸에 들어갈 알맞은 말을 고르거나 빈 칸을 채우세요.

회귀 계수는 다음 수치의 평균치로 볼 수 있다.

종속 변수의 변화량/설명 변수의 변화량
그런데 2차원 평면에서 편차 제곱의 합이 가장 작은 직선은, (종속) 변수와 (설명) 변수의 (평균)적인 추세를 나타낸다고 볼 수 있다. 그리고 직선의 기울기를 나타내는 식은 다

음과 같다.

세로축 값의 변화량/가로축 값의 변화량

또한 2차원 평면에서 가로축에는 (설명) 변수를, 세로축에는 (종속) 변수를 나타낸다. 따라서 직선이 (종속) 변수와 (설명) 변수의 (평균)적인 추세를 나타내는 경우, 해당 직선의 기울기인 (세로축 값의 변화량/가로축 값의 변화량)은, (종속 변수의 변화량/설명 변수의 변화량)의 평균치로 볼 수 있다.

11. ⑪을 보고 다음 물음에 답하세요.

(1) '직선'이 가리키는 대상('어떤' 직선인지)을 서술하세요.

편차 제곱의 합이 가장 작은 직선

(2) ⑪의 이유를 서술하세요.

직선이 점들의 위에 위치하는 경우, 그 직선을 아래로 내려서 점들과 보다 가깝게 하면 편차 제곱의 합이 작아진다. 반대로 직선이 점들의 아래에 위치하는 경우, 그 직선을 위로 올려서 점들과 보다 가깝게 하면 편차 제곱의 합이 작아진다. 따라서 편차 제곱의 합을 최소화하는 직선은 점들의 위에 위치하거나 아래에 위치할 수 없다.

12. 다음은 ⑫에 대한 설명입니다. 빈 칸에 들어갈 알맞은 말을 고르거나 빈 칸을 채우세요.

편차의 제곱은 2차원 평면에서 점과 직선의 거리를 제곱한 값이다. 이 때 거리는, 점과 직선 간의 (최단)거리가 아니라 점에서 (수직선)을 따라 직선까지 이동한 거리를 의미한다. 그런데 특정한 수치 a 의 제곱은, 해당 수치에 (정비례)해서 커지지 않는다. a 가 1 이상이면, a 의 증가율보다 a^2 의 증가율이 더 (크다). 따라서 직선과 멀리 떨어져 있는, 즉 직선과의 (거리)가 (큰) 점들일수록 편차의 제곱의 합을 더 (많이) 늘리게 된다. 결국 최소제곱법에서는 편차의 제곱의 합을 (최소화)하는 직선을 구해야 하므로, 그를 위해서는 직선에서 멀리 떨어져 있는 점들이 많을수록 직선을 (큰) 폭으로 수정할 필요가 있다.

13. '최소 제곱법'을 사용하는 경우 ⑬과 ⑭가 무엇에 해당하는지 각각 서술하세요.

⑬	설명 변수(독립 변수)
⑭	종속 변수

14. 다음은 ⑮의 의미를 설명한 것입니다. 빈 칸을 채우세요.

- (1) 소득은 (월간) 소득으로 산정한다.
(2) 2차원 평면에서의 수치 1은 (100만 원)을 나타낸다.

15. ⑯을 보고 다음 물음에 답하세요.

(1) '최소 제곱법'으로 구한 직선'이 어떤 직선인지 서술하세요.

편차 제공의 합이 가장 작은 직선

(2) ⑩의 이유를 1, 2문단에서 찾아서 서술하세요.

편차 제공의 합이 가장 작은 직선의 기울기가 설명 변수와 종속 변수의 관계를 나타내는 회귀 계수인데, 회귀 계수는 설명 변수가 변화할 때 종속 변수가 변화하는 평균치이기 때문이다.

16. ⑰을 보고 제기될 수 있는 의문을 서술하세요.

소득이 전혀 없는 사람은 물건을 구매할 돈이 없을 텐데, 식료품에 대한 지출을 할 수 있을까?

17. ⑱의 의미를 서술하세요.

식료품에 대한 지출을 할 수 있다(또는 식료품 지출액이 존재할 수 있다).

18. 다음은 ㉔의 의미를 정리한 것입니다. 빈 칸에 들어갈 알맞은 말을 고르세요.

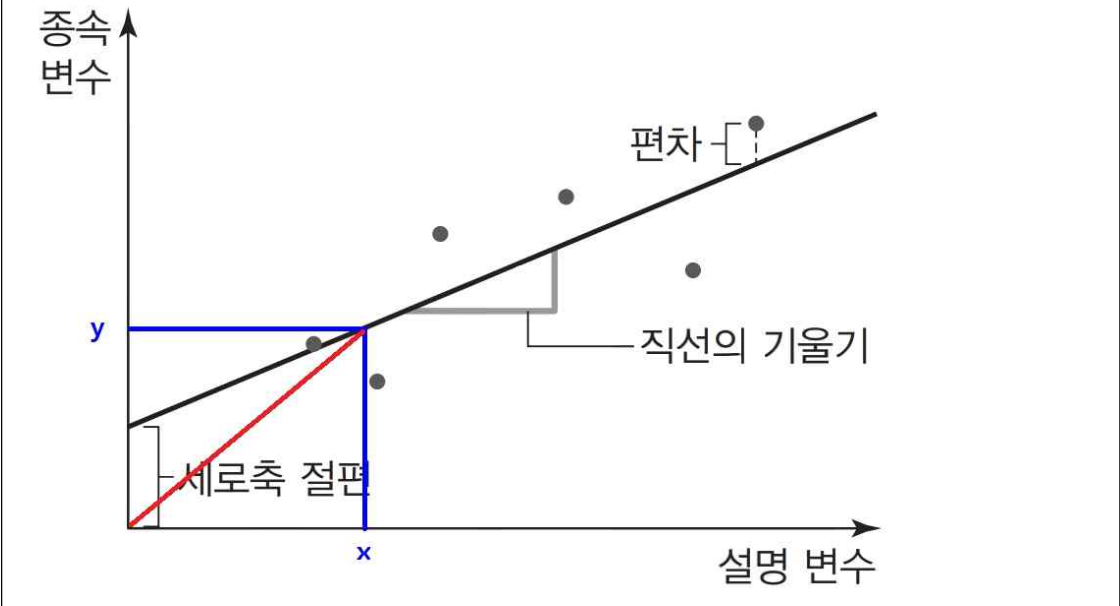
소득이 급격하게 (줄어들더라도), 식료품 지출액(은) 그만큼 급격하게 (줄어들지 않는다).

19. ㉔을 그림을 이용하여 설명하세요(본문의 그림이 ㉔에 제시된 그림에 해당하므로 참고하시면 됩니다.).

소득에서 식료품 지출액이 차지하는 비율은 다음과 같다.

식료품 지출액/소득

설명 변수를 가계의 소득, 종속 변수를 식료품 지출액으로 설정하여 최소제곱법을 적용하는 경우, 해당 비율은 다음과 같은 2차원 평면에서 빨간색 직선의 기울기와 같다(물론 소득이 설명 변수이므로 이때의 해당 비율은 소득의 크기, 즉 x의 값에 의해 결정된다.)



그런데 x값이 작아질수록 빨간색 직선의 기울기는 커진다. 즉, 소득에서 식료품 지출액이 차지하는 비율은 소득이 적을수록 높게 나타난다. 이는 다음과 같이 설명할 수 있다. y절편이 0보다 크므로, x값이 0에 가까운 경우에도 y값은 일정 수준 이상이므로 이 경우 빨간색 직선의 기울기는 매우 크다.

20. 다음은 '수요의 소득탄력성'에 대한 설명입니다.

수요의 소득탄력성은 소득이 변화할 때 특정 재화의 수요량이 얼마나 탄력적으로(민감하게) 변하는지를 나타내는 지표이다. 수식은 다음과 같다.

$$\text{수요의 소득탄력성} = \text{수요량의 변화율} / \text{소득의 변화율}$$

위 내용과 ㉔을 보고 다음 빈 칸에 들어갈 알맞은 말을 고르세요.

식료품은 문화 예술 분야의 상품에 비해 수요의 소득탄력성이 (작다).

21. ㉔의 내용을 서술하세요.

소득 대비 식료품 지출액의 비율이 높을수록, 삶의 질이 낮다.

22. ㉔가 ㉔에 해당하는 이유를 서술하세요.

복지 정책을 다음과 같이 시행할 수 있다. 우선적으로 최저 생계비를 정해놓는다. 실제 소득이 최저 생계비에 미달하는 가구(또는 사람)가 있다면, 그들은 인간다운 생활을 하기 어려운 상황에 놓인 것으로 볼 수 있다. 따라서 그들에 대해 우선적으로 생계비를 지원한다.

23. ㉔를 ㉔와 같이 산정하는 이유(근거)를 추론하여 서술하세요(글에 있는 정보만으로는 정확한 이유를 도출하기는 어려울 것입니다. 여러분이 생각하는 적절한 이유를 서술하시면 됩니다).

저소득층 가계의 소득에서 식료품 지출액이 차지하는 비율이 1/3 정도 된다면, 저소득층 가계의 소득은 저소득층 가계의 식료품 지출액의 약 3배이다. 그런데 저소득층은 일반적으로 최소한으로 필요한 식료품 지출액보다 약간 더 많은 금액을 식료품 지출액으로 사용할 것이다. 그렇다면 최소한으로 필요한 식료품 지출액의 3배는 저소득층의 일반적인 소득보다 약간 낮은 소득으로 볼 수 있고, 이는 인간다운 생활을 하기 위한 최소한의 소득에 근접한 수치로 판단할 수 있다.

24. 다음은 ㉔을 이용한 ㉔과 ㉔에 대한 판단입니다. 빈 칸에 들어갈 알맞은 말을 고르거나 빈 칸을 채우세요.

소득에서 식료품 지출액이 차지하는 비율이 지나치게 높은 경우, 소득이 너무 (낮거나) (즉, (충분)하지 않거나), 식료품 가격이 너무 (비싼) 것으로 볼 수 있다. 물론 둘 다일 수도 있다.