

변량

변량은 점수, 시간 같은 여러 자료를 수량으로 나타낸 것을 말해요. 그냥 자료를 쪽 적어놓은 거로 생각하면 쉬워요.

위에서는 수학 점수 92, 84, 88, 76, 96, ... 이렇게 쪽 쓰여 있는 게 변량이에요.

계급

계급은 변량을 일정한 구간으로 나눈 구간이에요.

70점대 몇 명, 80점대 몇 명 ... 이런 식으로 점수대별로 학생 수를 알아보려면 어떻게 했죠? 70 ~ 80, 80 ~ 90, 90 ~ 100 이렇게 점수를 나눴잖아요. 이렇게 점수별로 나누어 놓은 구간이 계급이에요. 위의 표에서 왼쪽에 있는 게 계급이에요.

계급의 크기라는 용어도 있어요. 계급의 크기는 계급의 간격(너비)을 말해요.

위 예에서 70 ~ 80이라는 계급이 있었어요. 여기서 계급의 크기는 10이에요. 70과 80 사이는 10의 차이가 있잖아요.

계급의 크기 = (계급의 큰 쪽 끝값) - (계급의 작은 쪽 끝값)

중요한 건 계급의 크기는 모두 같다는 거예요. 70 ~ 80이었으면 그다음 계급은 80 ~ 90이 되어야 해요. 80 ~ 85 이렇게 계급을 나누면 안돼요.

도수

도수는 각 계급에 속하는 변량의 개수예요.

60 ~ 70점에 해당하는 점수는 68점 하나네요. 70 ~ 80점에 해당하는 점수는 72, 76점으로 두 명이에요. 80 ~ 90점에 해당하는 점수는 84, 86, 84 세 명이고, 90 ~ 100점에 해당하는 점수는 92, 96, 92, 96 네 명이에요.

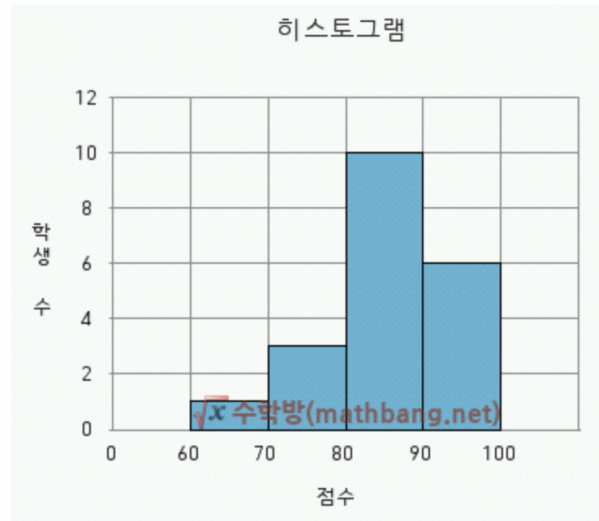
같은 값이 있어도 하나로 세지 않고 각각을 따로 세요.

여기서 60 ~ 70에 해당하는 점수가 하나니까 도수는 1, 70 ~ 80에 해당하는 점수는 두 개니까 도수가 2이고, 80 ~ 90에 해당하는 점수가 세 개니까 도수는 3, 90 ~ 100에 해당하는 점수는 네 개니까 도수가 4이에요. 위 표에서 오른쪽에 있는 게 도수지요.

즉 어떤 계급에 해당하는 자료가 몇 개인가가 바로 도수예요.

도수분포표

마지막으로 **도수분포표**는 주어진 전체 자료를 몇 개의 계급으로 나누고 각 계급에 속하는 도수를 조사하여 나타낸 표예요. 그러니까 위 표가 바로 도수분포표예요.



히스토그램 그리는 방법

위에서 설명한 것처럼 히스토그램의 가로축에는 도수분포표에서의 계급의 양 끝값을, 세로축에는 도수를 써요. 눈금과 눈금 사이가 아닌 눈금선이 있는 부분에 계급의 양 끝값과 도수를 써야 해요.

그리고 실제 사용하는 계급 앞과 뒤에 한 칸씩을 더 만드세요.

각 계급을 가로로, 도수를 세로로 하는 직사각형을 그려요. 주의할 건 눈금에 다 채워서 그려야 해요. 옆의 직사각형과 바로 붙도록 그립니다. 아래 그림처럼 직사각형 사이가 서로 떨어져 있으면 안 돼요. 앞의 그림은 제대로 된 히스토그램, 아래 그림은 잘못된 히스토그램입니다.



↑ 아래 그래프도
와장가지!!



히스토그램과 막대그래프의 차이

히스토그램은 얼핏 보면 막대그래프와 닮았어요. 그런데 왜 막대그래프가 아닌 히스토그램을 그릴까요?

막대그래프는 보통 연속되지 않는 자료들을 그래프로 그릴 때 사용해요. 사과는 몇 개, 수박은 몇 개, 이럴 때 사용하죠. 수박과 사과는 서로 연결할 수 없잖아요.

둘 다 시리클이 '도수(데이터의 개수)'를 의미한다.



이산형 변수를 그래프로 나타낼 때, '막대그래프'를 사용

광고 박효신의 음악을 소장하는 새로운 방법! 저작권 구매

뮤지코인

자세히 알아보기

히스토그램은 60 ~ 70점, 70 ~ 80점, ... 처럼 서로 연속된 자료를 나타낼 때 사용합니다. 첫 번째 계급의 끝값인 70점과 두 번째 계급의 70점이 서로 연결되잖아요.

이산형 변수를 그래프로 나타낼 때, '히스토그램'을 사용.

그래프를 보면 가장 눈에 띄는 게 있어요. 히스토그램은 막대가 서로 붙어 있고, 막대그래프는 떨어져 있어요. 위에서 설명한 연속이나 연속하지 않느냐의 차이 때문에 생기는 건데요. 60 ~ 70, 70 ~ 80은 연속하니까 죽 붙여서 그려야 하는 거지요.