

론적 결과와 거의 완벽히 일치했습니다.

그런데 이런 방식을 항상 사용할 수 있을까요? 만약 우리가 추정하고자 하는 값이 정규분포, 아니 알려진 그 어떤 확률분포도 따르지 않는다면 어떻게 할까요? 이럴 때는 중심극한정리를 사용할 수 없습니다. 그런데도 모수치를 추정하는 방법이 있을까요? 당연히 있습니다. 그중 하나가 다음 장에서 살펴볼 ‘부트스트랩’입니다. 이 기법은 현대에 와서 매우 강력해진 컴퓨터 시뮬레이션의 힘을 빌려 수학 공식 없이도 무식한(?) 컴퓨터의 힘으로 추정을 밀어붙이는 방법입니다. 워낙 강력하여 우리가 아는 대부분의 통계적 추정 문제에 대해 적용할 수 있습니다. 다음 장에서 부트스트랩에 대해 알아보겠습니다.



### 예측정확도의 역설

이번에 이야기할 주제는 확률이나 통계 분야에서 널리 알려진 진짜 ‘역설’은 아니지만 생각해볼 만한 지점이 있는 문제입니다. 바로 데이터과학 및 기계학습 등에서 이야기하는 ‘예측정확도(predictive accuracy)’에 관한 이야기입니다. 이를테면 어떤 기업에서 질병 진단 인공지능 알고리즘을 개발했는데 예측정확도가 99%에 달한다고 합시다. 그러면 이 알고리즘은 좋은 알고리즘일까요? 여러분이 만약 여윳돈을 좀 갖고 있다면, 이 회사의 주식을 사야 할까요? 정답은 ‘그때그때 달라요’입니다. 왜 그런지 지금부터 생각을 좀 해봅시다.

여기서 문제는 정확도를 어떻게 정의할 것이냐 하는 것입니다. 알기 쉽게, 질병이 있는 사람은 있다고 분류하고, 없는 사람은 없다고 분류한 비율을 정확도라고 칩시다. 얼핏 보기에는 별문제 없어 보이는 정확도입니다. 그런데 여기서 문제가 생깁니다. 우리가 진단하려고 하는 병이 사실 꽤 희귀한 병이어서, 전 인구의 0.5%만이 갖고 있는 질병이라고 해봅시다. 이런 상황에서 예측정확도 99%를 달

성했다는 것은 무슨 뜻일까요? 극단적인 경우, 질병이 없는 99.5%의 사람들 중 99%를 ‘질병 없음’으로 분류하고, 나머지 질병이 없는 0.5%와 질병이 있는 0.5%를 잘못 분류해도 이 분류기는 ‘정확도 99%’를 달성하게 될 것입니다. 더 심한 경우를 생각해볼까요? 어떤 인공지능 분류 알고리즘은 사실 기업에서 만든 사기 제품입니다. 이 분류기가 하는 것은 환자에 대한 어떤 정보가 들어오든 몽땅 다 ‘질병 없음’으로 분류하는 것입니다. 정보처리가 전혀 필요 없죠. 이 분류기의 정확도는 어떻게 될까요? 99.5%의 질병 없는 사람들을 음성으로, 나머지 0.5%의 질병 있는 사람들도 음성으로 분류하여 결과적으로 99.5%의 예측정확도를 보일 것입니다. 이런 것을 언론에 홍보해도 사실 엄밀히 말해 틀린 것은 아닙니다. 다만 우리가 흔히 생각하는 ‘정확도’의 정의가 이 상황에서 별 쓸모 없는 것일 뿐입니다.

일반적으로, 어떤 인공지능 알고리즘이 있을 때 그것이 좋냐 나쁘냐 하는 것은 정확도의 절대적 수치만 놓고 판단하지 않습니다. 비슷한 작업을 수행하는 다른 도구들, 이를테면 사람이나 다른 유형의 분류 알고리즘과 비교해야 합니다. 그리고 정확도를 정의하는 방법도 위와는 다른 방식으로 해야 합니다. 예를 들어 기계학습 분야에는 ‘F1 점수’라는 것이 있는데, 이것은 위에서 말한 음성과 양성에 대한 진단 결과 모두를 감안한 점수를 산출해줍니다(자세한 수식은 생략하겠습니다). 이런 방식으로 앞서 말한 ‘예측정확도’의 맹점을 보완할 수 있습니다.

굳이 사기를 치지 않더라도, 있는 자료를 보여주는 방식만 잘 바꾸어도 사람들을 속이는 것은 가능합니다. 사실 이런 유형의 기만은 데이터 시각화에서 자주 등장합니다. 예를 들어 세로축을 임의로 자른다거나, 도표 축 간격을 일정하지 않게 한다거나 하는 것들 말이죠. 하지만 숫자 자체도 예쁘게 꾸미려면 못할 것은 아닙니다. 세상에는 이런 것을 이용하여 사람들을 현혹해서 돈을 벌려는 사람이나 기업들이 존재하는데, 이런 것들까지 신경 쓰며 살아야 한다니 역시 사는 건 만만치 않은 일인 것 같습니다.



## 확률의 세계와 인간의 편향

마음이 어떻게 외부 정보를 처리하는지를 연구하는 학문을 ‘인지심리학(cognitive psychology)’이라고 부릅니다. 인지심리학의 주된 연구대상 중 하나는 인간이 확률을 어떻게 처리하는가 하는 것입니다. 확률이라는 게 대단한 것 같지만 사실 따지고 보면 별게 아닙니다. 어떤 말이 그럴듯하다고 생각하느냐, 어떤 진술이 참이라는 것을 얼마나 믿느냐 하는 것도 확률의 렌즈로 바라볼 수 있습니다. 사실 확률을 아예 ‘믿음의 정도(degree of belief)’로 정의하는 통계학 유파가 따로 존재하기도 합니다. 이를 ‘베이즈 통계(Bayesian statistics)’라 부릅니다. 여기서 이에 대해 자세하게 이야기 하지는 않겠지만, 아무튼 인간의 인지과정과 확률 사이에는 밀접한 관계가 있습니다.

심리학 연구에서 일관적으로 발견되는 것은 사람들이 확률 판단을 포함하여 논리적 오류를 흔히 저지른다는 것입니다. 이를 ‘인지적 편향(cognitive bias)’이라 부릅니다. 인지적 편향 중 대표 격인 ‘확증편향(confirmation bias)’이라는 말은 10년 전쯤 까지만 해도 그리 대중적인 말이 아니었는데, 요즘에는 누구나 사용하는 말이 됐 습니다. 주로 자신과 생각이 다른 사람들을 공격하는 데 쓰이는 것 같다는 게 옥 에 티이긴 하지만요. 아무튼 확률에 관련된 인지적 편향으로는 ‘결합오류(conjunction fallacy)’라는 게 있습니다. 긴말할 것 없이 먼저 다음 묘사를 읽어봅시다.

김수진 씨는 31세의 여성으로, 서울 소재 4년제 상위권 대학을 졸업하고 현재 IT기업에서 근무하고 있다. 소득은 그리 많지는 않지만 동년배들에 비하면 꽤 괜찮은 편이다. 수진 씨는 현재 독신이며 각종 진보 의제에 관심이 많다. 진보 정당의 당원으로 가입해 있으며, 가끔 집회 현장에 직접 나가기도 한다. 몇 해 전 있었던 강남역 살인사건 당시에도 수진 씨는 집 회 현장에서 목소리를 높였다.

이제 질문입니다. 다음의 둘 중 무엇이 더 가능성이 높다고 생각하십니까?

① 수진 씨는 사무직이다.

② 수진 씨는 사무직이면서 페미니스트이다.

사실 정답은 정해져 있습니다. 심지어 김수진 씨에 대한 그 어떤 추가 정보 없이도 말입니다. 답은 ①입니다. 왜냐하면 ②를 만족하면 반드시 ①도 만족하지만, ①을 만족한다고 해서 ②를 만족한다는 보장은 없기 때문입니다. 다시 말해 ①을 만족하는 가상적인 김수진의 집합이 ②를 만족하는 가상적인 김수진의 집합을 포함한다는 것입니다. 따라서 ①의 확률이 무조건 더 ②의 확률과 같거나 더 높게 되어 있습니다. 이는 논리적으로 참입니다.

그런데 이 문제를 사람들에게 주면 ②를 고르는 사람이 꽤 많습니다. 글에서 묘사된 김수진 씨의 모습이 ①보다는 ②에 더 ‘가까워’ 보이기 때문입니다. 이를 심리학에서는 ‘대표성 어림법(representativeness heuristic)’이라 부릅니다. 사람들은 어떤 사건의 확률을 평가할 때, 그 사건이 얼마나 대표성이 있는지에 기반을 둔다는 것입니다. 실제로 그 사건이 얼마나 그럴듯한지가 아니라 말이죠. 그러니까 ②가 ①에 비해 김수진 씨를 더 잘 대표한다고 생각한다는 것입니다.

이와 같이 사람들이 확률 관련 추론을 할 때 각종 논리적 오류를 흔히 저지른다는 것은 인지심리학 분야에서 매우 잘 알려져 있는 이야기입니다. 수학자가 몬티홀 문제에서 오류를 저질렀다는 일화도 있었죠. 역시 애초부터 확률은 사람이 이해하기에는 너무 난해한 대상인지도 모르겠습니다.