

병원 검사는 얼마나 정확할까?

통계학, 빅데이터를 잡다(조재근 저) 중에서

1. 민감도와 특이도

민감도(sensitivity) : 병이 있는 사람이 양성인 비율

위음성률 : 병이 있는데, 음성이 나올 비율 \Rightarrow 민감도를 a 라면, 위음성률은 $1-a$

특이도(specificity) : 병이 없는 사람이 음성이 나올 비율

위양성률 : 병이 없는데, 양성인 비율 \Rightarrow 특이도가 b 라면, 위양성률은 $1-b$

2. 양성인데, 병에 걸렸을 확률(기저율의 오류)

어떤 병의 민감도를 a , 특이도를 b 라 하고,

병에 걸리는 사람의 비율을 X 이고, N 명을 검사하였다면,

① 실제 병에 걸린 사람은 NX , 병에 걸리지 않은 사람은 $N-NX$ 임

② 검사 결과가 양성인 나오는 사람

= 병이 있으면서 양성인 사람 + 병이 없는데 양성인 사람

= $NXa + (N-NX)(1-b)$

③ 양성인데 병에 걸릴 확률

$$= \frac{\text{병에 걸린 비율}}{\text{양성이 나오는 사람}} = \frac{NX}{NXa + (N-NX)(1-b)} = \frac{X}{Xa + (1-X)(1-b)}$$

④ 기저율의 오류(base-rate fallacy) : 진단 결과 양성을 나와도 질병확률(X) 자체가 낮아서 실제로 병에 걸렸을 확률이 낮게 되는 것

\Rightarrow 희귀 질환일수록 양성인데, 병에 걸렸을 확률이 낮아짐

3. 예시

○ 40대 여성이 검사 결과 양성인면서 실제로 유방암에 걸릴 확률

- 40대 여성이 유방암에 걸릴 확률 $X = 1\% (= 0.01)$

- 유방암에 걸렸으면서 양성인 나올 확률 $a = 90\% (= 0.9)$

- 유방암에 걸리지 않았는데 양성인 나올 확률 $1-b = 9\% (= 0.09)$

(1) 계산에 의한 값

$$\text{확률 } P = \frac{X}{Xa + (1-X)(1-b)} = \frac{0.01}{0.01 \times 0.9 + 0.99 \times 0.09} = \frac{1}{9.81} \simeq 10\%$$

(2) 직관적으로 확인

100명이 있었고, 1명이 유방암에 걸렸음

유방암에 걸린 사람 1명은 양성이며, 유방암이 아닌데 양성인 나온 사람은 9명 정도

즉, 양성인 나온 사람은 대략 10명인데, 실제 걸린 사람은 1명이므로, 확률은 10%

4. 두 가지 검사를 시행하는 경우

① 두 가지 모두 양성일 때 질병으로 판단할 경우

- 질병이 있으면, 두 가지 검사 모두 양성인 나와야 함 \Rightarrow 민감도가 낮아짐

- 질병이 없으면, 한가지라도 음성이 나오면 됨 \Rightarrow 특이도는 올라감

② 두 가지 중 한 가지만 양성이면 질병으로 판단하는 경우

- 질병이 있으면, 한가지라도 양성인 나오면 됨 \Rightarrow 민감도가 올라감

- 질병이 없으면, 두가지 모두 음성이 나와야 함 \Rightarrow 특이도는 내려감

5. 결론

① 검사 결과 양성을 받으면, 병에 걸렸을 확률은 올라감

\Rightarrow 위의 유방암의 예에서 검사 전에는 1%에서 양성 후에는 10%가 됨

② 민감도가 좋은 검사에서 양성을 받았더라도 생각보다 확률이 높진 않음

③ 음성을 받았다고 하더라도 완전히 안심할 수는 없음