

## 오즈비는 언제 잘못 해석될 수 있는가? 를 읽고

2개의 범주를 갖는 독립변수  $x$ 와 2개의 범주를 갖는 반응변수  $y$ 가 있다고 가정합시다. 독립변수  $x$ 의 범주는 각각 ‘충족’, ‘미충족’이라고 하고, 반응변수  $y$ 의 범주는 각각 ‘성공’, ‘실패’입니다.

이러한 경우에  $x$ 를 충족한 사례 100개와 충족하지 못한 사례 100개에 대해  $y$ 값이 어떠한지 조사하고 이를 이차원분할표로 작성했다고 하면 다음과 같은 결과를 얻을 수 있습니다.

	성공	실패	합계
충족	$a$	$100-a$	100
미충족	$b$	$100-b$	100

오즈비는  $= \frac{a/(100-a)}{b/(100-b)}$ , 상대위험도  $= \frac{a}{b}$  로 나타낼 수 있습니다. 두 식을 연립하면

오즈비 = 상대위험도  $\times \frac{100-b}{100-a}$  임을 알 수 있습니다.

독립변수가 어떤 범주인지와 관계없이 성공과 실패가 결정되기 때문에  $a$ 와  $b$ 가 같은 값을 가지게 된다면, 이런 경우에는 오즈비와 상대위험도 역시 같은 값을 가지게 됩니다. 문제는  $a$ 와  $b$ 의 값이 달라질 때 발생합니다.

$a$ 가  $b$ 보다 큰 값을 가지게 될 때, 즉 반응변수  $x$ 를 충족하면 성공할 확률이 높아지는 경우 “오즈비>상대위험도”가 성립합니다.  $100-a$ 의 값이  $100-b$ 의 값보다 작기 때문입니다. 반대로  $a$ 가  $b$ 보다 작은 값을 가지게 될 때, 즉 반응변수  $x$ 의 충족이 성공에 방해가 되는 요인일 때는 “오즈비<상대위험도”가 성립하게 됩니다.

이 때  $b=50$ 으로 고정되어 있다고 할 때  $a=25$ 인 경우와  $a=75$ 인 경우 오즈비와 상대위험도 값은 다음과 같습니다.

$a=25, b=50$  인 경우 오즈비 = 0.333, 상대위험도 = 0.5

$a=75, b=50$  인 경우 오즈비 = 3, 상대위험도 = 1.5

이 결과값을 보면 상대위험도가 오즈비보다 직관적으로 이해하기 쉽다는 것을 알 수 있습니다. 따라서 사건이 발생할 확률이 0에 가까운 값인 경우 오즈비와 상대위험도가 근사하다는 사실과 별개로 오즈비를 사용하는 것은 신중해야 합니다.  $a=75, b=50$ 일 때 “반응변수  $X$ 가 충족되었을 때의 오즈값이  $X$ 가 미충족되었을 때의 오즈값의 3배이다”라고 해석하면 많은 사람들이 반응변수  $X$ 가 충족되면 성공할 확률이 3배 높다고 이해할 수 있습니다.

「When can odds ratios mislead?」는 이와 같은 사실을 강조하는 문헌이라고 할 수 있습니다. 저자에 따르면 오즈비는 사례-통제 연구(case-control studies)와 로지스틱 회귀 분석에서만 안전하게 사용할 수 있습니다.