

# Discrete\_hw1 : 분할표에서의 Fisher's exact test

Na SeungChan

2024-10-08

## 실험의 배경

과거 마술사 제임스 랜디는 '자칭 초능력자'들이 실제 초능력을 사용할 수 있는지 확인하기 위한 실험을 진행한 적 있었다. 이는 초능력자들이 주로 마술 트릭을 통해 자신이 초능력자인 것처럼 선전하는 것을 알기에, 그들이 마술 트릭을 사용하지 못하게 통제된 뒤에도 해당 초능력을 사용할 수 있는지 확인하여 그들이 실제 초능력자인지 검증한 것이다.

## 실험설계

이는  $2 \times 2$  차원 분할표에 표시할 수 있는 실험의 형식으로 표현할 수 있다. 구체적인 사례를 통해 확인하자. 어떤 사람이 투시 능력을 가지고 있어 사람이 신장 이식을 받았는지 여부를 알 수 있는 초능력자라고 주장한다고 가정하자. 이 사람이 다른 마술 트릭을 쓰지 못하도록 하고,  $n_{0+}$  명의 '신장 이식을 받지 않은 사람'과  $n_{1+}$  명의 '신장 이식을 받은 사람'을 데려오고, 해당 자칭 초능력자가 해당 사람들 각각을  $n_{+1}$  명의 신장 이식을 받은 사람과  $n_{+0}$  명의 받지 않은 사람으로 구분하면 된다. 이렇게 하면, '실제로 신장 이식을 받지 않았고, 자칭 초능력자도 신장 이식을 받지 않았다고 판단한 사람'의 도수인  $n_{00}$ 은 초기하분포를 따르고,  $n_{00}$ 이 결정됨으로써 분할표의 다른 값도 결정되어 Fisher's exact test의 상황을 만족하게 된다.

## 분할표의 구성과 분포

즉, 이와 같은 실험은 다음과 같은 분할표와 제약조건에 의해 분포가 결정되는 Fisher's Exact test에 해당한다. 분할표에서 X 변수는 '신장 이식을 받았는지 여부'를 나타내고, Y는 초능력자가 '해당 사람이 신장 이식을 받았다고 판단했는지 여부'를 나타낸다.

X   Y	0	1	(sum)
0	$n_{00}$	$n_{01}$	$n_{0+}$
1	$n_{10}$	$n_{11}$	$n_{1+}$
(sum)	$n_{+0}$	$n_{+1}$	$n$

이 경우,  $n_{00}$ 은 총  $n$ 개의 샘플 중  $n_{0+}$ 개의 성공이 존재하는 모집단에서  $n_{+0}$ 개의 샘플을 비복원추출한 경우 성공하는 경우의 수와 동일하다. 이는 아래와 같은 초기하분포를 따르므로, 이 분할표에서의 확률분포는 다음과 같이 모델링할 수 있다.

$$n_{00} \sim \text{HyperGeo}(n, n_{0+}, n_{+0})$$

한편,  $n_{00}$ 이 정해지면 다음과 같이 분할표의 나머지 수도 결정된다.

$$n_{01} = n_{0+} - n_{00}, n_{10} = n_{+0} - n_{00}, n_{11} = n - n_{00} - n_{01} - n_{10}$$