

1강 정보를 얻으면 확률이 바뀐다

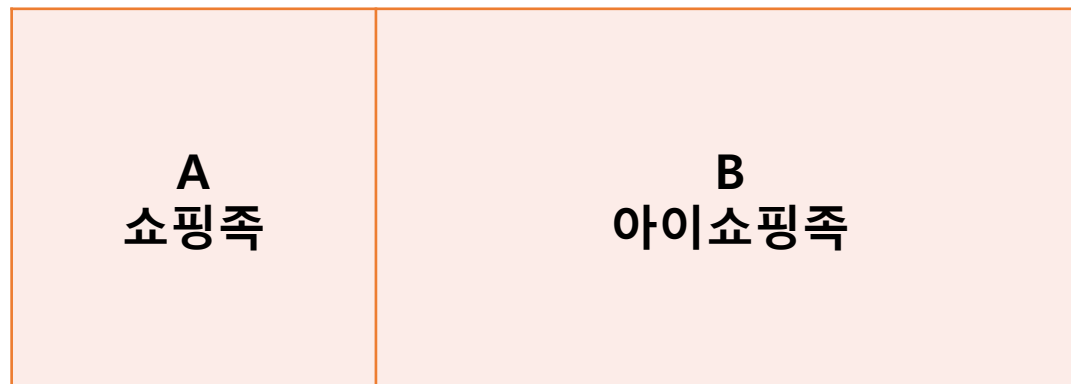
출처 : 세상에서 가장 쉬운 베イズ통계학 입문

1-1 베イズ 추정으로 쇼핑족과 아이쇼핑족을 판별한다.

- 베イズ 추정의 전형적인 사용법은 소개
- 직감적으로 손님이 타임을 꿰뚫어볼 줄 아는 점원도 많고 이것은 바로 점원의 테크닉임.
- 직감적인 판단을 수치화하여 계산할 수 있도록 메뉴얼화하면 신입사원 교육에 활용되고 **AI(인공지능)**처럼 사용할 수 있음.

1-2 [1단계] 경험에서 '사전확률'을 설정한다.

- 추측을 위해 가장 먼저 해야 할 일은 손님의 두가지 타임(쇼핑족과 아이쇼핑족)에 대한 비율을 각각 몇인지 수치 배정
- 이 타입에 대한 확률(비율) => 베이즈통계학에서는 **사전확률**
- 쇼핑족 확률 0.2, 아이쇼핑족 확률 0.8 <= 경험적 수치
- 사전 분포를 직사각형으로 분할,
- 확률은 전부 더해서 **1**이 되도록 설정 => **정규화 조건**



<= 가능세계

1-3 [2단계] 타입별로 '말거는' 행동을 하는 '조건부 확률'을 설정한다.(1/2)

- 손님의 타입에 따라 어느 정도의 확률로 점원에게 "말걸기" 행동을 하는가를 설정
- "타입의 차이에 의거한 행동의 확률 " 은 경험, 실증, 실험에 기반한 수치

타입	말을 걸 확률	말을 걸지 않을 확률	
쇼핑족	0.9	0.1	-> 1
아이쇼핑족	0.3	0.7	-> 1
	1.2	0.8	

- 조건부 확률이며, 타입을 한정한 경우 각 행동의 확률임.

1-3 [2단계] 타입별로 '말거는' 행동을 하는 '조건부 확률'을 설정한다.(2/2)

- 네 개로 분기된 세계

		쇼핑족 0.2	아이쇼핑족 0.8		
말을 건다	0.9	쇼핑족이 말을 건다.	아이쇼핑족이 말을 건다.		0.3
			아이쇼핑족이 말을 걸지 않는다.		0.7
말을 걸지 않는다.	0.1				

↑ 쇼핑족이 말을 걸지 않는다.

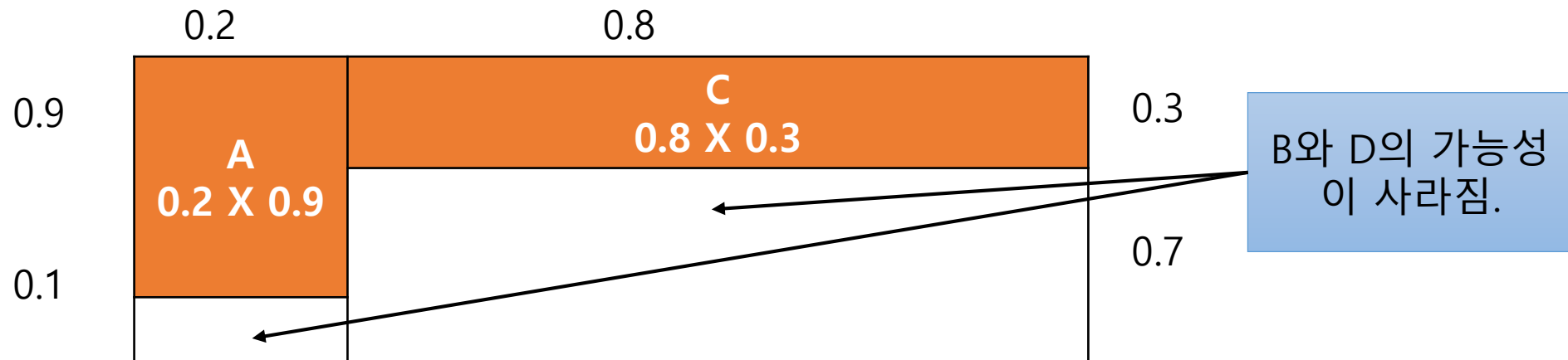
A 0.2×0.9 $= 0.18$	C $0.8 \times 0.3 = 0.24$
	D $0.8 \times 0.7 = 0.56$

↑ B $0.2 \times 0.1 = 0.02$

$$A + B + C + D = 1$$

1-4 [3단계] 관측한 행동에서 "가능성이 사라진 세계"를 제거한다.

- 점원이 당신이 "손님이 말을 걸었다"는 현실에 직면 <= 추가적인 정보 획득



1-5 [4단계]'쇼핑족'의 '베이지 역확률'을 구한다. (1/2)

- 말을 걸지 않는 세계의 소멸

쇼핑족

아이쇼핑족

말을 건다.

	0.24
0.18	

말을 건다

- 정규화 조건을 회복시켜 사후확률(=베이지역확률)을 구한다.

쇼핑족

아이쇼핑족

말을 건다.

$0.18 / (0.18 + 0.24) = 3/7$	$0.24 / (0.18 + 0.24) = 4/7$
	더하니 다시 1이 됨.

말을 건다

1-5 [4단계]'쇼핑족'의 '베이지 역확률'을 구한다. (2/2)

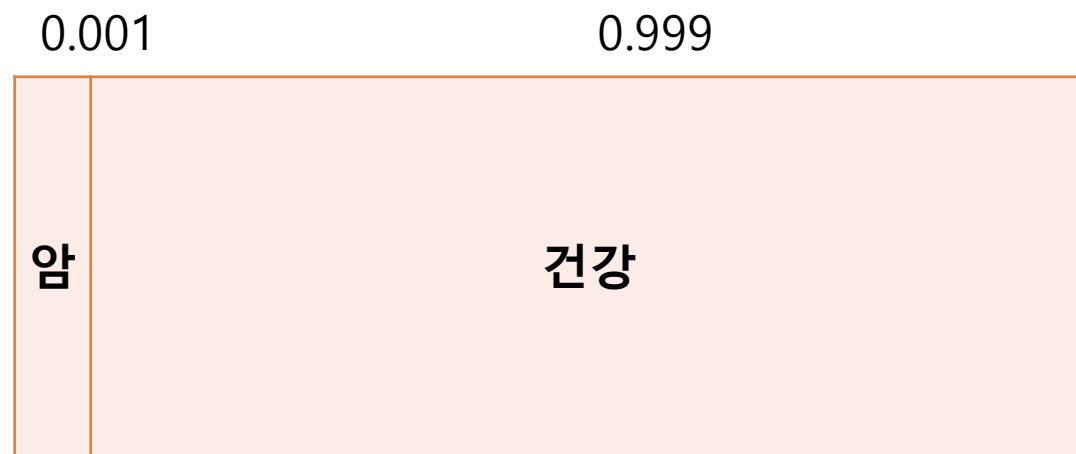
- [2단계]에서 손님의 타입이 2종류이며 각각의 타입이 '말을 건다', '말을 걸지 않는다'의 2가지 행동의 확률을 구함
=> **원인**으로부터 **결과**를 파악
- [4단계]에서 말을 건다는 행동의 결과로부터 타입이라는 원인을 거슬러 올라감.
=> **결과**으로 부터 **원인**의 확률을 구함 => 역확률

2강 베이지 추정은 때로
직감에 크게 반한다.

2-1 암에 걸려 있을 확률을 계산한다.

- 어느 특정 암에 걸릴 확률이 **0.1%**(0.001) 이라고 하자.
- 이 암에 걸려 있다면, **95%**의 확률로 양성진단을 받고
- 건강한 사람을 양성으로 오진을 받을 확률은 **2%**(0.002)인 검사가 있다.
- 검사 결과 양성 판정이 나왔다고 하자.
- 이때 당신은 자신이 그 암에 걸려 있을 확률이 95%라고 판단해야 할까 ??
- 답은 차차 설명함..~~~~~

2-2 의료데이터를 근거로 '사전확률'을 설정한다.



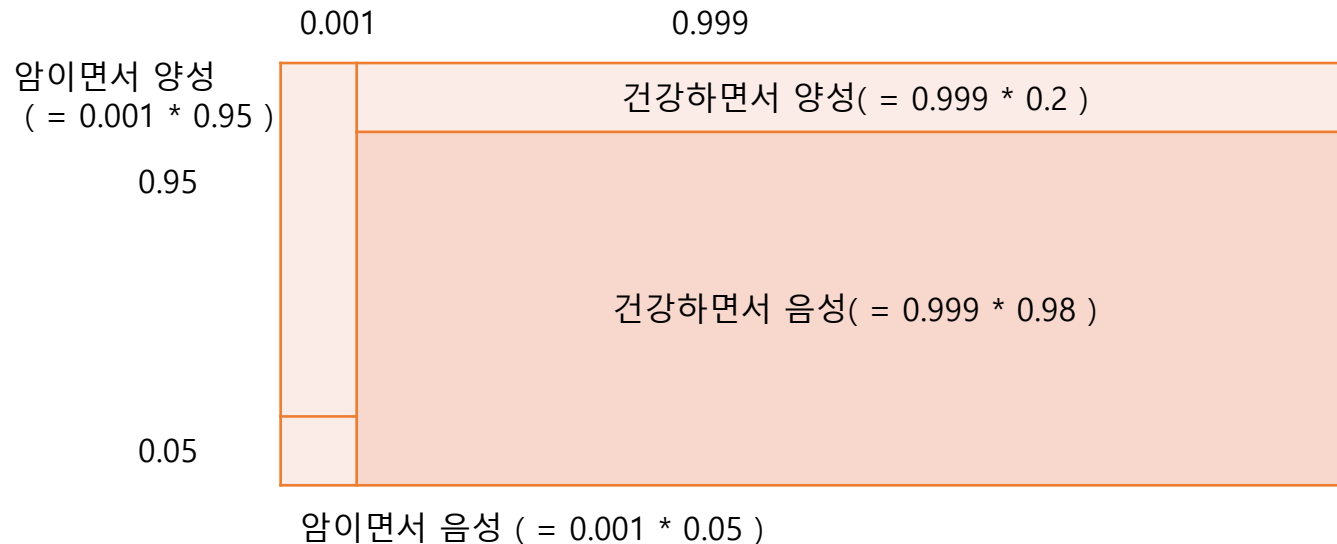
1000명 중 1명이
암에 걸려 있음.

2-3 검사의 정밀도를 근거로 '조건부 확률'을 설정한다.

- 검사 정밀도에 따른 조건부 확률

타입	양성일 확률	음성일 확률
암 환자	0.95	0.05
건강한 사람	0.02	0.98

- 네 개로 분기된 세계 각각의 확률



2-4 검사결과가 양성이므로 "일어날 가능성이 없는 세계"를 소거한다.

- 정보에 따라 가능성이 한정됨.

	암	건강
양성	0.095 %	1.998 %
		음성의 세계를 소거

2-5 당신이 암일 것이라는 "베イズ 역확률"을 구한다.

- 정규화에 따라 사후 확률을 구한다.

	암	건강
양성	0.095 %	1.998 %

- 양성이라는 검사 결과를 받았을때, 당신의 암에 걸려 있을 사후 확률을 얼마인가 ????