12장 베이즈 추정에서는 정보를 순차적으로 사용할 수 있다.

출처 : 세상에서가장 쉬운 베이즈 통계학 입문

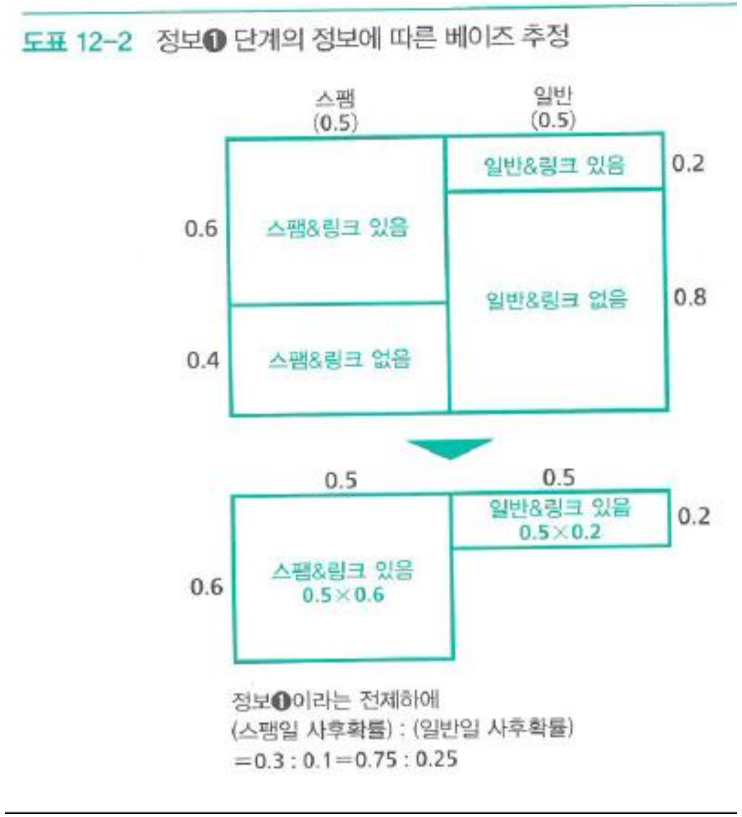
12-1 베이즈 추정에서는 이전 정보를 잊어도 앞뒤가 들어맞는다.

도표 12-1 두 가지 정보에 의한 베이즈 추정

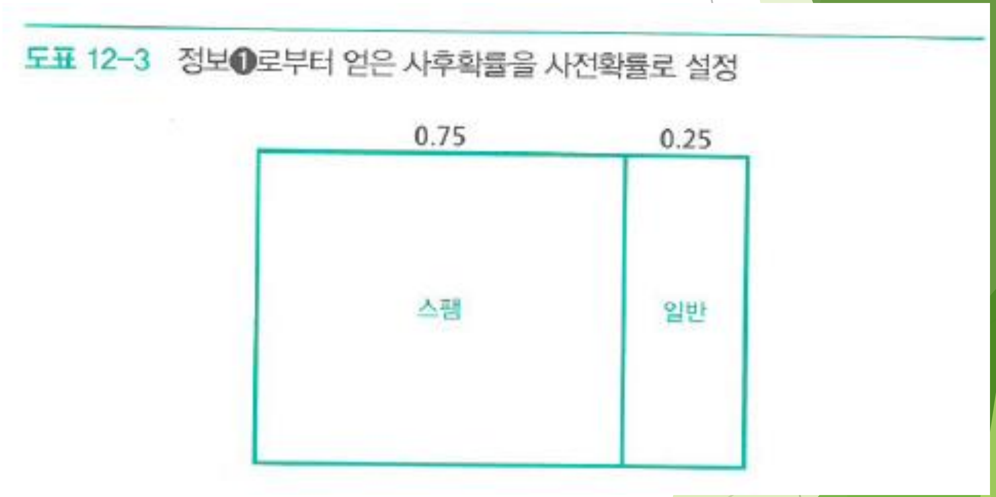
정보1 에서 타입에 대한 확률을 개정하면, 정보2 를 사용할때 앞의 정보1은 잊어도 된다.

=> 축자합리성

12-2 정보1 로부터 얻은 사후확률을 ‘사전’ 확률로 설정한다.



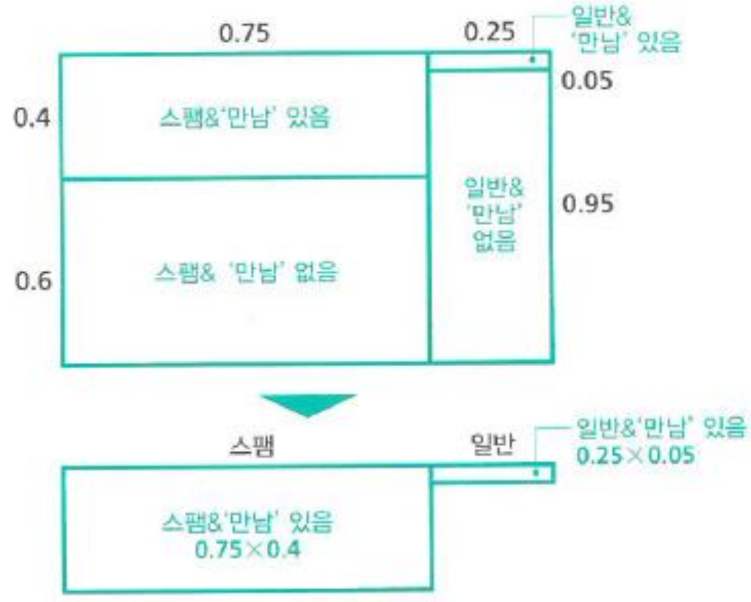
사후확률을
사전확률로
재설정



“이유는 잊어버렸지만, 사전확률이 그렇게 설정되어 있다”고 생각하는 것과 같음.

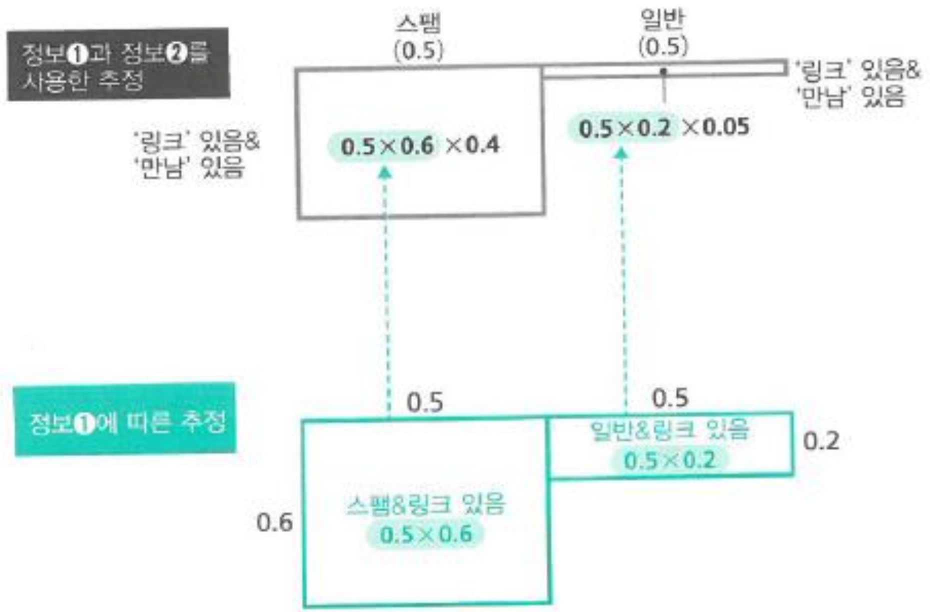
12-3 정보2 를 사용하여 베이즈 갱신을 한다.

도표 12-4 정보❷를 사용한 베이즈 추정에 의한 사후확률



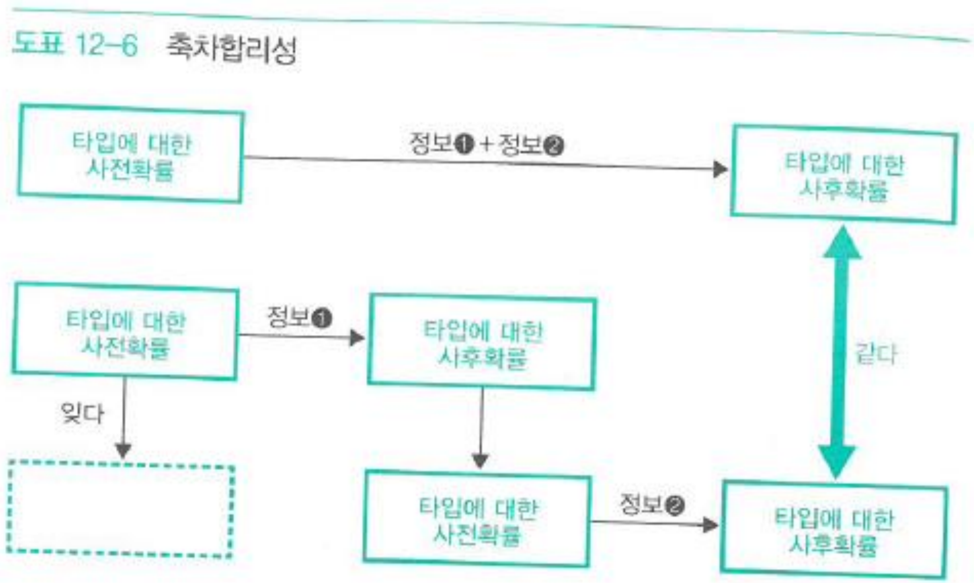
(스팸메일일 사후확률):(일반메일일 사후확률)
= 0.75 x 0.4 : 0.25 x 0.05
= 3 x 8 : 1 x 1
= 12 : 1
= 12/25 : 1/15

도표 12-5 두 가지 정보에 따른 개정과 축차적 개정이 일치하는 이유



"정보1 로부터 수정한 사후확률을 사전확률로 사용하고 여기에 정보2를 결합하여 구한 사후확률"과 "정보1과 정보2를 한번에 사용해서 구한 사후확률이 일치"

12-4 베이즈 추정은 인간다운 추정이다.



이전에 사용한 정보는 잊어도 관계없이, 계산된 사후확률을 사전확률인양 취급하여 새로이 추정을 해도 결과는 달라지지 않음.

방대한 정보를 사용하여 확률적 추측할때, 한번 사용한 정보는 버려도 현재의 추정에 완전히 반영하면 매우 효율적임

이것은 일종의 '**학습기능**' 이라고 할 수 있음.
베이즈 추정은 정보를 입수하면 자동적으로 똑똑해지는 기능은 갖춤.

요약

① 두 가지 정보를 한꺼번에 사용해서 구한 사후확률과 첫번째 정보로 얻은 사후확률을 사전확률로 재설정하여 두 번째 정보를 이용해 개정한 사후확률은 항상 일치한다.

② 1의 성질을 축차합리성이라 부른다.

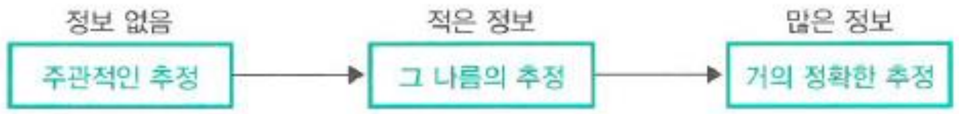
③ 축차합리성은 학습 기능의 일종으로 간주할 수 있다.

④ 베이즈 추정에서 일단 추측에 사용한 정보는 버려도 문제되지 않는다.

13장 베이즈 추정은 정보를 얻을수록 더 정확해진다

13-1 '적당적당'한 추측에서 '더 정확한' 추정으로 만들려면

도표 13-1 정보가 많으면 많을수록 더 정확한 추정이 이루어진다



베이즈 추정의 학습기능에는 "정보가 많아질수록 더 정확한 추정을 한다" 는 성질이 있음.

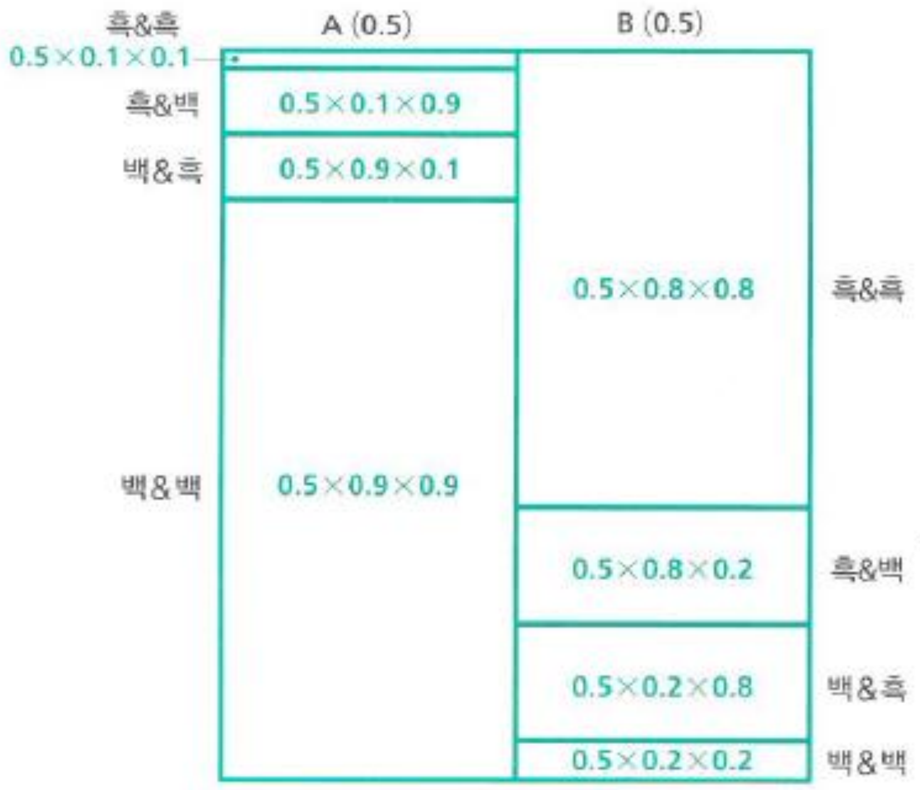
13-2 단지 문제에서 공을 두 개 꺼낸다.

문제설정

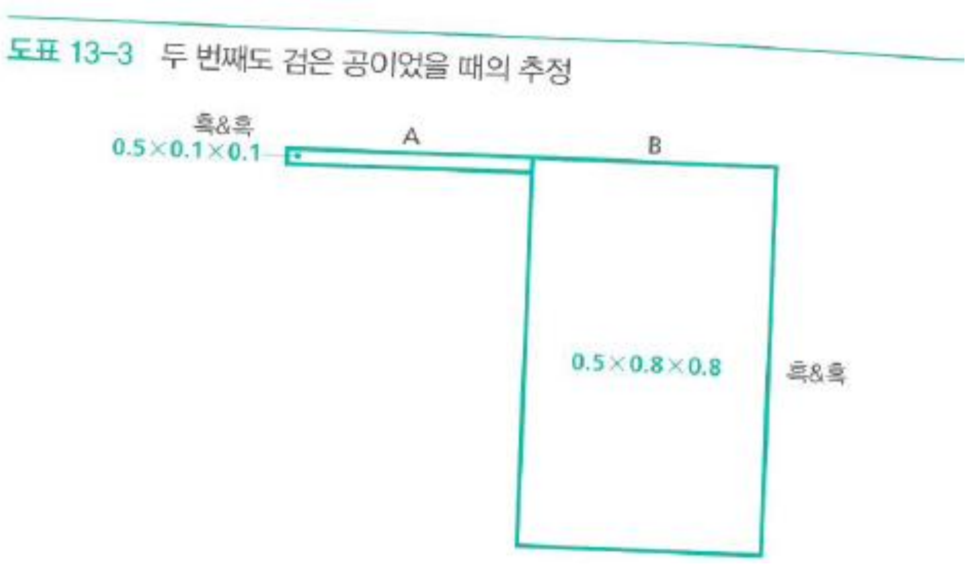
동일한 모양의 항아리 2개 있음.
항아리 A에는 흰공 9개, 검은공 1개
항아리 B에는 흰공 2개, 검은공 8개

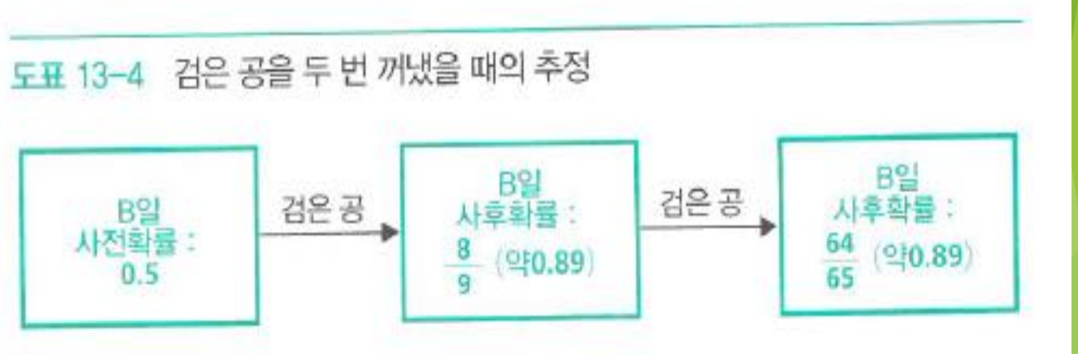
각각의 항아리에서 처음 꺼낸 공
공을 다시 단지에 넣고 새로이 공을
한 개 뽑은 경우의 추정해보자.

도표 13-2 두 가지 정보로 인해 세계는 여덟 개로 나뉜다.



13-3 두 번째도 검은 공이었을때의 추정





(흑&흑일때 A의 사후확률) : (흑&흑일때 B의 사후확률)
= 0.5 x 0.1 x 0.1 : 0.5 x 0.8 x 0.8
= 0.01 : 0.64
= 1 : 64
= 1/65 : 65/65

B일 사후확률이 0.89에서 0.98로 높아지므로, 항아리 B일 가능성이 한층 높아짐.

13-4 첫번째는 검은공, 두번째가 흰공이었을때의 추정

도표 13-5 두 번째가 흰 공이었을 때의 추정

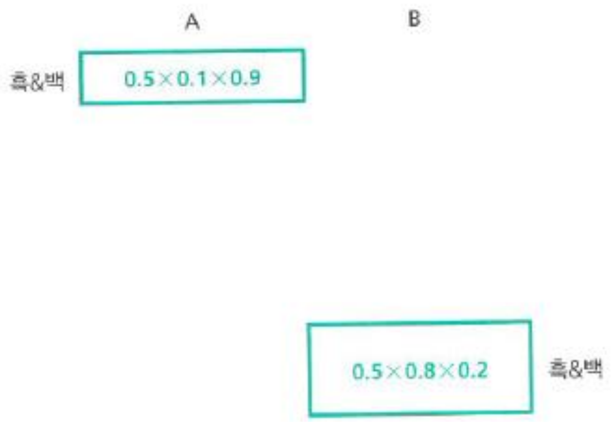
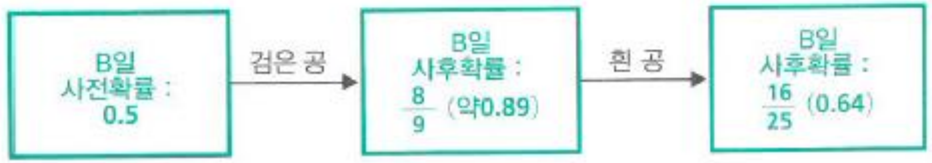


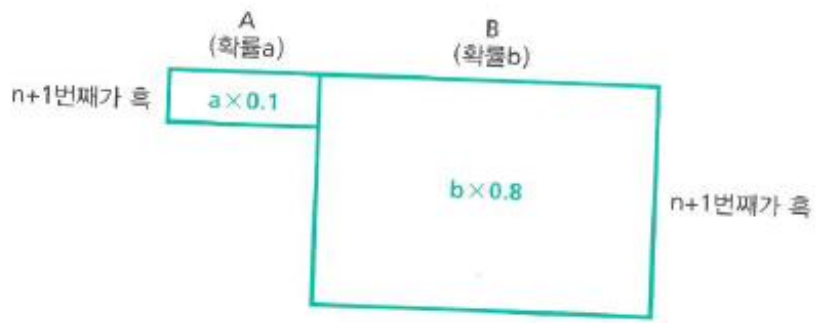
도표 13-6 첫 번째에 검은 공, 두 번째에 흰 공을 뽑았을 때의 추정



13-5 최신 관측결과에 따라 결론이 달라진다.

도표 13-7 정보로부터 추정 결과가 어느 쪽으로 기우는가

도표 13-8 n + 1번째가 검은 공이었을 때의 변화



n + 1번째 공을 관측한 뒤의 사후확률을 a’ , b’라고 하면,

a’ : b’
= a x 0.1 : b x 0.8
= a : 8b
a’는 a보다 작아지고, b’는 b 보다 커짐.

13-6 여러 번 관측할수록 추측은 진실에 가까워진다.

도표 13-9 검은 공의 관측 횟수와 사후확률과 발생확률

검은 공 횟수	0	1	2	3	4
사후확률 b	$8.62 \times \frac{1}{10^{14}}$	$3.00 \times \frac{1}{10^{12}}$	$1.10 \times \frac{1}{10^{10}}$	$4.00 \times \frac{1}{10^{9}}$	$1.40 \times \frac{1}{10^{7}}$
발생확률	$1.05 \times \frac{1}{10^{14}}$	$8.00 \times \frac{1}{10^{13}}$	$3.20 \times \frac{1}{10^{11}}$	$8.00 \times \frac{1}{10^{10}}$	$1.30 \times \frac{1}{10^{8}}$

검은 공 횟수	5	6	7	8	9	10
사후확률 b	$5.22 \times \frac{1}{10^{6}}$	0.0002	0.007	0.1957	0.898	0.9968
발생확률	$1.66 \times \frac{1}{10^{7}}$	$2.00 \times \frac{1}{10^{6}}$	0.00001	0.00009	0.0005	0.002

검은 공 횟수	11	12	13	14	15	16
사후확률 b	0.9999	1	1	1	1	1
발생확률	0.0074	0.0222	0.0545	0.109	0.1746	0.2182

검은 공 횟수	17	18	19	20
사후확률 b	1	1	1	1
발생확률	0.2054	0.1369	0.0576	0.0115

공을 20회 관측했을때 검은 공이 나온 횟수에 대응하여 ‘항아리 B일 사후확률’

검은 공이 6회미만일때는 항아리B 일 사후확률이 매우 작음.

검은 공이 9회일때는 항아리B 일 사후확률은 0.898로 높아지고

10회이상일때 항아리 B라고 단정해도 크게 위험이 없음.

요약

① 베이즈 추정은 정보에 따라 판단이 흔들리는 상태를 묘사한다.

② 검은 공이 관측되면 검은 공이 많은 단지 쪽으로 판단이 기울고 흰공이 관측되면 흰공이 많은 단지로 판단이 기운다.

③ 베이즈 추정에서는 정보가 있으면 올바른 결론을 내릴 수 있다.