

CH7 주관적 확률-숫자에 대한 믿음

2021010702 진서영

1. 백워터인베스트먼트 사는 여러분의 도움이 필요합니다.

백워터인베스트먼트 사

- > 개발도상국 시장의 잘 알려지지 않은 투자처를 찾아서 수익을 내려고 하는 기업
- > 남들이 좀처럼 이해할 수 없거나 찾을 수조차 없는 투자처를 찾아냄
- > 분석가들의 전문지식에 의존함

1. 백워터인베스트먼트 사는 여러분의 도움이 필요합니다.

백워터인베스트먼트 사의 분석가?

> 엄청난 판단력 + 충분한 인맥

=> 백워터인베스트먼트사가 가지고 있는 좋은 투자건에 대한 판단에 필요한 정보 얻을 수 있음

2. 분석가들이 서로 심하게 다투고 있습니다.

백워터인베스트먼트 사의 분석가

> 분석가들 사이에서 각국의 정세에 대한 인식의 차이 ○

> 다양한 논점에 대해 인식의 차이 ○

=> 심각한 문제!

2. 분석가들이 서로 심하게 다투고 있습니다.

이 이메일에는 확률 용어가 많이 사용되고 있습니다.

From : 수석 연구 분석가 (백워터인베스트먼트 사)
To : 헤드 퍼스트
제목 : 베트남에 대해 말하고 싶은 것은

지난 6개월 동안 나는 베트남 정부가 올해 석유를 줄일 것이라는 의견을 계속해서 주장해 왔습니다. 그리고 현장의 사람들을 통해 본 것과 뉴스 보도 등 모든 정보가 그것을 증명하고 있습니다.

하지만 백워터인베스트먼트 사의 '분석' 커뮤니티의 다른 사람들은 이것을 알도 안 된다고 생각하고 있는 것 같습니다. 상사들은 저를 음모상이라고 여기고, 베트남 정부의 그런 짐새나 노력은 '거의 없다'고 말합니다. 무슨 근거로 이런 판단을 내리게 된 거죠? 정부는 분명히 외국의 투자를 장려하고 있습니다. 감세 정책을 펴고 개인 투자자들이 몰려들 것이고, 우리는 그 전에 베트남 진출을 늘릴 필요가 있습니다.

From : 경제 부문 부문, 부사장 (백워터인베스트먼트 사)
To : 헤드 퍼스트
제목 : 이 사람들은 러시아에 가본 적이 있나요?

경제 부문의 분석은 계속 널리 이용되고 있고 우리는 러시아 및 정부에 대한 분석의 질을 높이는 작업을 계속하고 있습니다. 다른 부서는 러시아의 내부 역학을 놀라운 만큼이나 무시하고 러시아가 유럽 항공을 인수할 가능성은 매우 낮고, 다음 분기의 산업에 대한 지원은 지금까지의 결정 중 가장 머뭇거리는 것 같습니다.

From : 정치 분석가 (백워터인베스트먼트 사)
To : 헤드 퍼스트
제목 : 불투명한 지역에 대한 투자 : 성명서

러시아, 인도네시아, 베트남. 백워터인베스트먼트 사의 임직원들은 이 3개국에 집착하고 있습니다. 하지만 우리의 질문에 대한 답은 너무나도 명확하지 않아요? 러시아는 지금까지와 마찬가지로 다음 분기에도 계속해서 석유에 대해 보조금을 줄 것이고, 심중팔구 다음 분기에 유럽 항공을 인수할 것 같습니다. 베트남은 올해 감세를 할지도 모르지만 외국인 투자를 장려할 것 같지는 않습니다. 인도네시아는 아마 올해는 생태 관광에 투자할 것으로 보이지만, 이것은 별로 도움이 되지 않습니다. 관광 산업은 확실히 완전히 붕괴될 겁니다.

반대하는 사람들이나 말쑥구거기를 하고하지 않는다면 이 회사는 문을 닫는 것이 좋을지도 모릅니다...

From : 초급 연구원 (백워터인베스트먼트 사)
To : 헤드 퍼스트
제목 : 인도네시아

본사에서 탁상공론만 하는 분들의 이야기에는 귀를 기울이지 않아야 합니다. 현지에서 보면 올해의 관광 산업은 확실히 성장할 가능성이 있고 인도네시아는 생태 관광에 주력하고 있습니다. 본사에서 탁상공론만 하는 분들은 아무것도 모릅니다. 전 제 정보가 경쟁업체들에게서 더 잘 사용될 것이라고 생각하기 시작했습니다.

1) 의견 충돌의 원인이 되고있는 주요 논점

: (1) 러시아가 다음 분기에 석유산업에 보조금을 지원하는가?

(2) 러시아가 유럽 항공을 인수하는가?

(3) 베트남이 올해 감세를 하는가?

(4) 올해 베트남 정부가 외국인 투자를 장려하는가?

(5) 올해 인도네시아 관광 산업은 확대되는가?

(6) 인도네시아 정부는 생태 관광에 투자하는가?

2) 보낸 사람들이 사용하는 '확률 용어'

: 그들은 아마, 매우 가능성이 낮다, 매우 가능성이 높다, 어쩌면, 모른다, 확실히, 전망이, 좋은 기회...등 등 여러가지 사용.

2. 분석가들이 서로 심하게 다투고 있습니다.

짐 : 우리가 개입해서 누가 옳고 누가 틀렸다는 것을 모두에게 말할 필요가 있을까요? 데이터만 보면 되니까 그건 어렵지 않을 것 같은데요.

프랭크 : 그렇게 쉬운 일이 아니에요. 분석가들은 그냥 평범한 사람들이 아니죠. 그들이 담당하는 국가에 관해서는 전문 교육도 받았고 업무 경험도 풍부한 상당한 지역 전문가들이거든요.

조 : 그렇죠. CEO는 그들이 필요한 모든 데이터를 가지고 있다고 하던데. 그들은 세계 최고의 정보에 접근할 수 있어요. 기밀 정보를 손에 넣을 수도 있고, 정부에서 나온 정보를 얻을 수도 있고, 현지에서 조사해서 보고하는 인력도 있어요.

프랭크 : 그리고 정세는 매우 불확실한 요소 투성이랍니다. 그들은 지금 과거의 빈도에 대한 데이터가 없는 단일 사건에 대한 예측을 하고 있지요. 빈도 데이터만 있다면 그것을 보고 더 많은 예측을 할 수 있는데, 그들은 다양한 경로를 통해 데이터를 수집하고 꽤 높은 지식을 기반으로 예측을 하고 있어요.

짐 : 그럼 분석가들이 우리보다 똑똑하기 때문에 우리가 이 논쟁을 해결하기 위해 할 수 있는 것은 아무것도 없다고 말하는 건가요?

조 : 우리가 우리 나름의 데이터 분석 결과를 제공하는 게 그저 논쟁에 기름을 붓는 것처럼 될 수도 있다는 것이죠.

프랭크 : 사실 이 논쟁은 전부 다른 나라에서 무슨 일이 일어났 것인가에 대한 가설을 포함하고 있기 때문에 확률을 표현하는 단어들을 접하면서 분석가들이 혼란스러워하고 있어요. '아마?' 나 '좋은 기회?' 등의 표현은 대체 무슨 의미인가요?

짐 : 그럼 그들의 느낌을 나타낼 수 있는, 더 좋은 단어를 찾는 것을 돕고 싶다는 말인가요? 어이쿠, 그건 시간 낭비인 것 같은데...

프랭크 : 단어를 찾는 건 아닐 겁니다. 비록 누군가의 주관적인 견해를 판단하는 거지만 그 판단을 좀 더 명확하게 표현할 수 있는 무언가를 찾을 필요가 있다는 말입니다.

확률을 표현하는 단어를

좀 더 명확하게 하기 위해
서는

어떻게 해야하지?

3. 주관적 확률은 분석가의 믿음을 나타냅니다.

주관적 확률

> 어떤 대상에 대한 믿음의 정도를 숫자로 표현한 것

> 분석을 통제하기 위한 좋은 방법

=> 단일 사건을 예측하는 경우에 유용!

3. 주관적 확률은 분석가의 믿음을 나타냅니다.

모두 이렇게 말하고 있지만...

석유 산업에 대한 러시아의
지속적인 지원은
매우 가능성이 높아요.



... 실제로는 어떤 의미일까요?

러시아가 석유 산업에 대한 지원을
계속할 가능성은 60%라고 생각합니다.



... 가능성은 70%



... 가능성은 80%



... 가능성은 90%

이것들은 주관적 확률입니다.

이러한 숫자는 분석가가 자신의
생각을 나타내는 데 사용한
단어들보다 훨씬 명확합니다.

1) '모두 이렇게 말하고 있지만...' 말풍선

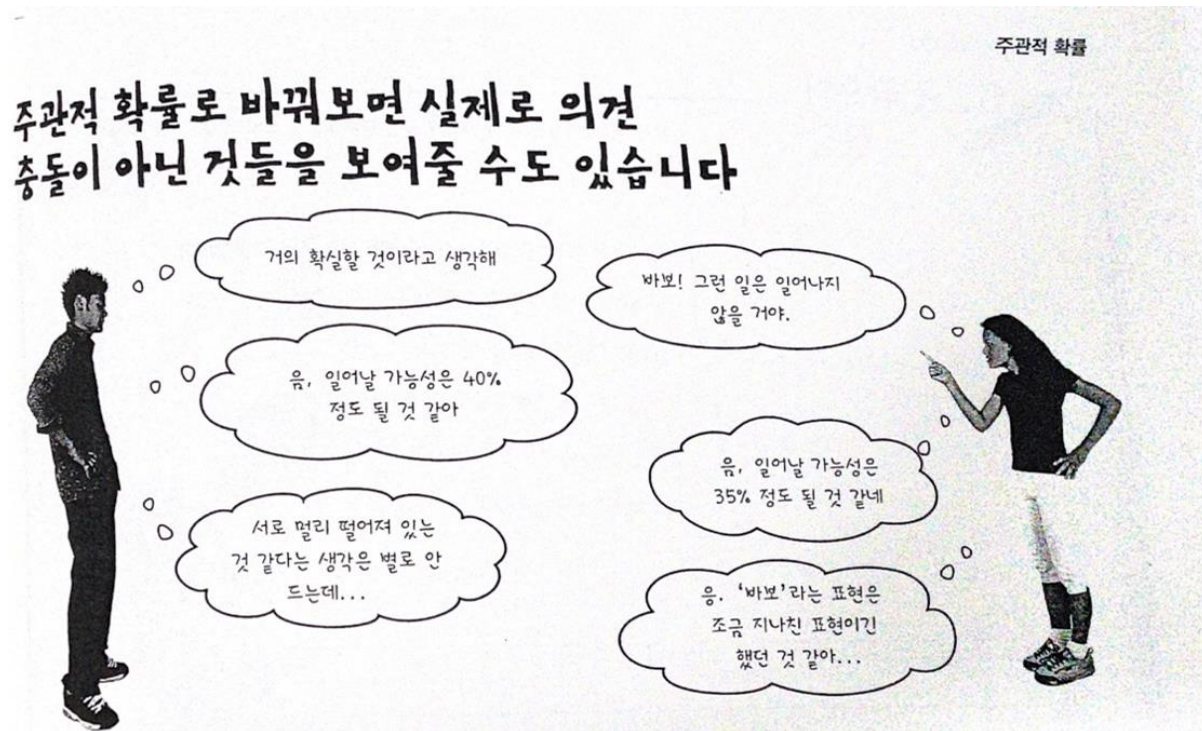
확률 용어 사용

2) '...실제로는 어떤 의미일까요?' 말풍선

주관적 확률 사용

=> 확률 용어보다 훨씬 더 명확함

4. 주관적 확률로 바꿔보면 실제로 의견 충돌이 아닌 것들을 보여줄 수도 있습니다.



4. 주관적 확률로 바꿔보면 실제로 의견 충돌이 아닌 것들을 보여줄 수도 있습니다.

표 형식으로 시각화하다



연필을 꺾으며
정답

분석가들에게서 받은 주관적 확률을 나타내는 스프레드시트는 어떨까요?

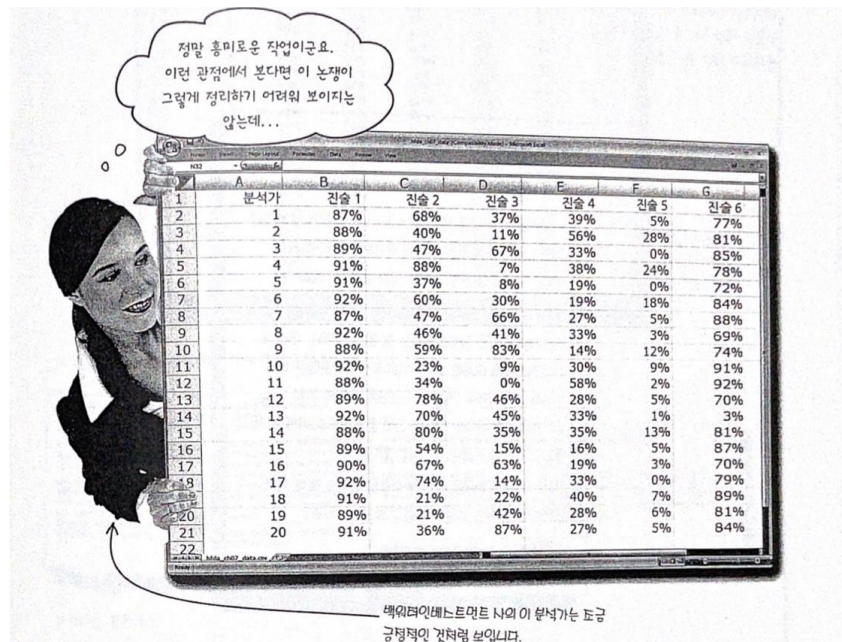
이 표의 맨 왼쪽에 6개의
진동수를 기재합니다.

- 러시아는 다음 분기에 석유 산업에 보조금을 지원합니다.
- 러시아는 다음 분기에 유럽 항공을 인수합니다.
- 베트남은 올해 경제를 합니다.
- 올해는 베트남 정부가 외국인 투자를 장려합니다.
- 올해는 인도네시아 관광 산업이 확대됩니다.
- 인도네시아 정부는 생태 관광에 투자합니다.

각 진동에 대해 각각의 분석가가 생각하는 것을
여기에 있는 빈 칸에 적정할 수 있습니다.

분석가	진동 1	진동 2	진동 3	진동 4	진동 5	진동 6
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						

5. 분석가는 그들의 주관적 확률에 대해 말해주었습니다.



이제 좀 진전을 보이는 것 같군요.

모든 의견 충돌을 해결하는 방법은 아직 찾아내지
못했지만 의견 충돌이 있는 부분을 명확하게
보여주는 것에는 확실히 성공한 것 같습니다.

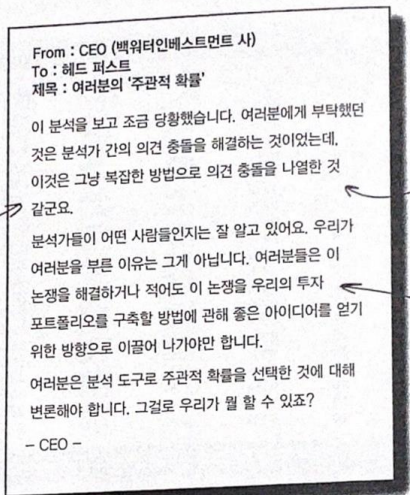
또 몇 가지 데이터를 보면 몇몇 논점은 결국 그렇게
큰 의견 충돌이 있는 것 같지는 않네요.

CEO는 이 데이터에 대해 무엇을 말하고 싶어 하는가?

6. CEO는 여러분의 의도를 모릅니다.

CEO는 이 결과가 분석가들의 의견 충돌을 해결하는 데 도움을 준다고 생각하지 않는 것 같습니다.

CEO는 주관적 확률
누가 도움이 된다고
생각하지 않습니다.



CEO가 부탁한것: 분석가 간의 의견 충돌을 해결하는것

실제로 한 것: 분석가 간의 의견 충돌을 복잡한 방법으로 나열한 것

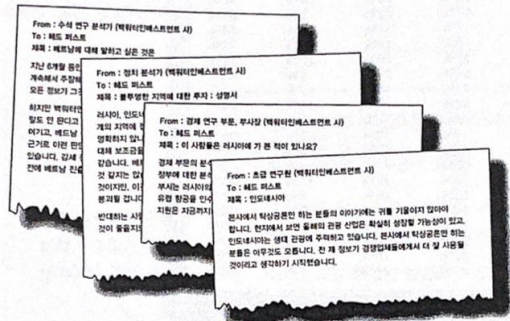
> 데이터를 모은 이유를 CEO에게 설명하고 변론해야함

6. CEO는 여러분의 의도를 모릅니다.

주관적 확률이 정리된 표



... 이 분노 섞인 이메일들보다 분석에 더 도움이 되나요?



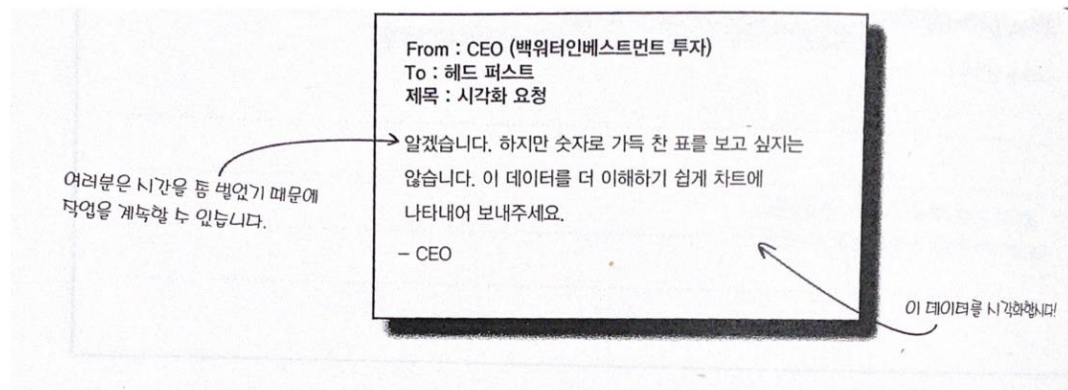
	A	B	C	D	E	F
1	분석가	인물 1	인물 2	인물 3	인물 4	인물 5
2	1	87%	68%	37%	39%	18%
3	2	88%	40%	11%	56%	7%
4	3	89%	47%	47%	33%	18%
5	4	91%	88%	7%	33%	0%
6	5	91%	37%	8%	19%	0%
7	6	92%	60%	30%	19%	0%
8	7	87%	47%	66%	27%	18%
9	8	92%	46%	41%	33%	3%
10	9	88%	59%	83%	14%	0%
11	10	92%	23%	9%	30%	17%
12	11	88%	34%	8%	58%	2%
13	12	89%	78%	46%	28%	5%
14	13	92%	70%	45%	33%	1%
15	14	88%	80%	35%	35%	17%
16	15	89%	54%	15%	20%	6%
17	16	90%	67%	63%	19%	3%
18	17	92%	74%	14%	33%	0%
19	18	91%	21%	22%	40%	0%
20	19	89%	21%	42%	28%	6%
21	20	91%	38%	87%	27%	5%
22						

- 주관적 확률을 보면 몇몇 분야에서는 생각했던 정도의 의견 충돌이 없는 것을 확인할 수 있음.

- 주관적 확률은 어디서 의견 충돌이 일어났고, 그 차이가 얼마나 났는지 명확하게 보여줌.

- 분석가는 주관적 확률을 활용하여 과제를 해결하기 위해 집중해야 할 부분을 알아낼 수 있음

6. CEO는 여러분의 의도를 모릅니다.



데이터 시각화!

6. CEO는 여러분의 의도를 모릅니다.

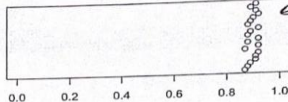
분포를 시각화하다

연필을 꺾으며
정답

여러분이 차트에 표시한 점을 살펴보세요.
분석가의 주관적 확률이 어떻게 분포되어
있나요?

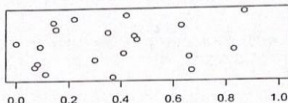
이 차트에 대해서는 사실 어느
정도의 차이가 있는 것 같습니다.

진술 1
러시아는 다음 분기에 석유 산업에 보조금을 지원할 것입니다.



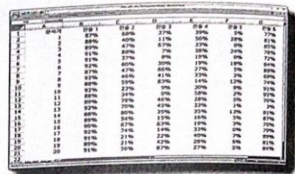
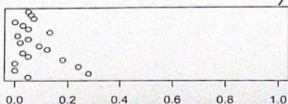
이 차트에 대해서는 분석가의
의견이 곳곳에 분산되어 있습니다.

진술 3
베트남은 올해 감세를 할 것입니다.

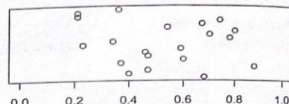


여기에서도 부분적인 양의를 볼 수 있습니다.

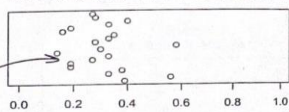
진술 5
올해는 인도네시아 관광 산업이 확대될 것입니다.



진술 2
러시아는 다음 분기에 유럽 항공을 인수할 것입니다.

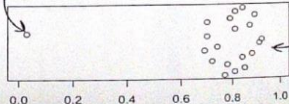


진술 4
올해는 베트남 정부가 외국인 투자를 장려할 것입니다.



여기서는 한 명을 제외하고 20%의
범위 내에 있습니다.

진술 6
인도네시아 정부는 생태 관광에 투자할 것입니다.



진술1: 러시아는 다음 분기에 석유 산업에 보조금을 지원할 것입니다.

진술2: 러시아는 다음 분기에 유럽 항공을 인수할 것입니다.

진술3: 베트남은 올해 감세를 할 것입니다.

진술4: 올해는 베트남 정부가 외국인 투자를 장려할 것입니다.

진술5: 올해는 인도네시아 관광 산업이 확대될 것입니다.

진술6: 인도네시아 정부는 생태 관광에 투자할 것입니다.

7. CEO가 여러분의 작업을 마음에 들어 합니다.

From : CEO (백워터인베스트먼트 투자)

To : 헤드 퍼스트

제목 : 감사합니다!

정말 많은 도움이 되었습니다. 자세한 정보를 얻기 위해
자원을 집중해야 할 분야를 알 수 있게 되었어요. 또
실제로는 의견 충돌이 아닌 것 같은 것들을 알게 된 것은
매우 훌륭했습니다.

앞으로는 주관적 확률(또는 가능하면 객관적 확률)로
표현하지 않은 분석가의 보고는 듣고 싶지 않습니다.

이러한 논점을 의견 충돌의 정도에 따라 순위를
매겨주시지 않겠습니까? 구체적으로 어떤 진술에 대한
의견이 가장 차이가 있는지 알고 싶습니다.

- CEO

주관적인 확률

> 누구나 이해할 수 있지만 충분히 사용되지 않고 있음

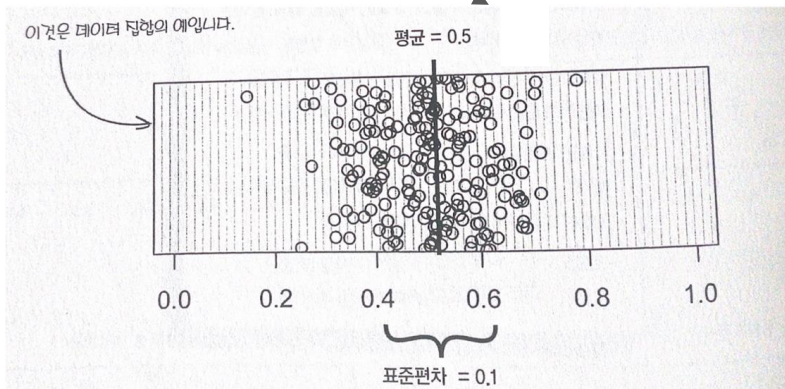
> 자신의 생각과 믿음을 다른사람에게 정확하게 전달하는 계몽
적인 방법



1. CEO는 주관적 확률(또는 객관적 확률)로 표현하지 않은 분석가의 보고는 안듣고 싶어 함.
2. 어떤 진술에 대한 의견이 가장 차이가 있는지 알고 싶어 함.

8. 표준편차는 평균에서 점들이 얼마나 떨어져 있는지를 측정합니다.

데이터 집합의 대부분의 관측점은
산술 평균의 표준편차 1개 분
이내의 범위에 들어감



표준편차

- > 데이터 집합의 표준값이 산술 평균에서 얼마나 떨어져 있는지를 측정함
- > 표준편차의 단위는 측정하는 대상에 따라 달라짐
- > 표준편차를 사용해서 의견 충돌의 정도를 측정할 수 있음

8. 표준편차는 평균에서 점들이 얼마나 떨어져 있는지를 측정합니다.

러시아가 사업에 대한 자신감 저하를
언급하며 유전을 전부 매각할 것이라고
발표했습니다.

충격적인 행보입니다. 러시아 대통령이 국책 산업을
포기했네요.

러시아 대통령은 모스크바에서 오늘 새벽 “음..우린
석유를 포기합니다”라고 말해 기자들을 깜짝 놀라게
했습니다. “우린 그저 이 산업에 대한 자신감을
잃었을 뿐이고, 그래서 더 이상 자원을 캐는 데에도
관심이 없어졌습니다.”

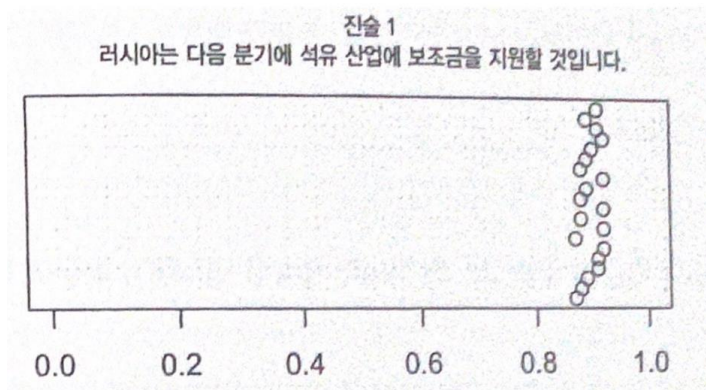
러시아, 국책 산업 포기!

끔찍하군요! 우리 모두 러시아가
계속 이 사업에 대해 자신감을 가
지고 있다고 예상했는데.



[분석가]

9. 이 뉴스는 전혀 예상하지 못했던 것입니다.



분석가들 우려O

> 석유에 대한 지속적인 지원이 있을 것이라는 전망에 대해서는 의견의 일치가 있었기 때문에 백워터인베스트먼트 사는 러시아 석유에 거액의 투자를 하고 있었음

> 하지만 이 뉴스를 본 사람들은 갑자기 러시아 석유 산업에 큰 문제가 있을거라고 생각하기 때문에 이 투자는 가치가 급락할 가능성이 있음

> 러시아의 전략일 수 있고 실제로는 유전을 매각하지 않을 수도 있음

10. 베이즈 정리는 주관적 확률 검토에 제격입니다.

이 증거를 전제로 한
가설의 확률

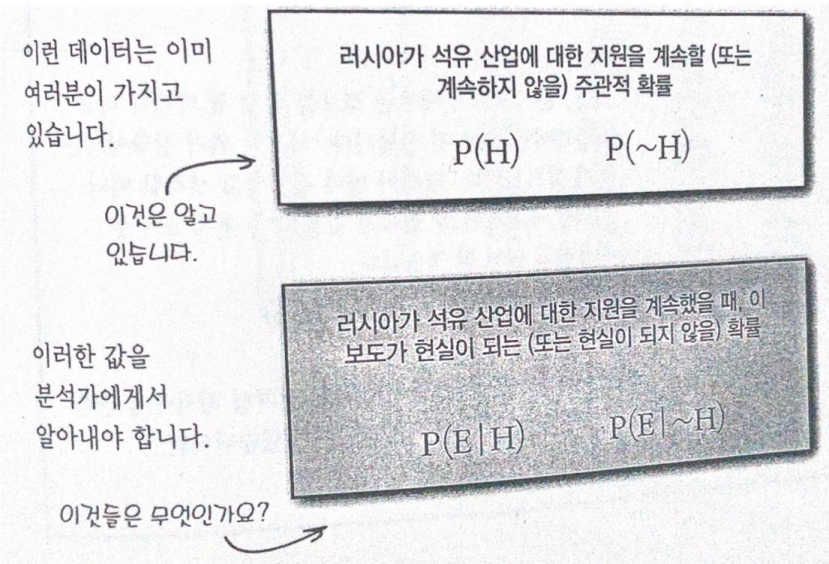
이것이 여러분이 알고
넣어 하는 것입니다.

$$P(H|E) = \frac{P(H)P(E|H)}{P(H)P(E|H) + P(\sim H)P(E|\sim H)}$$

가설이 참일 경우
이 증거가 실현될 확률

가설이 거짓일 확률

가설이 거짓일 경우
이 증거가 실현될 확률



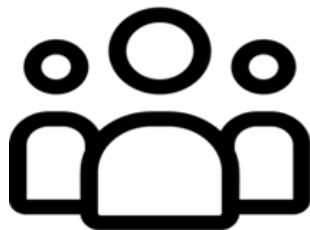
베이즈 정리

> 새로운 증언을 가설에 대한 믿음과 통합할 수 있기 때문
에 도마뱀 독감 뿐만 아니라 주관적 확률을 검토하는 데에
도 제격임.

> H: 가설(or 기준 비율)

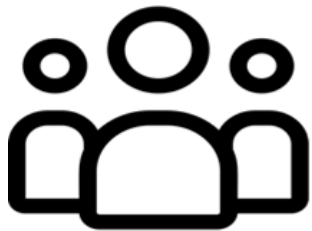
> E: 새로운 증거

> 베이즈 정리와 주관적 확률을 이용하면 **가설이 참일 경우
증거가 실현될 확률**을 구할 수 있음



왜 이런 고생을 하는거죠? 분
석가에게 돌아가 이 사건에 대
한 그들의 반응을 기반으로 새
로운 주관적 확률을 물어보면
안되나요?

10. 베이즈 정리는 주관적 확률 검토에 제격입니다.



왜 이런 고생을 하는거죠? 분석가에게 돌아가 이 사건에 대한 그들의 반응을 기반으로 새로운 주관적 확률을 물어보면 안되나요?

[베이즈 정리를 사용하는 이유]

> 분석가들이 새로운 정보를 알게될 때 마다 베이즈 정리를 사용해야 하는것은 아님

> 위기 상황에서는 베이즈 정리가 필요함
=> ex) 병에 걸렸다고 생각될때, 거액을 투자해야 할 필요가 있을때

10. 베이즈 정리는 주관적 확률 검토에 제격입니다.

이 식은 분석가의 기존 비율과 새로운 데이터에 대한 판단을 결합하여 새로운 평가 결과를 이끌어냅니다.

$$= (B2 * D2) / (B2 * D2 + C2 * E2)$$

이것이 결과입니다.

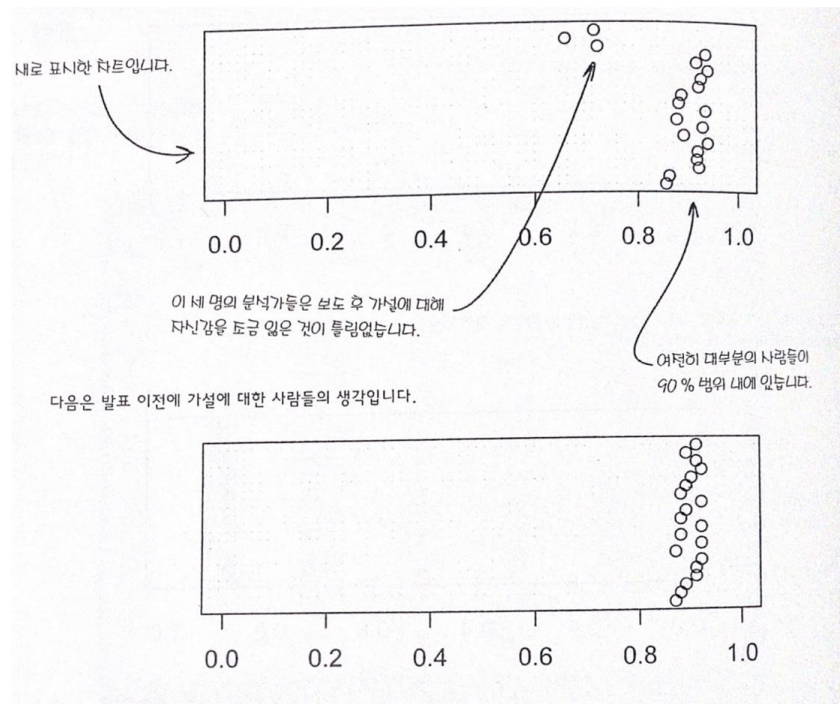
분석가들이 수집한 두 개의 새로운 주관적 확률이 나열되어 있는 스프레드시트

	A	B	C	D	E	F
1	분석가	P(S1)	P(~S1)	P(E S1)	P(E ~S1)	P(S1 E)
2	1	87%	13%	54%	61%	86%
3	2	88%	12%	57%	67%	86%
4	3	89%	11%	55%	39%	92%
5	4	91%	9%	58%	54%	92%
6	5	91%	9%	58%	53%	92%
7	6	92%	8%	64%	49%	94%
8	7	87%	13%	65%	54%	89%
9	8	92%	8%	50%	45%	93%
10	9	88%	12%	53%	55%	88%
11	10	92%	8%	62%	51%	93%
12	11	88%	12%	56%	56%	88%
13	12	89%	11%	59%	62%	89%
14	13	92%	8%	61%	62%	92%
15	14	88%	12%	66%	40%	92%
16	15	89%	11%	54%	29%	94%
17	16	90%	10%	69%	58%	91%
18	17	92%	8%	67%	55%	93%
19	18	91%	9%	14%	55%	72%
20	19	89%	11%	22%	93%	66%
21	20	91%	9%	16%	65%	71%

- 1) $P(E|S1)$: 러시아가 석유 산업에 대한 지원을 계속할 경우 (S1)에 유전 매각을 발표할 가능성(E)에 대한 각 분석가들의 주관적 확률
- 2) $P(E|\sim S1)$: 러시아가 석유 산업에 대한 지원을 계속하지 않을 경우 (~S1)에 유전의 매각을 발표할 가능성(E)에 대한 각 분석가들의 주관적인 확률

10. 베이즈 정리는 주관적 확률 검토에 제격입니다.

<러시아가 석유 산업을 지원할 것이라는 가설에 대한 믿음의 분포>



발표 이전에 가설에 대한 사람들의 생각 VS 발표 이후 가설에 대한 사람들의 생각

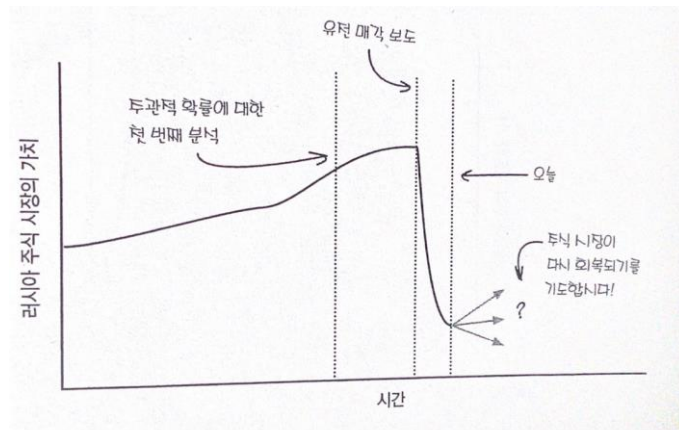
- 1) 발표 이전보다 발표 이후에 주관적 확률의 분포는 조금 넓어졌고, 발표 이후에 가설의 주관적 확률이 크게 낮아진 사람은 3명!
- 2) 대부분의 사람들은 러시아가 유전 매각을 발표했음에도 불구하고 석유 산업에 한 지원을 계속 할 가능성은 90% 정도 됨

11. CEO는 이 새로운 정보를 가지고 무엇을 해야 할지 명확히 알았습니다.

모든 회사들이 러시아 자산을 매각하고
있지만 우리 분석가의 생각에 대한 이 새로운
데이터들을 볼 때 우리의 자산을 유지하는 것이
 좋을 것 같네요. 잘 되길 빕시다!



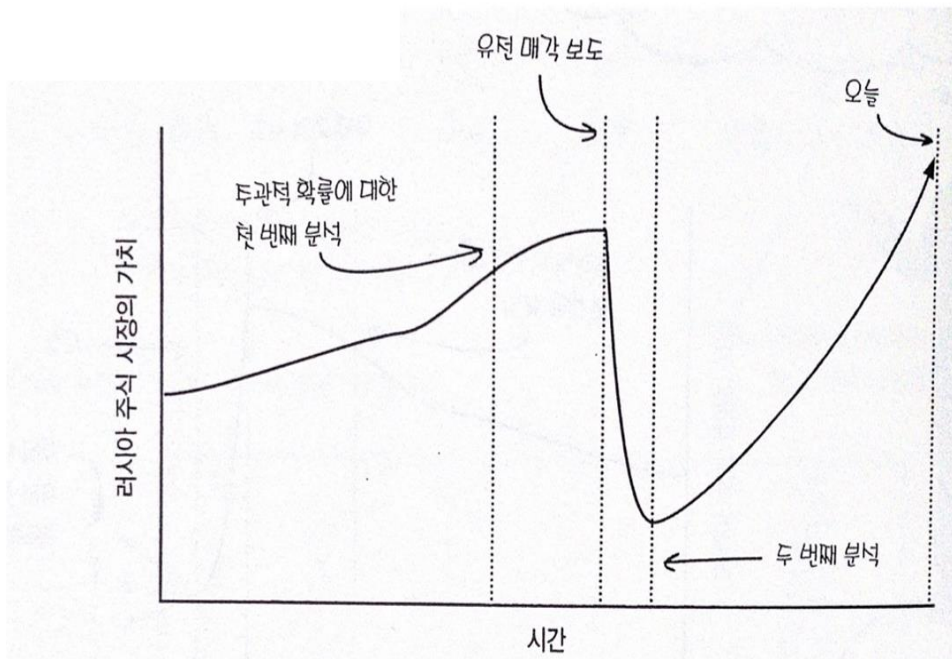
[백워터인베스트 사의 CEO]



- 검토 결과, 분석가들은 러시아의 석유
산업에 대한 지원 여부와 상관없이 유전 매각
보도를 할 것이라고 판단함

하지만, 분석가들은 옳은 판단을 한 걸까요?

12. 러시아 주식 보유자는 매우 기뻐하고 있습니다!



분석가들이 옳았음

> 러시아는 유전 매각에 대해 잘못된 정보를 흘려 사람들을 속임

=> 사실이 밝혀지자, 시장이 활황 장세로 돌아섬

=> 백워터인베스트먼트 사에 매우 이득이 됨

Thank You