

Frequentism vs. Bayesianism (Frequentist vs. Bayesianist)

소프트웨어 끈대 강의

노기섭 교수

(kafa46@cju.ac.kr)

Motivation for this Video

여론 조사 결과가 나왔네?

그러면 000 후보가 당선될 확률이
신뢰수준 95%에서
70% 확률로 당선 되겠군...



뭐라고!!!
왜 표본 크기가 1,000명인 것이
문제가 되는 거지?
이건 통계적으로 정확한거야!!!

전국 1000명을 대상으로
차기 선거 적합도 조사를 했다는 군...

5천만 인구 중에서 1,000명 설문이
정확하겠어?
빅데이터 분석을 하는게 훨 좋을 거야



보다 많은 데이터를 기반으로
예측하는게 더 좋지 않아?

Paradigm in Probability

공통적 목표: 데이터를 이용하여
추론 (Inference)

■ Frequentist (frequentism)

- Ronald Fisher (1890~1962)



Long-term relative frequency

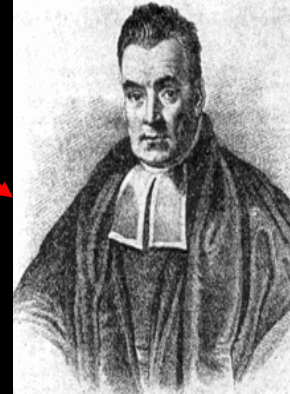
Fixed parameters

Random data

- Sometimes called
 - Traditional (classical) statistics
- Pros
 - Ease of interpretation & computation
 - No prior knowledge, Specific tool of test
- Cons
 - Depending on the number of times conducted
 - Not informative for other hypothesisises

■ Bayesianist (Bayesianism)

- Thomas Bayes (1701~1761)



Update beliefs based on new information

Degree of uncertainty (Unknown parameters)

Fixed data

- Sometimes called
 - Data-driven, evidence/information basis
- Pros
 - Higher precision than classical approaches
 - More informative
- Cons
 - Intensive computation (not popular in past)
 - Require prior knowledge & very complex

Problems in Academic Fields

■ Frequentism

- Widely taught, but very poorly
- 확률/통계 공식을 활용한 교육
 - 비판적 사고 X
 - 자신만의 방법론 활용 자체가 어려움
- 구석기 시대의 접근법을 여전히 사용
 - 구석기 도구 선호: z-score, p-value, t-test, table look-up, some charts, SPSS, ...
 - 최신 도구를 사용하지 않음: scipy, numpy, or many computational libraries

■ Bayesianism

- Not widely taught
- 복잡한 수학 공식을 도입
- 대부분 추론(inference) 방법론에서 대부분 교육 중단
- 이미 사용하지 않는 오래전 기법들 위주로 교육

Hypothesis

■ Hypothesis (가설)

- A proposed explanation for a phenomenon.
 - 하나의 현상에 대하여 제안된 설명
 - 현실적으로 증명하기 어려운 현상에 대한 예측

From wikipedia: <https://en.wikipedia.org/wiki/Hypothesis>

■ Hypothesis 특징

- 가설은 맞든지, 틀리든지 → 결론은 True or False
- True/False 증명을 위해 무한대로 실험할 수는 없다.

■ Statistical Inference

- 과학적으로 완전하게 설명할 수 없는 관찰결과에 대하여 가설을 수립
- Determine statistical significance of observed data
 - 확보한 데이터의 통계적 유의성 측정

Frequentist Approach - Beginning of Problem

■ 주사위 던지기 실험 데이터에서 3이 나올 확률?

- 이론적 값

$$\frac{1}{6} \cong 0.167 (16.7\%)$$

데이터에서 얻은 값이
이론값과 완전 다를 경우,
데이터로부터 추론한 것이 맞나?



- 6번 실험한 데이터로부터 확률 값을 구할 경우?

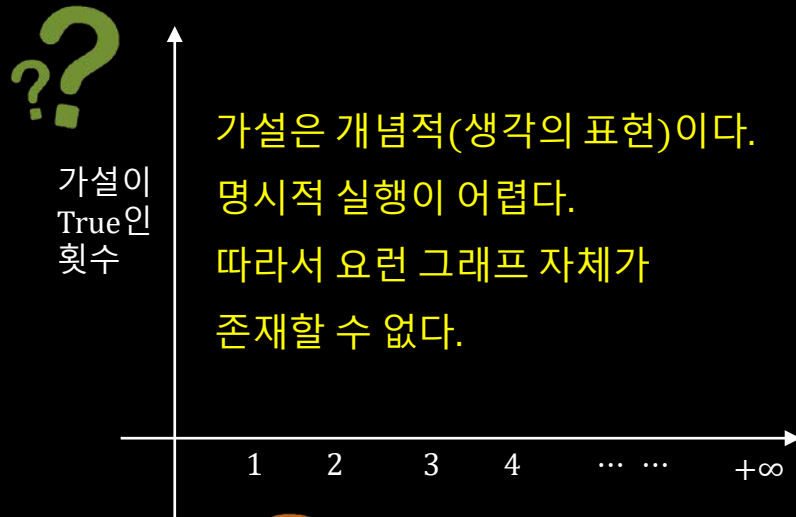
데이터 종류	데이터 수집	3이 나올 확률
Dataset 1	\Rightarrow \Rightarrow \Rightarrow \Rightarrow \Rightarrow	$\frac{1}{6} = 0.167(16.7\%) ?$
Dataset 2	\Rightarrow \Rightarrow \Rightarrow \Rightarrow \Rightarrow	$\frac{0}{6} = 0.0 (0.00\%) ?$
Dataset 3	\Rightarrow \Rightarrow \Rightarrow \Rightarrow \Rightarrow	$\frac{6}{6} = 1.0 (100\%) ?$

Frequentist Approach - Possible Solution

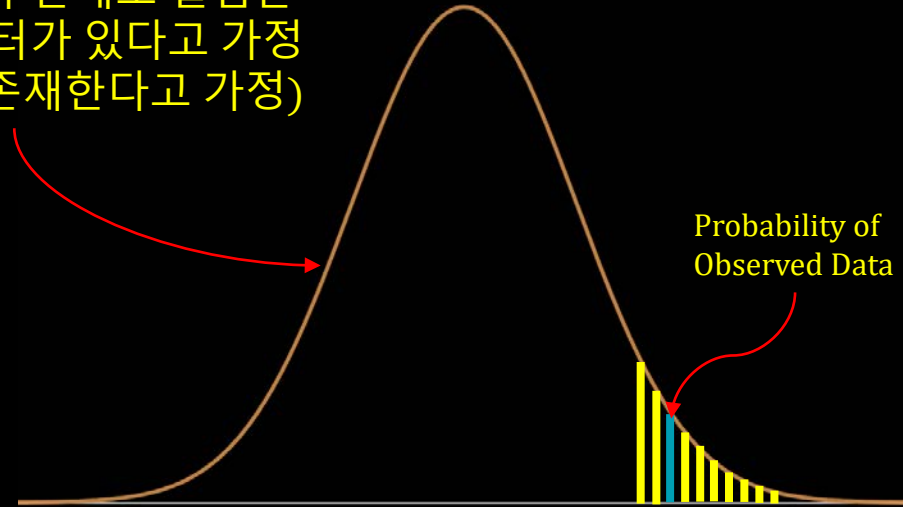
■ Frequentists' Idea



Frequentist Approach - Obstacles & Solution



무한대로 실험한
 데이터가 있다고 가정
 (어떤 분포가 존재한다고 가정)



Calculating the probability of
 observed data occurring

가설을 실행한 횟수

대표적 tool
 (p - value)

만약 관측 데이터의 출현 확률이
 가정했던 분포와 명확히 차이가 난다면,
 → 우리 가설이 틀린(false) 것으로 판정



"귀무가설(H_0 , null hypothesis)
 기각한다."

만약 관측 데이터의 출현 확률이
 가정했던 분포의 어떤 기준 범위에 있다면,
 → 우리 가설이 맞는(True) 것으로 판정



"귀무가설(H_0 , null hypothesis)
 채택한다."

약간 유식한 표현으로...

Bayesian in Our Life

아내가 임신했어요 ^^.
그런데 아들일까? 딸일까?



예비 아빠
노기섭 교수님



아들일 확률
50% 일거야!



통계학과
홍길동 교수님



청주지역 남아
출생률은 65% 입니다.



산부인과 의사
진찰해 박사님

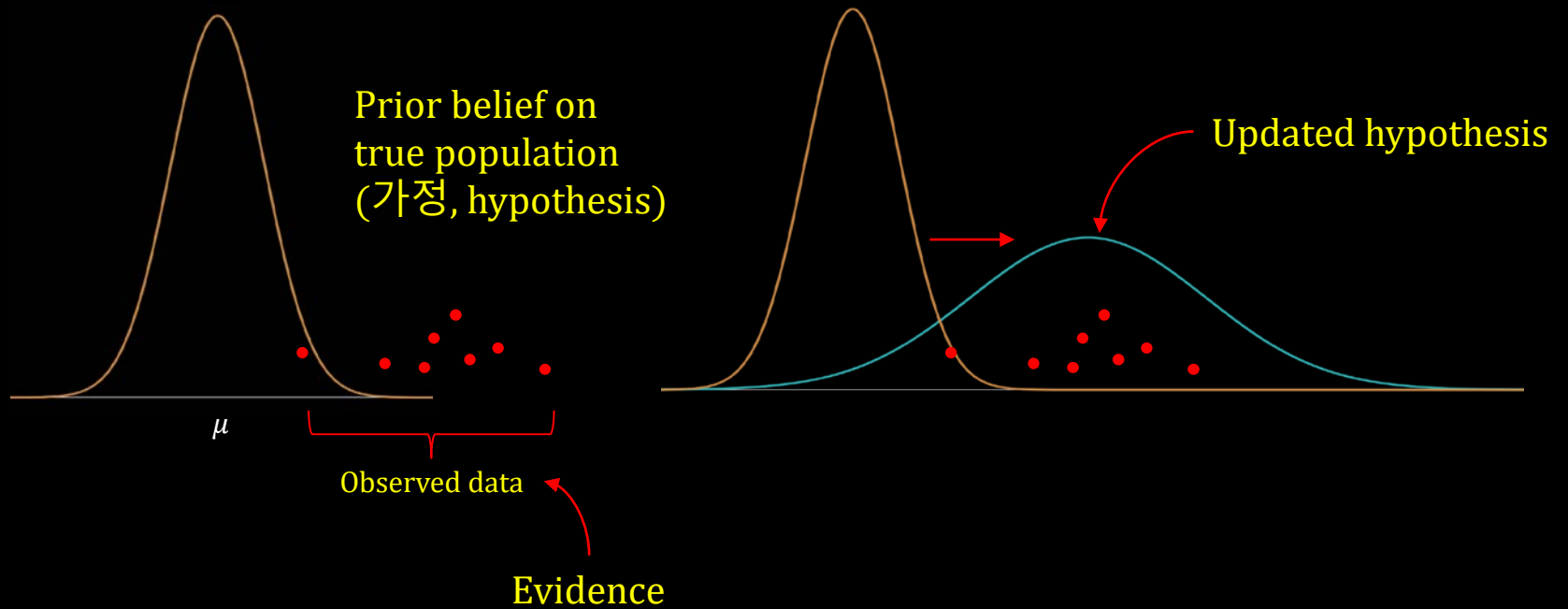


초음파 검사 해보니
아들 확률이 98% 입니다.

Bayesian 관점: 모두 맞는 답이 될 수 있다.

“Probability는 각 개인의 belief 또는 uncertainty level이다.”

Bayseian Approach



Bayesian: Degrees of certainty
→ No relative frequency!

Bayes Theorem in Math

Conditional Probability

■ Join Probability (결합 확률)

- 두 사건이 동시에 발생할 확률
- (Example) 주사위를 던지는 실험
 - Sample space: $S = \{1, 2, 3, 4, 5, 6\}$
 - 사건 A: 짝수가 나오는 사건 $\rightarrow A = \{2, 4, 6\}$
 - 사건 B: 3의 배수가 나올 확률 $\rightarrow B = \{3, 6\}$
 - 사건 C: 짝수이면서 3의 배수가 나오는 사건 $\rightarrow C = \{6\}$
 - Random variable X : 주사위를 던졌을 때 나온 값

S	X	x
1	\rightarrow	1
2	\rightarrow	2
3	\rightarrow	3
4	\rightarrow	4
5	\rightarrow	5
6	\rightarrow	6

$$P(A) = P(X = 1 \text{ or } X = 2 \text{ or } X = 3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6}$$

$$P(B) = P(X = 3 \text{ or } X = 6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}$$

$P(C)$ 를 A 와 B 로 표현하려면?

$$P(C) = P(A, B) = P(A \cap B) = P(X = 6) = \frac{1}{6}$$

Conditional Probability

■ $P(C)$ 를 A 와 B 로 표현하려면?

- Sample space: $S = \{1, 2, 3, 4, 5, 6\}$
- 사건 A: 짝수가 나오는 사건 $\rightarrow A = \{2, 4, 6\}$
- 사건 B: 3의 배수가 나올 확률 $\rightarrow B = \{3, 6\}$
- 사건 C: 짝수이면서 3의 배수가 나오는 사건 $\rightarrow C = \{6\}$

A와 B가 동시에 발생할 확률

동시에 발생(곱하기)

$$P(C) = P(A \cap B) = P(A, B) = P(A) \times P(B|A) = P(B|A) \times P(A)$$

A가 발생할 확률

A가 발생한 상황에서 B가 발생 확률

■ 조건부 확률의 탄생!

$$P(A, B) = P(A) \times P(B|A) \rightarrow P(B|A) = \frac{P(A, B)}{P(A)}$$

Deriving Bayes Theorem

■ Bayes theorem (베이즈 정리) 탄생!

$$P(A|B) = \frac{P(A, B)}{P(B)}$$



두개가 똑같음

$$P(B|A) = \frac{P(A, B)}{P(A)}$$



$$P(A, B) = P(A|B) \times P(B)$$

$$P(A, B) = P(B|A) \times P(A)$$

따라서, 요 두 녀석도 똑같음

$$P(A|B) \times P(B) = P(B|A) \times P(A)$$

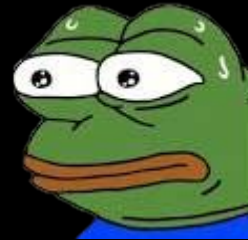
깔끔하게
다시 정리하면
요게 베이즈 정리
^^

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

$$P(B|A) = \frac{P(A|B) \times P(B)}{P(A)}$$

Interpreting Bayes Theorem

■ 베이즈 정리를 해석하는 방법



헐 뭘 말인지
도대체...

Likelihood
(우도, 가능성)

Prior probability
(사전 확률)

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Posterior probability
(사후 확률)

Evidence(증거)

Summary

■ 다시 한번 가설(Hypothesis)과 증거(Evidence)로 정리하면,

$$P(H|E) = \frac{P(E \cap H)}{P(E)} = \frac{P(E|H) \times P(H)}{P(E)} = \frac{P(E|H)}{P(E)} \times P(H)$$

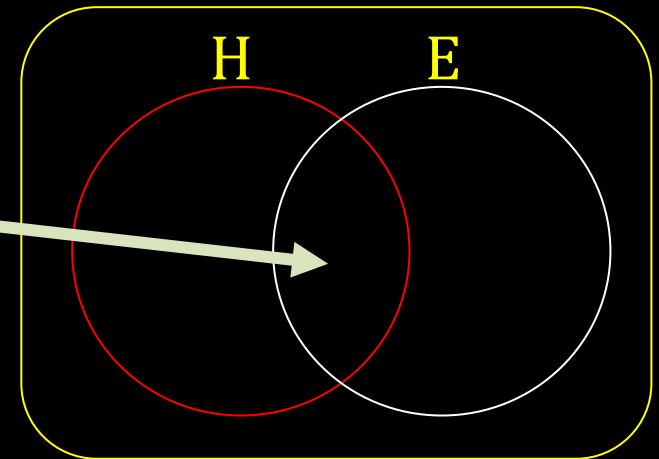
$P(H|E) = \alpha \cdot P(H)$
Updating Factor

Update

Normalization Constant

$$= \frac{P(E|H) \times P(H)}{P(E|H)P(H) + P(E|H^c)P(H^c)}$$

$$\begin{aligned} P(E \cap H) &= P(E|H) \times P(H) \\ &= P(H|E) \times P(E) \end{aligned}$$



Excellent References

- **CrashCourse, "You Know I'm All About that Bayes: Crash Course Statistics #24"**
 - YouTube link: <https://www.youtube.com/watch?v=9TDjifpGj-k>
- **Adam Pierce, "Breakthrough Junior Challenge 2020 || Probability Theory: Frequentist Vs. Bayesian"**
 - YouTube link: https://www.youtube.com/watch?v=CaU1V-C_iZU
- **Woody Lewenstein, "Introduction to Bayesian Statistics - A Beginner's Guide"**
 - YouTube link: <https://www.youtube.com/watch?v=NIqeFYUhSzU&t=2123s>
- **edanz learning lab, "Frequentist vs. Bayesian Statistics – Which should you use?"**
 - Web link: <https://learning.edanz.com/frequentist-bayesian-statistics/>
- **Ravin Kumar, "Bayesian Vs Frequentist: Which should you be"**
 - Web link: <https://youtu.be/foSPfzYs4yY?si=4pC0ReFgGe1BjPu>
- **Why we cannot computing the integral of $P(x)$ in Bayesian Inference analytically? 적분의 어려움, StackExchange**
 - Web link: <https://stats.stackexchange.com/questions/510193/why-we-cannot-computing-the-integral-of-px-in-bayesian-inference-analytically>
- **Bayesian Neural Network (베이지안 뉴럴 네트워크) 내용 정리, Korean Blog**
 - https://gaussian37.github.io/dl-concept-bayesian_neural_network/
- **데이터 사이언스 스쿨 - 베이즈 정리**
 - <https://datascienceschool.net/02%20mathematics/06.06%20%EB%B2%A0%EC%9D%B4%EC%A6%88%20%EC%A0%95%EB%A6%AC.html>



수고하셨습니다 ..^^..