차등정보보호란 무엇인가
(Differential Privacy 맛보기)

박민정 / 통계청 통계개발원 통계방법연구실

개인정보 노출?

예를 들어 생각해 봅시다.

Alice와 Bob은 대학교수입니다. 그들은 연구를 위해 학교 데이터베이스에 접근이 필요합니다. 그 데이터베이스에는 학생에 관한 개인 정보도 포함되어 있습니다. 접근 권한을 얻기 위해 Alice와 Bob은 **기밀 유지 교육**을 받고, 데이터베이스에서 얻은 개인 정보의 사용 및 공개를 금지하는 **계약서에 서명**했습니다.

3월에 Alice는 이 데이터베이스 분석 결과를 기반으로 한 기사를 발표합니다. "대학 신입생은 **3,005 명**이며, 이중 **202 명**은 연간 35만 달러 이상의 수입을 올리는 가정입니다."라는 내용이 있습니다. **3,000 명이 넘는 집단에 대한 통계**이므로, Alice는 어떤 개인 정보도 공개되지 않았다고 생각합니다.

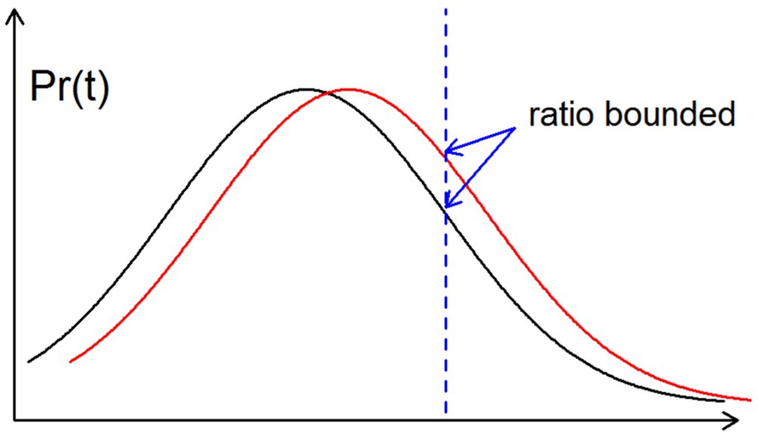
4월에 Bob은 자료를 분석하고, 다음 통계를 포함하는 별도의 기사를 게시합니다. "대학 신입생 **3,004 명 중 201 가구**가 연간 35 만 달러를 초과하는 가계 소득을 가지고 있습니다." Alice와 Bob은 둘 다 비슷한 정보를 발표했음을 알고 있지 않습니다.

영리한 학생 Eve가 이 두 기사를 읽었습니다. Eve는 3월과 4월 사이에 한 신입생이 대학을 그만두었으며, 학부모는 연간 35만 달러 이상의 수입을 얻는다고 결론을 내립니다. Eve는 주위에 묻고, John이라는 학생이 3월 말쯤에 중퇴했다고 듣습니다.
Eve는 학급 친구들에게 **John의 부모님의 수입은 아마도 1년에 35만 달러 이상일 것**이라고 알려줍니다.

차등정보보호 정의

레코드(응답자 한 명의 정보) 하나만 다른 두 개의 데이터베이스 D와 D'를 분석한 결과($K(D)$) 분포의 비율의 차이를 ε으로 제한!

$$\left| \log \frac{\Pr[\,K(D) = t\,]}{\Pr[\,K(D') = t\,]} \right| \leq \varepsilon$$



$\varepsilon \downarrow\downarrow$ 결과물 분포의 차이가 작음 : **강한** 정보보호
$\varepsilon \uparrow\uparrow$ 결과물 분포의 차이가 큼 : **약한** 정보보호

차등정보보호 구현

D 및 D' : 레코드 하나만 다른 두 개의 데이터베이스
f : 이용자가 요청하는 쿼리(통계)
K_f : 메커니즘 (차등정보보호를 만족하는 쿼리에 대한 결과값)

$K_f = f + noise$: 메커니즘 (DP를 만족하는 쿼리에 대한 결과값)
$noise \sim Laplace(\Delta f/\varepsilon)$, $\Delta f = max_{D,\,D'} |f(D) - f(D')|$

Privacy? 생각 뒤집기

(과거) 정보보호에 대한 생각

Dalenius의 명제 (1977년) : 이상적인 정보보호
"데이터베이스에 접근해서 얻을 수 있는 개인에 대한 정보는 데이터베이스에 대한 접근 없이도 알 수 있어야 한다."

(예) Terry Gross 의 키

(민감정보) 어떤 개인의 키가 얼마인지
(보조정보) Terry는 리투아니아 여성 평균키보다 2센치 작다
(데이터베이스) 인종별 여성 평균키

데이터베이스 접속하지 않은 외부인의 지식 → Terry는 좀 작다
데이터베이스 접속한 외부인의 지식 → Terry의 정확한 키

(결론) 데이터베이스 접속 여부에 따라 개인정보 노출 발생!

차등정보보호 Differential Privacy (Dwork, 2006)

보조정보로 **이상적인 정보보호는 불가능!** → 수학적 증명을 제시

<노출 판단 기준 전환>
외부인의 데이터베이스 접근 → 응답자의 데이터베이스 참여

(예) Terry Gross 의 키

Terry의 **데이터베이스 미참여** → Terry의 정확한 키 노출
Terry의 **데이터베이스 참여** → Terry의 정확한 키 노출

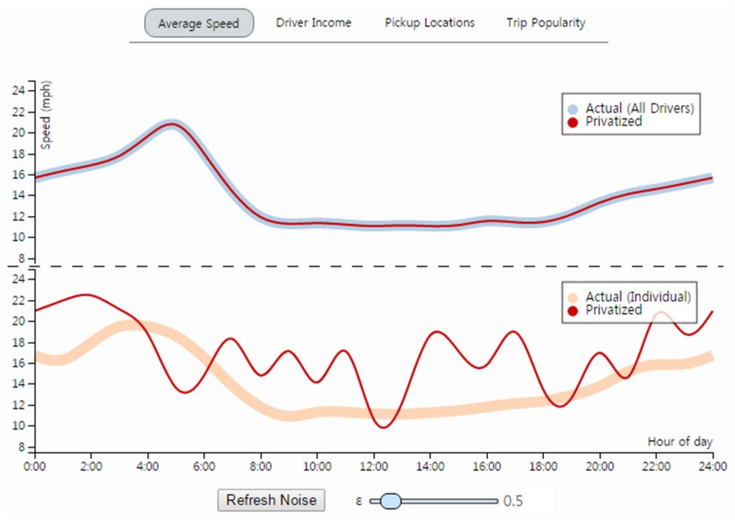
(결론) **정보노출 결과가 동일** → 차등정보보호 달성!

"새로운 노출위험"의 **개념과 수학적 정의**를 제시!!

적용한다면?

교통대책 수립을 위해 택시의 평균 운행속도가 필요하다고 상상하자.

강한 수준으로 차등정보보호를 적용할 때 ($\varepsilon = 0.5$)



차등정보보호 적용과 상관없이, 택시운전기사 평균 속도가 정확하게 제공됨

강한 수준으로 적용하면, 택시기사 1명의 속도를 정확하게 알 수 없음

약한 수준으로 차등정보보호를 적용할 때 ($\varepsilon = 4.12$)

차등정보보호 적용과 상관없이, 택시운전기사 평균 속도가 정확하게 제공됨

약한 수준으로 적용하면, 운전자 1명의 속도를 좀 더 정확하게 파악할 수 있음 (과속에 대한 벌금 부과?)

참고문헌

- 박민정, 이용희, 권성훈(2018), 차등정보보호에 관한 연구, 통계개발원. (+ 보고서에 수록된 참고문헌)