인공지능을 위한 수학
SECTION 4-1

윤정인
2020531001

[6주차 발표]

01-1

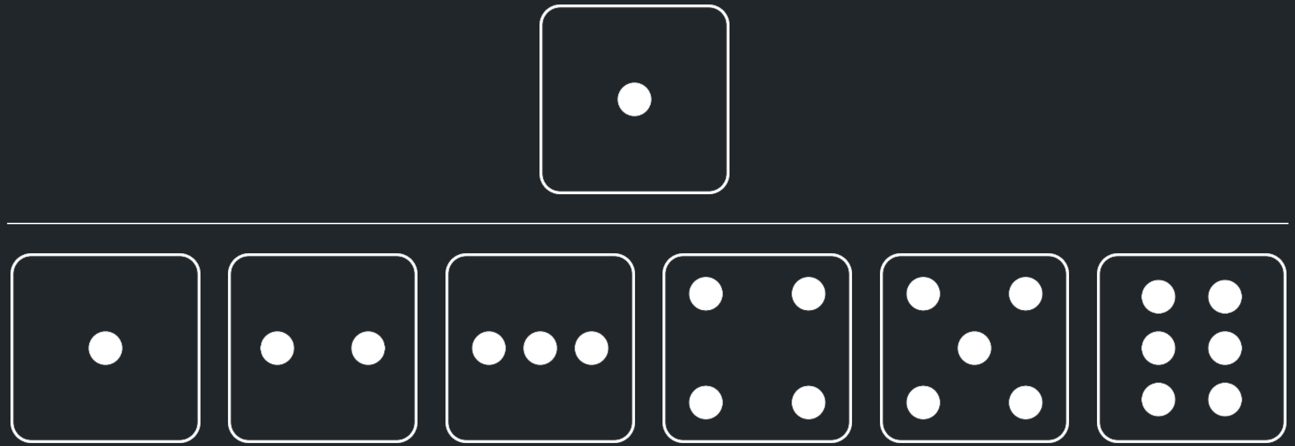
확률 구하기

$$\text{확률} = \frac{\text{어떤 사건이 발생할 수 있는 경우의 가짓수}}{\text{모든 경우의 가짓수}}$$

Probability

0% ~ 100%

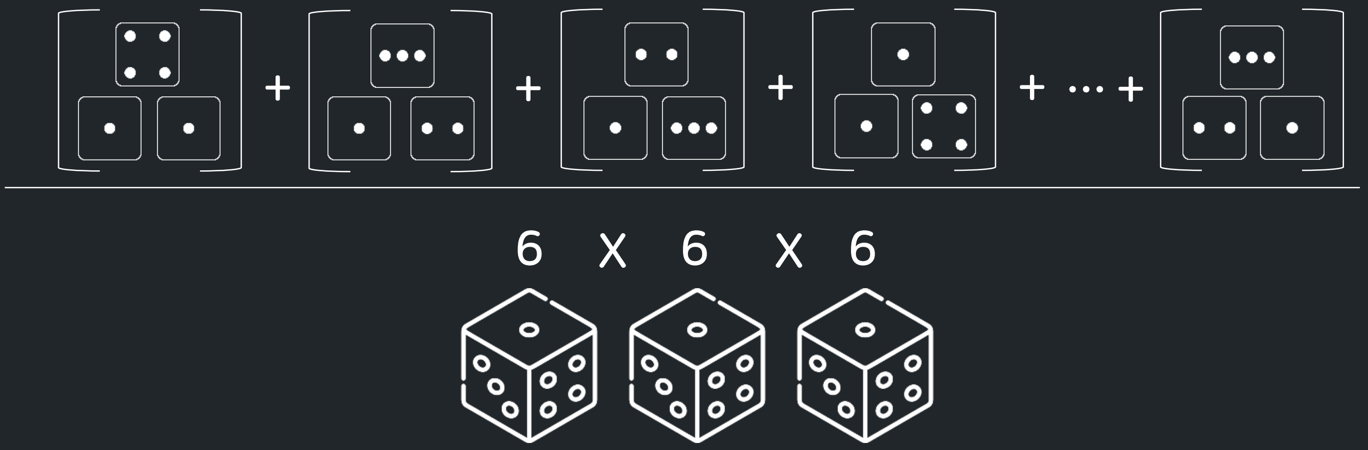
주사위를 던졌을 때
1이 나올 확률 =



$$= \frac{1}{6}$$

주사위를 3개 던져
눈의 합이 6이 될 확률

=



$$= \frac{\overset{5}{\cancel{10}}}{\underset{108}{\cancel{216}}}$$

4.63 %

01-2

조합의 공식

조합의 공식

$$_{n}C_{k} = \frac{n \cdot (n-1) \cdots\cdots (n-k+1)}{1 \cdot 2 \cdots\cdots (k-1) \cdot k}$$

Q. 트럼프에는 하트, 다이아몬드, 스페이드, 클로버와 같은 총 4가지 무늬(♥, ◆, ♠, ♣)의 카드가 있는데, 각 무늬는 1에서 10까지, 그리고 J, Q, K까지 포함해 총 13장의 카드를 한 벌로 가지고 있습니다. 트럼프의 무늬는 4가지 이기 때문에 4x13 = 52, 즉 합계 52장의 카드가 들어있습니다. 이 트럼프 더미에서 다섯 장의 카드를 동시에 뽑았을 때, 다섯 장 모두가 하트(♥)인 경우는 몇가지인지 답 하시오.

$$ {}_{n}C_{k} = \frac{n \cdot (n-1) \cdots\cdots (n-k+1)}{1 \cdot 2 \cdots\cdots (k-1) \cdot k} $$

↓

$$ {}_{13}C_{5} = \frac{13 \cdot (13-1) \cdots\cdots (13-5+1)}{1 \cdot 2 \cdots\cdots (5-1) \cdot 5} = \frac{13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 1287 $$

01-3

여사건

여사건

사건 A가 발생할 확률이 P일 때,
사건 A의 여사건 $\overline{A}$가 발생할 확률은 다음과 같다

$$P(\overline{A}) = 1 - P$$

Q. 52장의 트럼프 더미에서 4장의 카드를 동시에 뽑았을 때, 적어도 1장이 스페이드(♠)일 확률을 구하시오.

52장 – (♠)13장 = 39

$$1 - \frac{{}_{39}C_4}{{}_{52}C_4} = \frac{14498}{20825} \fallingdotseq 69.6\%$$

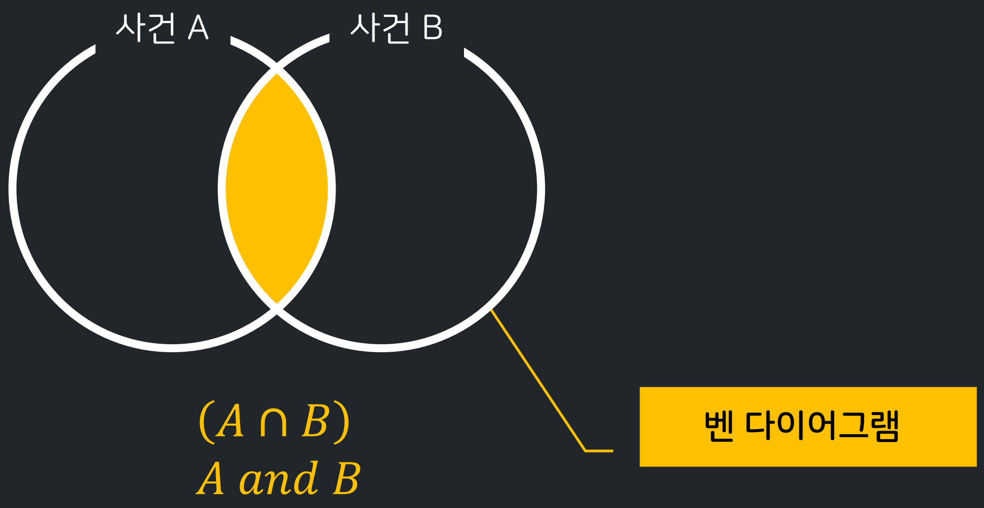
52장의 카드 중 4장

01-4

사건에 대한 조합

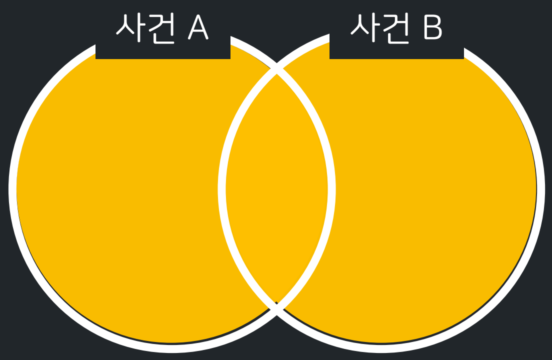
사건 A와 사건 B가 동시에 발생하는 사건 : $A \cap B$ $(A\ and\ B)$

$$P(A \cap B) = P(A)P(B)$$



사건 A와 사건 B 중에서 어느 한쪽이 발생할 사건 : $A \cup B\ (A \text{ or } B)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



$(A \cup B)$

$A\ or\ B$

Q. 52장의 트럼프 더미에서 다섯 장의 카드를 동시에 뽑은 후 원래대로 돌려놓습니다.
트럼프 더미의 카드를 뒤섞은 다음, 다시 한번 다섯 장의 카드를 동시에 뽑습니다.
이때, 두 번 연속으로 다섯 장의 카드 모두가 하트(♥) 무늬이거나 다이아몬드(◆) 무늬가 되는 확률을 구하시오.

A = (다섯 장 모두 ♥) ∪ (다섯 장 모두 ◆)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

하트(♥) 13장

P(다섯 장 모두 ♥) = $\dfrac{{}_{13}C_5}{{}_{52}C_5} = \dfrac{\frac{13\cdot 12\cdot 11\cdot 10\cdot 9}{\cancel{1}\cdot\cancel{2}\cdot\cancel{3}\cdot\cancel{4}\cdot\cancel{5}}}{\frac{52\cdot 51\cdot 50\cdot 49\cdot 48}{\cancel{1}\cdot\cancel{2}\cdot\cancel{3}\cdot\cancel{4}\cdot\cancel{5}}} = \dfrac{33}{66640} \fallingdotseq 0.0495\%$

다이아몬드(◆)도 동일

BUT '다섯 장 모두 ♥' 인 경우와 '다섯 장 모두 ◆' 인 경우는
동시에 발생 불가능

$\boldsymbol{P(A)}$ = P(다섯 장 모두 ♥) + P(다섯 장 모두 ◆) - P((다섯 장 모두 ♥) ∩ (다섯 장 모두 ◆))

$\boldsymbol{P(A)}$ = P((다섯 장 모두 ♥) + P(다섯 장 모두 ◆)) − 0 = $\dfrac{33}{33320} \fallingdotseq 0.0990\%$

발생 불가능 = 0

Q. 52장의 트럼프 더미에서 다섯 장의 카드를 동시에 뽑은 후 원래대로 돌려놓습니다.
트럼프 더미의 카드를 뒤섞은 다음, 다시 한번 다섯 장의 카드를 동시에 뽑습니다.
이때, 두 번 연속으로 다섯 장의 카드 모두가 하트(♥) 무늬이거나 다이아몬드(◆) 무늬가 되는 확률을 구하시오.

A_1 = 첫 번째 시도에서 다섯 장의 카드가 하트(♥) 무늬이거나, 다이아몬드(◆) 무늬인 사건

&

A_2 = 두 번째 시도에서 다섯 장의 카드가 하트(♥) 무늬이거나, 다이아몬드(◆) 무늬인 사건

↓

$A_{1\&2}$ = 두 번 연속으로 다섯 장의 카드가 하트(♥) 무늬이거나, 다이아몬드(◆) 무늬인 사건

↓

$$A_{1\&2} = A_1 \cap A_2$$

Q. 52장의 트럼프 더미에서 다섯 장의 카드를 동시에 뽑은 후 원래대로 돌려놓습니다.
트럼프 더미의 카드를 뒤섞은 다음, 다시 한번 다섯 장의 카드를 동시에 뽑습니다.
이때, 두 번 연속으로 다섯 장의 카드 모두가 하트(♥) 무늬이거나 다이아몬드(◆) 무늬가 되는 확률을 구하시오.

$$\boldsymbol{A_{1\&2}} = \boldsymbol{A_1} \cap \boldsymbol{A_2}$$

↓

$$\mathrm{P}(\boldsymbol{A_{1\&2}}) = \boldsymbol{P}(\boldsymbol{A_1}) \cdot \boldsymbol{P}(\boldsymbol{A_2}) = \left(\frac{33}{33320}\right)^2 ≒ 0.000098\%$$

02-1

확률변수 (이산확률변수, 연속확률변수)

확률 변수

무작위 실험을 했을 때, 특정 확률로 발생하는 각각의 결과를 수치적 값으로 표현하는 변수

Ex.

주사위를 던졌을 때 각 숫자가 나올 확률이 1/6으로 똑같다고 가정 ↔ 주사위의 각 숫자를 1/6 확률로 나오게 할 수 있다

어떤 변수 X를 $\mathrm{P}(X)$의 확률로 나오게 할 수 있다 ↔ 어떤 변수 X를 사용할때 $\mathrm{P}(X)$의 값을 구할 수 있다

이산확률변수

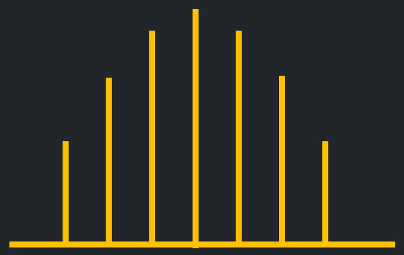
값이 연속되지 않고 셀 수 있을 만큼
뿔뿔이 흩어져 있음

Ex.

주사위 면에 쓰인 숫자

어떤 사건이 일어나는 시행 횟수

뿔뿔이 흩어진 값들



연속확률변수

값이 특정 범위 내에서 실수 형태로 존재하며,
소수점 이하까지 내려 감

Ex.

키나 몸무게

경과 시간

끊임없이 연속적으로 이어지는 값들

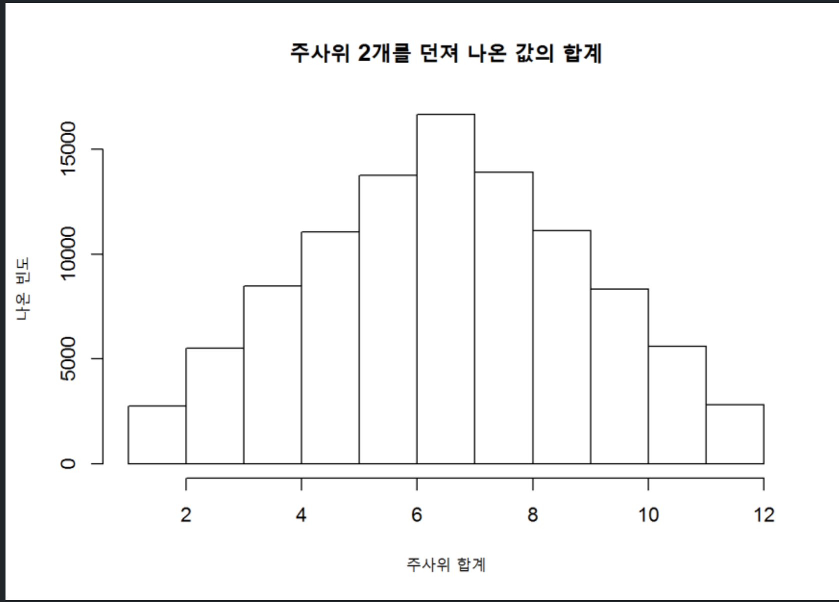
이산확률변수

$$\mathrm{P}(X) = \mathrm{f}(x)$$

어떤 사건의 이산확률 변수가 X일 때, 그에 대한 확률 P는 이산확률분포 $\mathrm{f}(x)$를 따른다

주사위 2개를 던져 숫자의 합을 확률변수 X_2라고할 때 이산확률분포 $\mathrm{f}(x)$는 다음과 같다

X_2	2	3	4	5	6	7	8	9	10	11	12
$P(X_2)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$



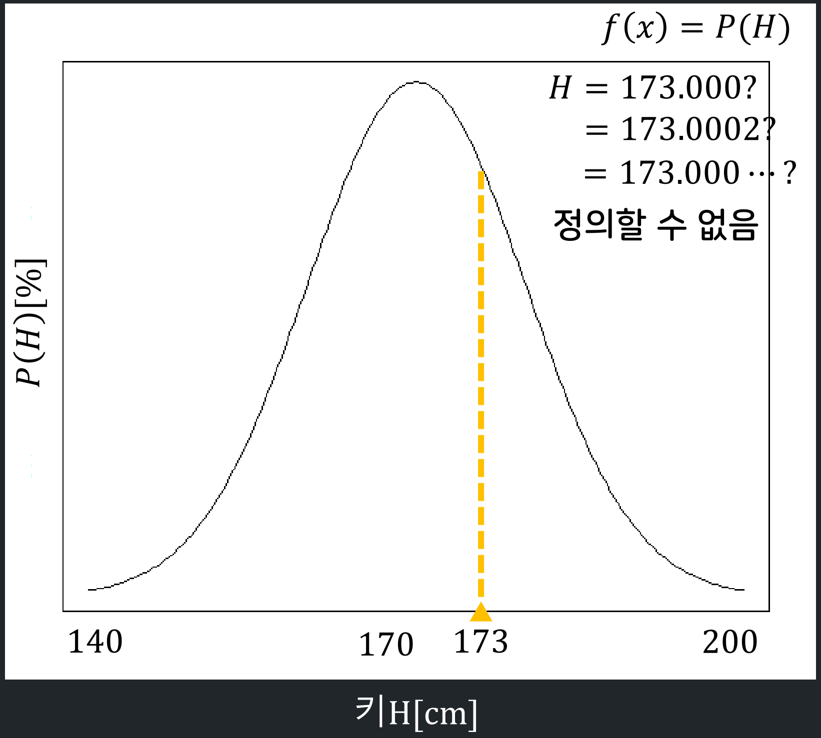
주사위의 개수를 늘리면 이산확률분포는 종형곡선의 모양에 가까워 짐

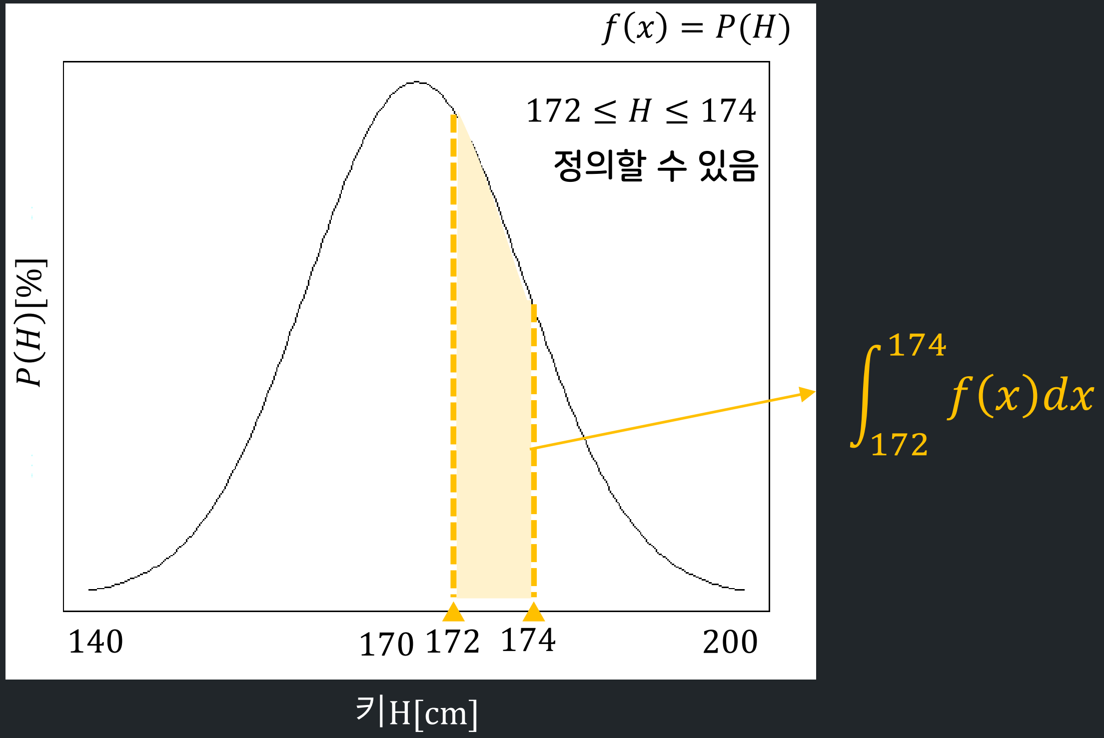
주사위를 던져보는 시행을 무한으로 하는 극한에서는 종형곡선을 정규분포라고 부름

연속확률변수

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

어떤 사건의 연속확률 변수가 X 일 때, 그 에 대한 확률 P는 연속확률 분포 f(x)를 지정한X의 구간 안에서 적분한 값과 같다.





03-1

결합확률과 조건부확률

결합확률

사건 A와 사건 B가 서로 독립된 사건일 때,
두 사건의 결합확률은 다음과 같다

$$P(A \cap B) = P(A, B) = P(A)P(B)$$

서로 독립적인 사건 A와 B가 동시에 일어날 확률

조건부확률

사건 B가 일어났을 때,
사건 A가 일어날 조건부 확률은 다음과 같다

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

사건 B가 일어난 후에 사건 A가 일어날 확률

Q. 하나의 주사위를 두 번 던진다고 가정합니다. 사건 A는 주사위를 두 번 던져서 나온 숫자의 합이 8 이상일 사건이고, 사건 B는 첫 번째 주사위를 던졌을 때 나오는 숫자가 5가 되는 사건이라고 할 때, 결합확률 $P(A,B)$와 조건부확률 $P(A|B)$ 를 구하시오

결합확률

$$P(A \cap B) = P(A,B) = P(A)P(B)$$

첫 번째 주사위를 던졌을 때 5가 나오고,
두 번째 주사위를 던졌을 때 3이상의 수가 나올 확률

$$P(A \cap B) = \frac{1}{6} \times \frac{4}{6} = \frac{1}{9}$$

조건부확률

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

첫 번째 주사위를 던졌을 때 5가 나오고,
두 번째 주사위를 던졌을 때 3이상의 수가 나올 확률

$P(B) =$ 첫 번째 주사위를 던졌을 때 5가 나올 확률

$$P(B) = \frac{1}{6}$$

$$P(A|B) = \frac{\frac{1}{9}}{\frac{1}{6}} = \frac{2}{3}$$

인공지능에서는?

인공지능 에서는?

조건부 확률

- 관찰되는 데이터에서 그것이 발생한 원인 혹은 그 데이터를 발생시킬 조건이 여러 개일 때 그 중 제일 가능성이 높은 것을 찾아냄
- 현실 세계에서 명확한 규칙을 만들어 내기가 어려워 통계적인 방법을 쓸 때 사용

통계 모델

- 100% 정확한 답을 주지는 않지만 여러 후보들 중에서 상대적으로 확률이 높은 것을 알 수 있기 때문에 더 복잡한 문제에 사용 가능
- 데이터가 모이고 모일수록 통계적 자료가 많아 짐에 따라 인공지능은 점점 더 정교해짐
- Ex. Google 자동번역시스템, Apple의 Siri

이산확률 분포

- 어떤 현상에 대한 관측 결과들을 이산확률변수로 취급하고, 이에 대한 이산확률분포를 구할 수 있다면, 다음에 일어날 사건에 대한 확률을 과거의 데이터로부터 추측할 수 있다.

연속확률 분포

- 적절한 연속확률분포를 선택한다면 적은 수의 시행만으로도 앞으로 일어날 사건의 확률을 상당히 높은 정확도로 추측할 수 있다.

[참고] "인공지능, 조건부 확률 by 우혁 김," *Prezi*, 2018년 11월 18일 수정, 2020년 4월 30일 접속,
https://prezi.com/bbsbvfpbsy4e/presentation/.

THANK YOU