STAT 131
12 May 20
(lecture)
①

$$P(\mathcal{I}_1 \in A_1 \overset{\text{and}}{,} \mathcal{I}_2 \in A_2, \dots, \mathcal{I}_n \in A_n) = (*)$$

if $\mathcal{I}_i$ are <u>not</u> independent

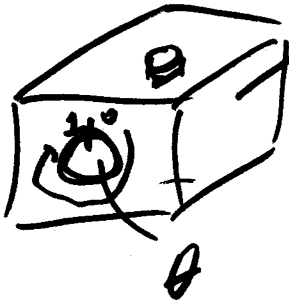
$$(*) = P(\mathcal{I}_1 \in A_1) \cdot P(\mathcal{I}_2 \in A_2 \mid \mathcal{I}_1 \in A_1) \cdot P(\mathcal{I}_3 \in A_3 \mid \mathcal{I}_1 \in A_1, \mathcal{I}_2 \in A_2) \cdots P(\mathcal{I}_n \in A_n \mid \mathcal{I}_1 \in A_1, \dots, \mathcal{I}_{n-1} \in A_{n-1})$$

if instead the $\mathcal{I}_i$ <u>are</u> independent,

$$(*) = P(\mathcal{I}_1 \in A_1)\, P(\mathcal{I}_2 \in A_2) \cdots P(\mathcal{I}_n \in A_n)$$

---

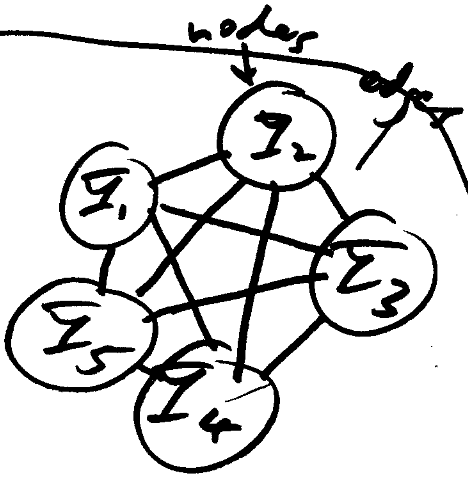
in machine learning, pretending that $(\mathcal{I}_1, \dots, \mathcal{I}_n)$ are independent when they're not is called <u>naive Bayes learning</u>.



1 or 0 (if know θ
IID
Bernoulli(θ))

$(X_i | \underline{\theta}) \overset{IID}{\sim}$

if θ not known, X_i <u>dependent</u>



Graph theory

Directed Acyclic Graph (DAG)

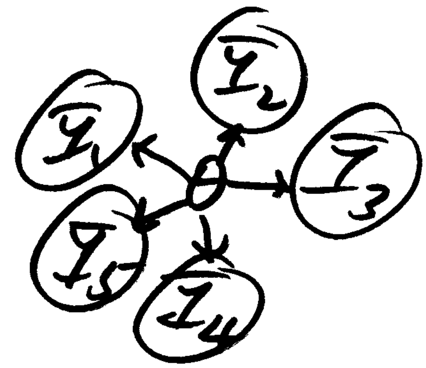
Case 1: θ unknown

X_i dependent

$n = 5$ nodes

$\binom{n}{2} = \binom{5}{2} = 10$ edges

$\frac{n(n-1)}{2} = \underline{\underline{O}}(n^2)$

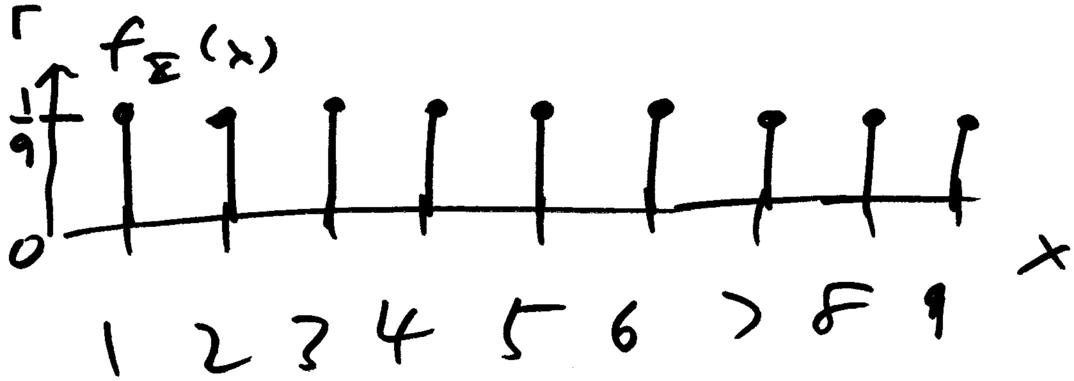


Case 2: θ known

X_i cond. ind.

$n = 5$ nodes

only $n = 5 = \underline{\underline{O}}(n)$ edges



$$F_Y(y) = 1 - F_X\left(\frac{1}{y}\right)$$

$$\underbrace{f_Y(y)}_{\text{PDF}} = \frac{d}{dy}\underbrace{F_Y(y)}_{\text{CDF}} = \frac{d}{dy}\left[1 - F_X\left(\frac{1}{y}\right)\right]$$

$$= -\underbrace{\frac{d}{dy}F_X\left(\frac{1}{y}\right)}$$

$$= -f_X\left(\frac{1}{y}\right)\underbrace{\overbrace{(-y^{-2})}^{y^{-1}}}_{\text{chain rule}}$$

$$f_Y(y) = f_X\left(\frac{1}{y}\right) \cdot \frac{1}{y^2}$$