Mathematics for AI

Lecture 22

Probabilistic Independence,

Question

- Pr(rain in 1hr) = .15
- Pr(rain in 1hr | cloudy now) = .5

- Pr(rain in 1hr) = .15
- Pr(rain in 1hr | I failed the math exam) = ?

Question

- Pr(rain in 1hr) = .15
- Pr(rain in 1hr | cloudy now) = .5

- Pr(rain in 1hr) = .15
- Pr(rain in 1hr | I failed the math exam) = .15

Probabilistic Independence

K. N. Toosi
University of Technology

- $\Pr(M = m \mid N = n) = \Pr(M = m)$ for all m, n

Probabilistic Independence

- $Pr(M = m \mid N = n) = Pr(M = m)$ for all m, n

  $P(M = m, N = n) / Pr(N = n) = Pr(M = m)$

Probabilistic Independence

- $Pr(M = m \mid N = n) = Pr(M = m)$ for all m, n

  $P(M = m, N = n) / Pr(N = n) = Pr(M = m)$

  $\Rightarrow P(M = m, N = n) = Pr(N = n)\ Pr(M = m)$

Probabilistic Independence

K. N. Toosi
University of Technology

- $Pr(M = m \mid N = n) = Pr(M = m)$ for all m, n

  $P(M = m, N = n) / Pr(N = n) = Pr(M = m)$

  $\Rightarrow P(M = m, N = n) = Pr(N = n)\ Pr(M = m)$

- What does independence mean?
  - does "having a rainfall" depend on "people using umbrellas"?

Probabilistic Independence

K. N. Toosi
University of Technology

- $\Pr(M = m \mid N = n) = \Pr(M = m)$ for all m, n

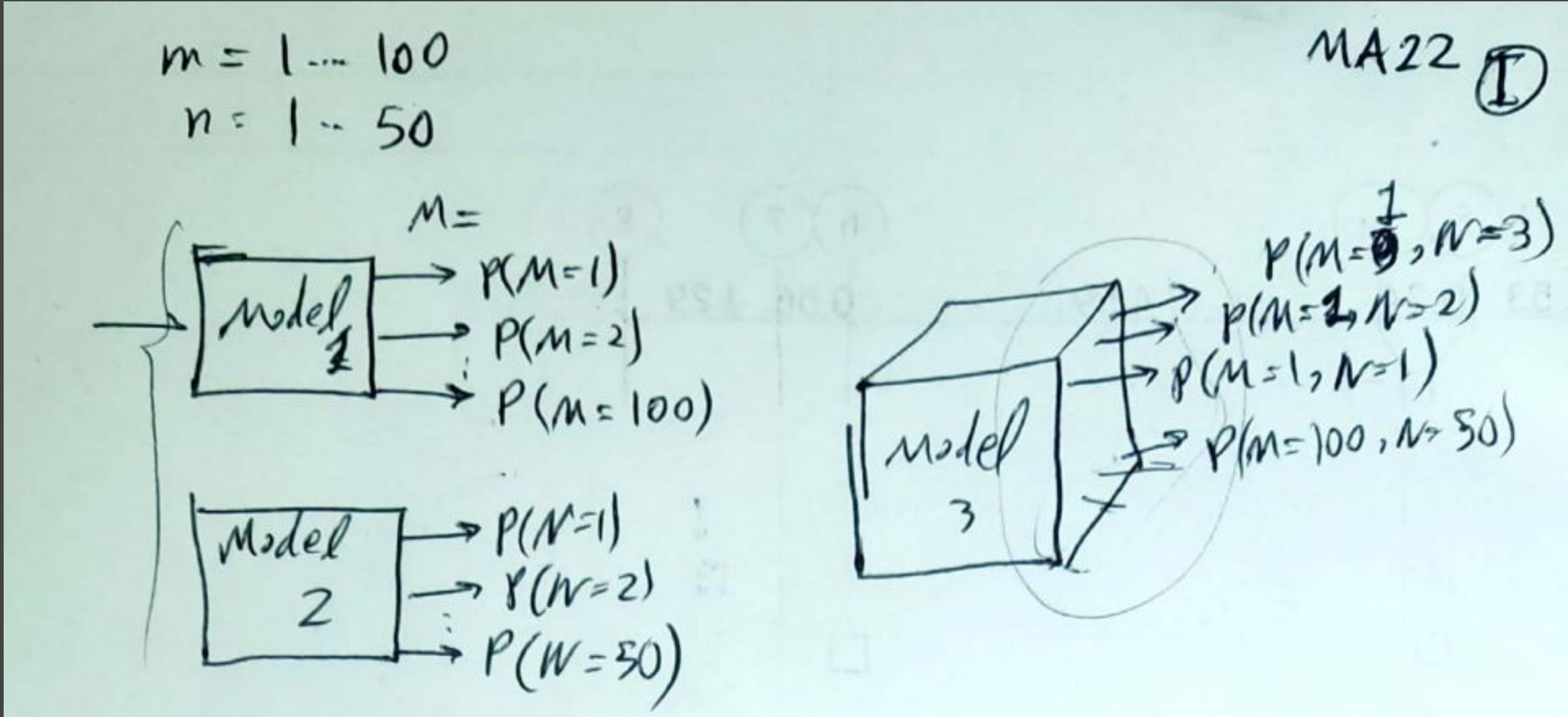
  $P(M = m, N = n) / \Pr(N = n) = \Pr(M = m)$

  $\Rightarrow P(M = m, N = n) = \Pr(N = n)\ \Pr(M = m)$

- Continuous case:

  $p(y \mid x) = p(y) \quad \Rightarrow \quad p(x,y) = p(x)\, p(y)$

Independence and model complexity



More than two variables

K. N. Toosi
University of Technology

- $p(x_1, x_2, x_3, \ldots, x_m)$
- Pairwise independence
  - Every pair of variables x_i, x_j are independent
- Mutual Independence
  - $p(x_i$ | any **subset** of other variables $) = p(x_i)$
  - $p(x_1, x_2, \ldots, x_m) = p(x_1)\, p(x_2) \ldots p(x_m)$

More than two variables

$$p(\underbrace{x_1, x_2, \ldots, x_n}_{Y}) = p(x_1 \mid x_2, x_3, \ldots x_n)\, p(x_2, x_3, \ldots, x_n)$$

$$= p(x_1 \mid x_2, \ldots\ x_n)\, p(x_2 \mid x_3 \ldots x_n)\, p(x_3 \ldots x_n)$$

$$= p(x_1 \mid x_2 \ldots x_n)\, p(x_2 \mid x_3 \ldots x_n) \cdot p(x_3 \mid x_4 \ldots x_n)$$

$$\ldots p(x_{n-1} \mid x_n)\, p(x_n)$$

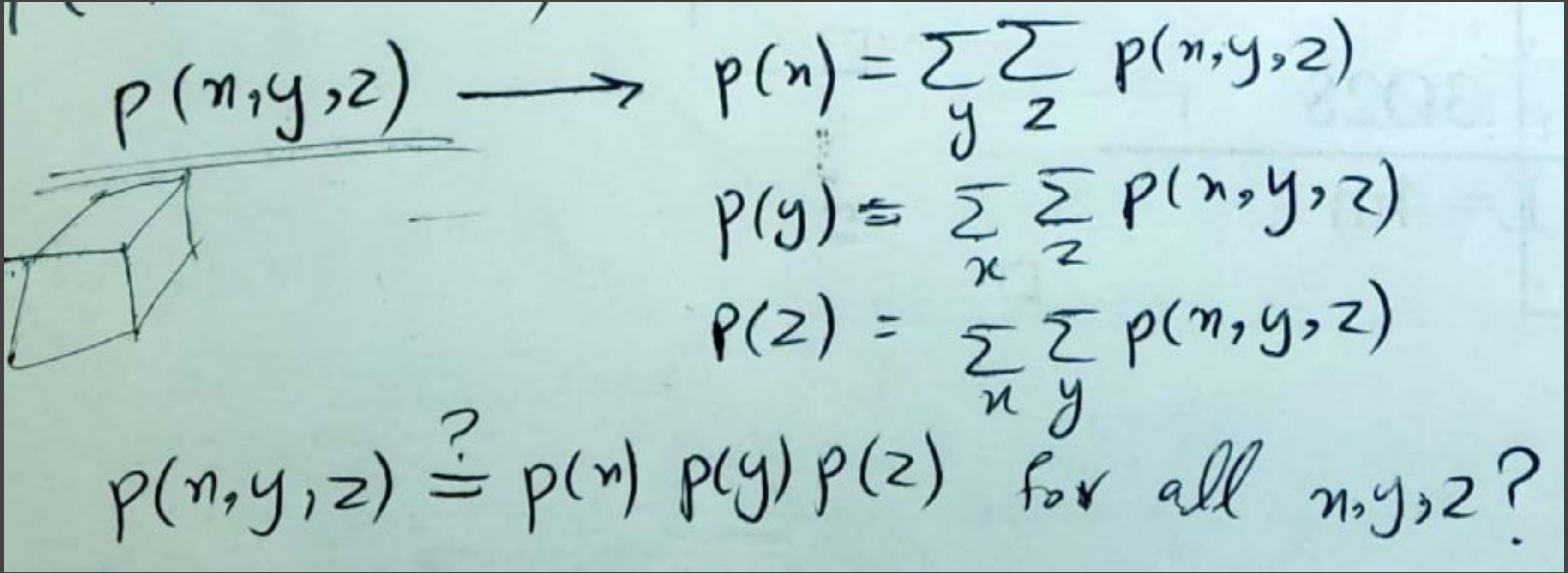
Chain Rule

$x_1, x_2, \ldots x_n$ idependent:

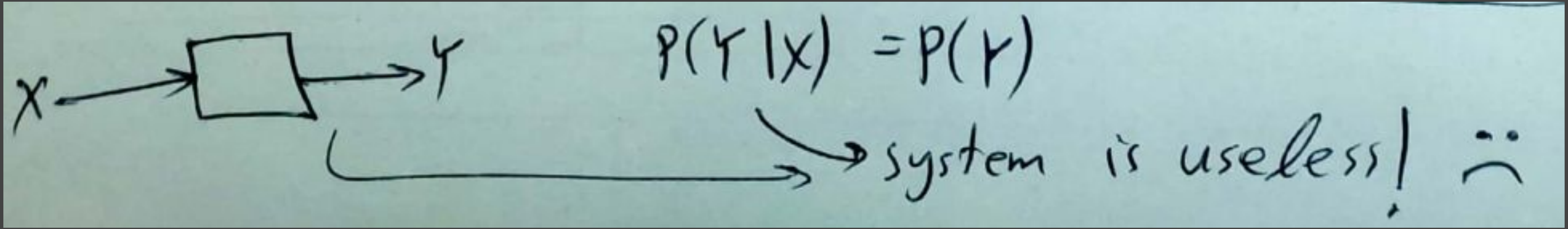
$$p(x_1, x_2, \ldots, x_n) = p(x_1)\, p(x_2) \ldots p(x_{n-1})\, p(x_n)$$

testing independence

K. N. Toosi
University of Technology



Sometimes dependence is desired



Example

- Are "Having a cloudy morning" and "getting wet" dependent?
- $P(W \mid C) = P(W)$?

Example

- Are "Having a cloudy morning" and "getting wet" dependent?
- $P(W \mid C) \neq P(W)$

Example

- Knowing that we had a rainfall
  - Are "Having a cloudy morning" and "getting wet" dependent?

Conditional Independence

- Knowing that we had a rainfall
  - Are "Having a cloudy morning" and "getting wet" dependent?
  - $P(W \mid R, C) = P(W \mid R)$

Conditional Independence

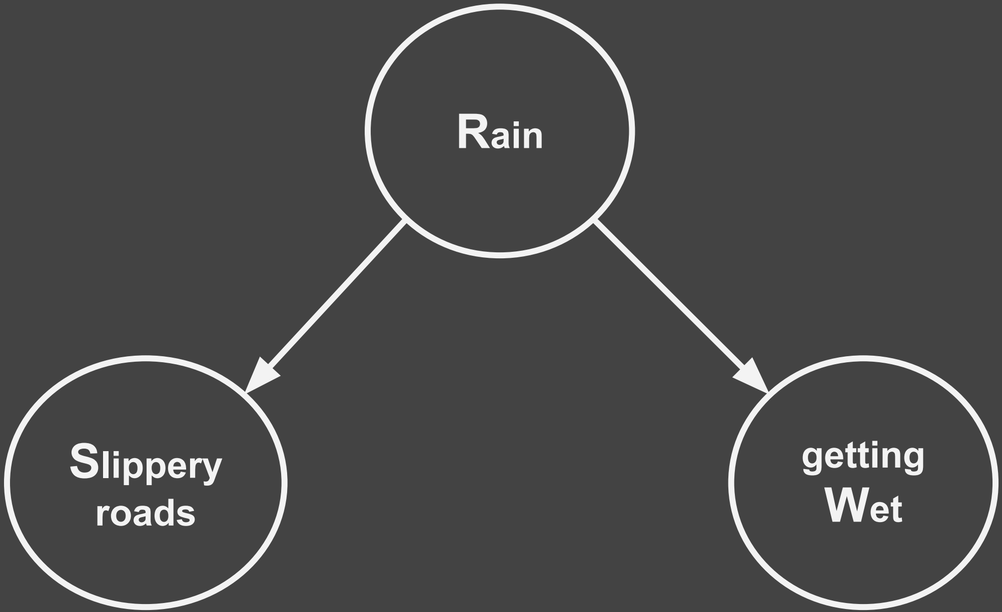
- Knowing that we had a rainfall
    - Are "Having a cloudy morning" and "getting wet" dependent?
    - $P(W \mid R, C) = P(W \mid R)$
    - $P(W, C \mid R) = P(W \mid R)\ P(C \mid R)$

Conditional Independence

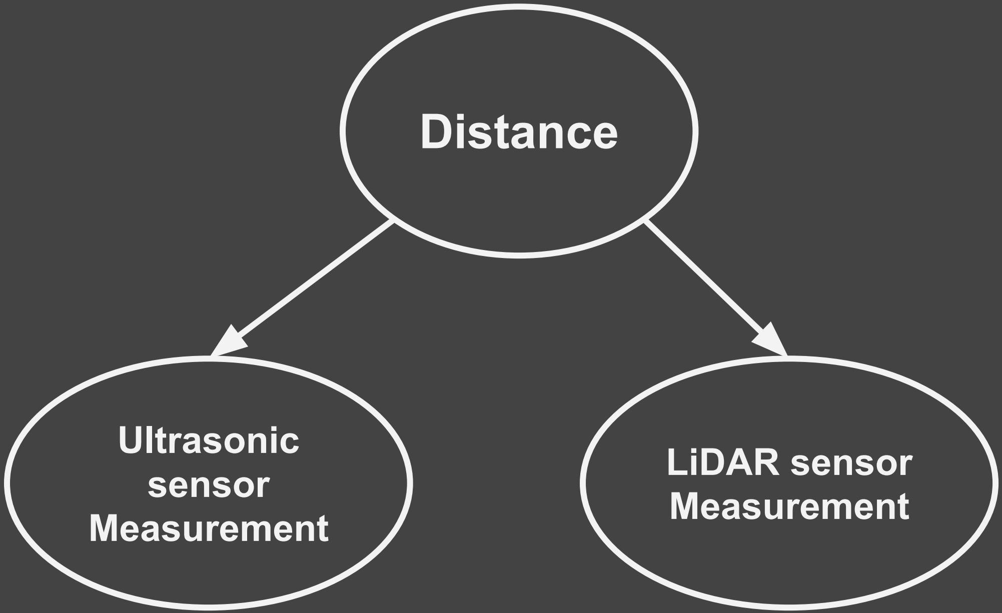
- Knowing that we had a rainfall
  - Are "Having a cloudy morning" and "getting wet" dependent?
  - $P(W \mid R, C) = P(W \mid R)$
  - $P(W, C \mid R) = P(W \mid R)\ P(C \mid R)$
- W and C are conditionally independent given R

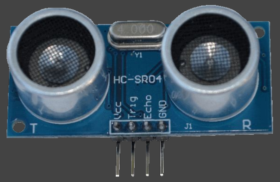
Conditional Independence

K. N. Toosi
University of Technology



Conditional Independence



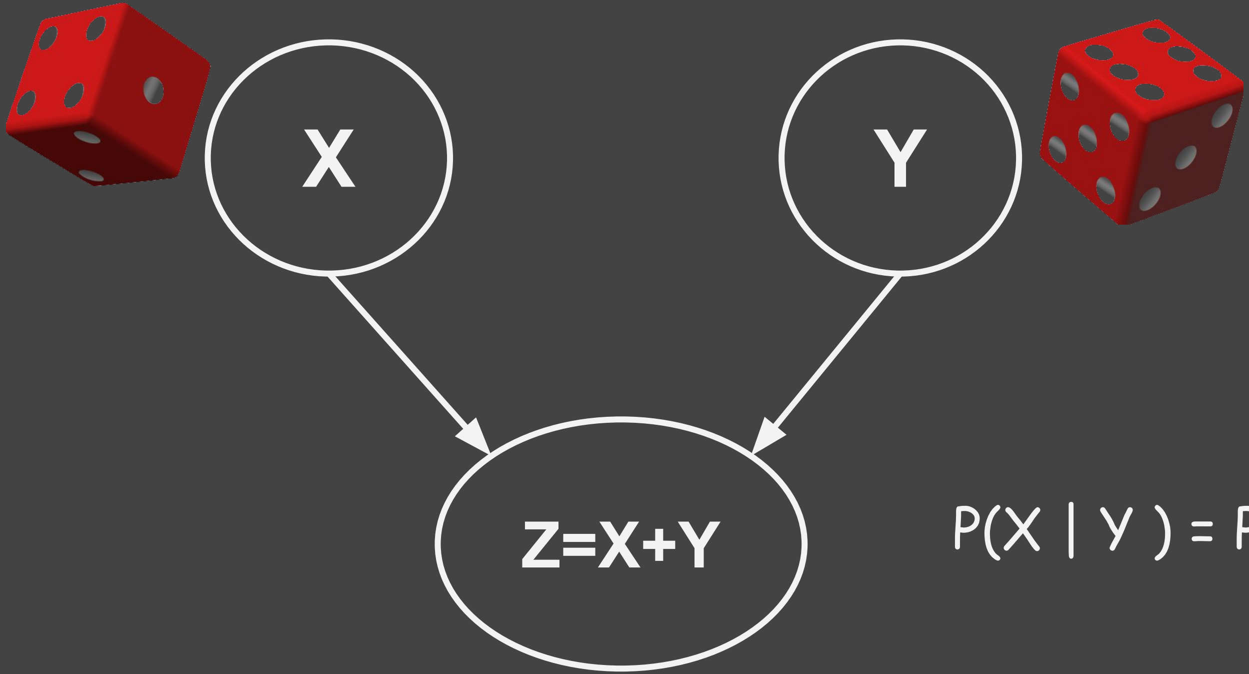


Conditioning can destroy independence

$$P(X \mid Y) = P(X)$$

Conditioning can destroy independence

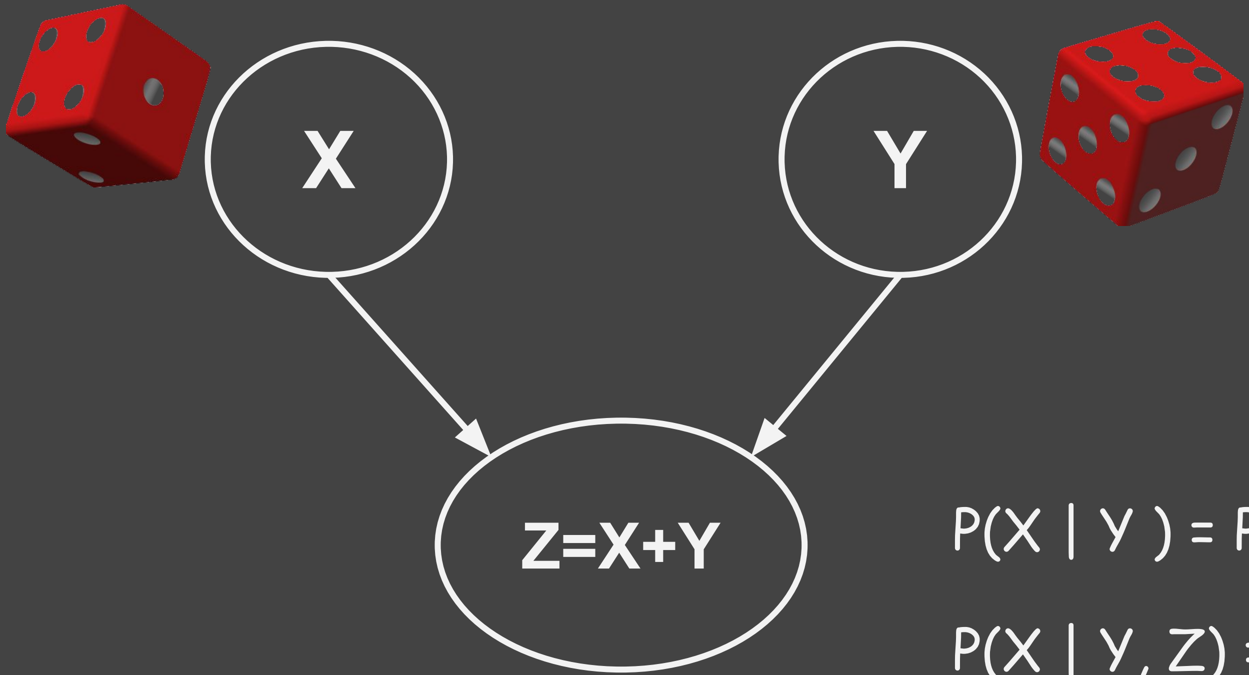
K. N. Toosi
University of Technology



$P(X \mid Y) = P(X)$

Conditioning can destroy independence

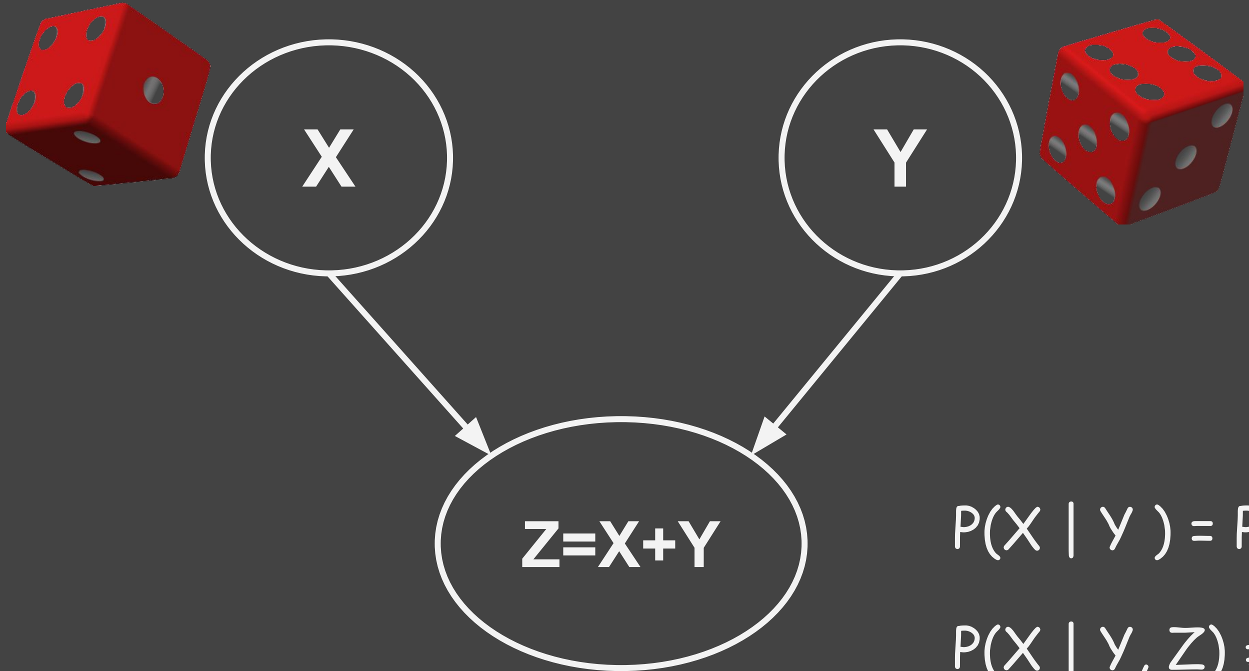
K. N. Toosi
University of Technology



$$P(X \mid Y) = P(X)$$

$$P(X \mid Y, Z) = P(X \mid Z)\ ?$$

Conditioning can destroy independence

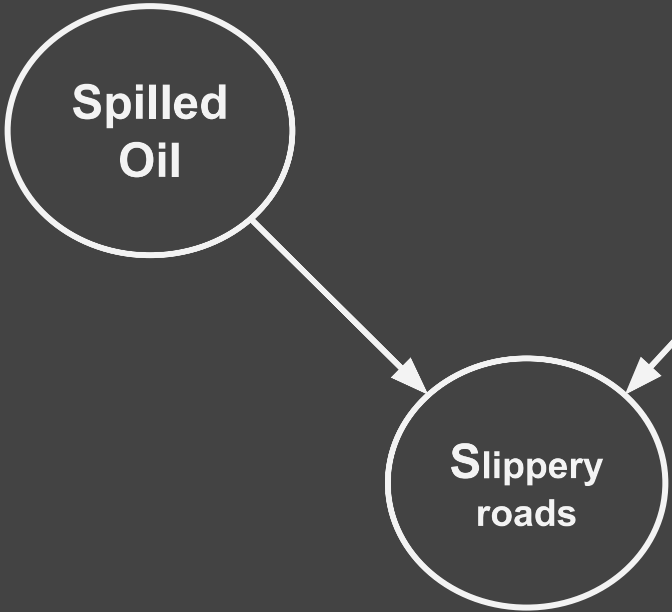


$$P(X \mid Y) = P(X)$$

$$P(X \mid Y, Z) \neq P(X \mid Z)$$

K. N. Toosi
University of Technology

Conditioning can destroy independence



Conditional independence reduce complexity?

- Tabular representation P(C,R,W)
- General case: how many independent parameters in general?

Conditional independence reduce complexity?

- Tabular representation P(C,R,W)
- General case: how many independent parameters in general? 7

Conditional independence reduce complexity?

- Tabular representation P(C,R,W)
- General case: how many independent parameters in general? 7
- Fully independent case: P(C,R,W) = P(C) P(R) P(W)
    - How many parameters?

Conditional independence reduce complexity?

- Tabular representation P(C,R,W)
- General case: how many independent parameters in general? 7
- Fully independent case: P(C,R,W) = P(C) P(R) P(W)
  - How many parameters? 3

Conditional independence reduce complexity?

K. N. Toosi University of Technology

- Tabular representation P(C,R,W)
- General case: how many independent parameters in general? 7
- Fully independent case: P(C,R,W) = P(C) P(R) P(W)
  - How many parameters? 3
- Conditionally independent: P(C,R,W): P(W | R, C) = P(W | R)
  - How many parameters?

Conditional independence reduce complexity?

- Conditionally independent: P(C,R,W) = P(W | R, C) = P(W | R)
  - How many parameters?
  - P(C,R,W) = P(W | C, R) P(C,R) = P(W | R) P(C,R)

Conditional independence reduce complexity?

- Conditionally independent: $P(C,R,W) = P(W \mid R, C) = P(W \mid R)$
  - How many parameters?
  - $P(C,R,W) = P(W \mid C, R)\, P(C,R) = P(W \mid R)\, P(C,R)$
  - $P(W \mid R)$:
  - $P(C,R)$:

Conditional independence reduce complexity?

- Conditionally independent: P(C,R,W) = P(W | R, C) = P(W | R)
  - How many parameters?
  - P(C,R,W) = P(W | C, R) P(C,R) = P(W | R) P(C,R)
  - P(W | R): P(W = 0| R = 0), P(W = 0| R = 1) : 2 parameters
  - P(C,R):

Conditional independence reduce complexity?

K. N. Toosi
University of Technology

- Conditionally independent: P(C,R,W) = P(W | R, C) = P(W | R)
  - How many parameters?
  - P(C,R,W) = P(W | C, R) P(C,R) = P(W | R) P(C,R)
  - P(W | R): P(W = 0| R = 0), P(W = 0| R = 1) : 2 parameters
  - P(C,R): 3 parameters

Conditional independence reduce complexity?

- Conditionally independent: $P(C,R,W) = P(W \mid R, C) = P(W \mid R)$
  - How many parameters?
  - $P(C,R,W) = P(W \mid C, R)\, P(C,R) = P(W \mid C)\, P(C,R)$
  - $P(W \mid R)$: $P(W = 0 \mid R = 0)$, $P(W = 0 \mid R = 1)$: 2 parameters
  - $P(C,R)$: 3 parameters
  - $P(C,R,W) = P(W \mid C, R)\, P(C,R)$: 5 parameters

Conditional independence reduce complexity?

- Tabular representation P(C,R,W)
- General case: 7 parameters
- Fully independent case: 3 parameters
- Conditionally independent: 5 parameters