REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mohamed Khider-BISKRA.

Faculté des sciences exactes et des sciences de la nature et de la vie

Département d'informatique

Machine Learning (ML)

Dr. TIBERMACINE Ahmed

Content Summary

Chapter 5:
Bayesian network

INTRODUCTION

- Bayesian networks= belief networks or probabilistic graphical models,
- Powerful tools for representing and reasoning about uncertainty and probabilistic dependencies in complex systems.
- Explore the fundamentals of BNs, their graphical representation, and how to perform probabilistic inference using them.

LEARNING OBJECTIVES

- Understand the fundamental concepts of BNs.
- Learn how to represent and manipulate BNs.
- Explore probabilistic reasoning and inference in BNs.
- Gain insight into applications of BNs in various fields

Bayesian network

- **Probabilistic reasoning in Artificial intelligence**
  - **Uncertainty**

- Certainty involves high confidence (close to 1 or 0) in a statement's truth.
- Uncertainty recognizes varying degrees of belief (probabilities between 0 and 1).
- In cases of uncertain knowledge, probabilistic reasoning is necessary.

Bayesian network

- **Probabilistic reasoning in Artificial intelligence**
  - **Probabilistic reasoning**
- Probabilistic reasoning involves representing knowledge by using probability to express uncertainty.
- In probabilistic reasoning, we combine probability theory with logic to handle the uncertainty.
- Certainty is often unconfirmed in various scenarios:
  - "It will rain today “;
  - "behavior of someone for some situations ";
  - "A match between two teams or two players “.

Bayesian network

- **Probability Basics**
  - **Probability**
- Probability can be defined as a chance that an uncertain event will occur.
- It is the numerical measure of the likelihood that an event will occur.
- The value of probability always remains between 0 and 1 that represent ideal uncertainties.

- $\text{probability of an uncertain event} = \frac{\text{Number of favorable outcomes}}{\text{Total number of possible outcomes}}$

Bayesian network

Probability Basics

Probability: Example 1

- Let's consider the rolling of a fair six-sided die. The event A is rolling a 4.
- $\mathrm{P(A)} = \frac{\text{Number of favorable outcomes for rolling a 4}}{\text{Total number of possible outcomes}}$
- There is only one way to roll a 4 (favorable outcome), and there are six possible outcomes (rolling a 1, 2, 3, 4, 5, or 6).
- $P(rolling\ \ 4) = \frac{1}{6}$
- So, the probability of rolling a 4 on a fair six-sided die is: 1/6, or approximately 0.167.

Bayesian network

- **Probability Basics**
  - **Probability: Example 2**
- The letters of the word "MATHEMATICS" are placed in a bag.
- What is the probability of drawing the letter "T" from the bag?
- **Solution:** The total number of letters = 11
- As there are two T's placed in the bag thus, the number of favorable outcomes = 2.
- P(T) = 2 / 11 = 0.182
- **Answer:** The probability of drawing the letter "T" is 0.182

Bayesian network

❖ Probability Basics

○ Probability concept:

- **Sample Space (Ω):**

Example: when rolling a fair six-sided die, the sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$.

- **Event (A):**

Example: "rolling an even number" can be represented as $A = \{2, 4, 6\}$.

- **Complement of an Event (A'):** represents all outcomes that are not in A.

In the example of rolling an even number, A' would represent rolling an odd number.

Bayesian network

- **Probability Basics**
  - **Probability concept:**
- **Random Experiments:**

Examples: rolling a die, flipping a coin.

- **Random variables:**
  - discrete (e.g., the outcome of a dice roll)
  - or continuous (e.g., the temperature in a given location).

Bayesian network

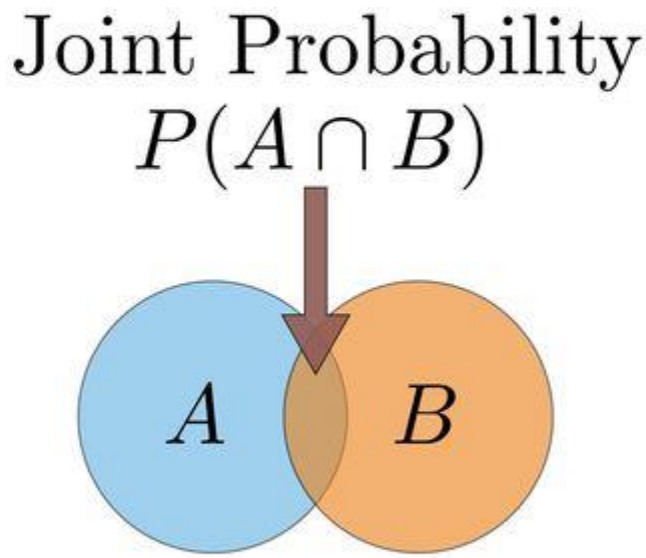
- **Probability Basics**
- **Marginal probability**

- Probability of a single event occurring on its own, without considering any other variables.
- **Example:**

if you're rolling a fair six-sided die, the margin probability of rolling a 4 would be 1/6 (P (4) = 1/6), as there is one favorable outcome (rolling a 4) out of six possible outcomes.

❖ Probability Basics

- **Joint probability** P(A and B)
- the probability of two or more events occurring simultaneously



- **Example:** if you want to find the joint probability of flipping a coin and rolling a die and getting both a "heads" (H) and a "4", it would be: P(H and 4) = P(H) * P(4) = (1/2) * (1/6) = 1/12.

Bayesian network

- **Probability Basics**
  - **Joint probability**
- The general formula for the joint probability of more than two events:

  - $P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$
- For more than three:

$$P(A1 \cap A2 \cap \ldots \cap An) = P(A1) \times P(A2|A1) \times P(A3|A1 \cap A2) \times \ldots \times P(An|A1 \cap A2 \cap \ldots \cap An-1)$$

- .

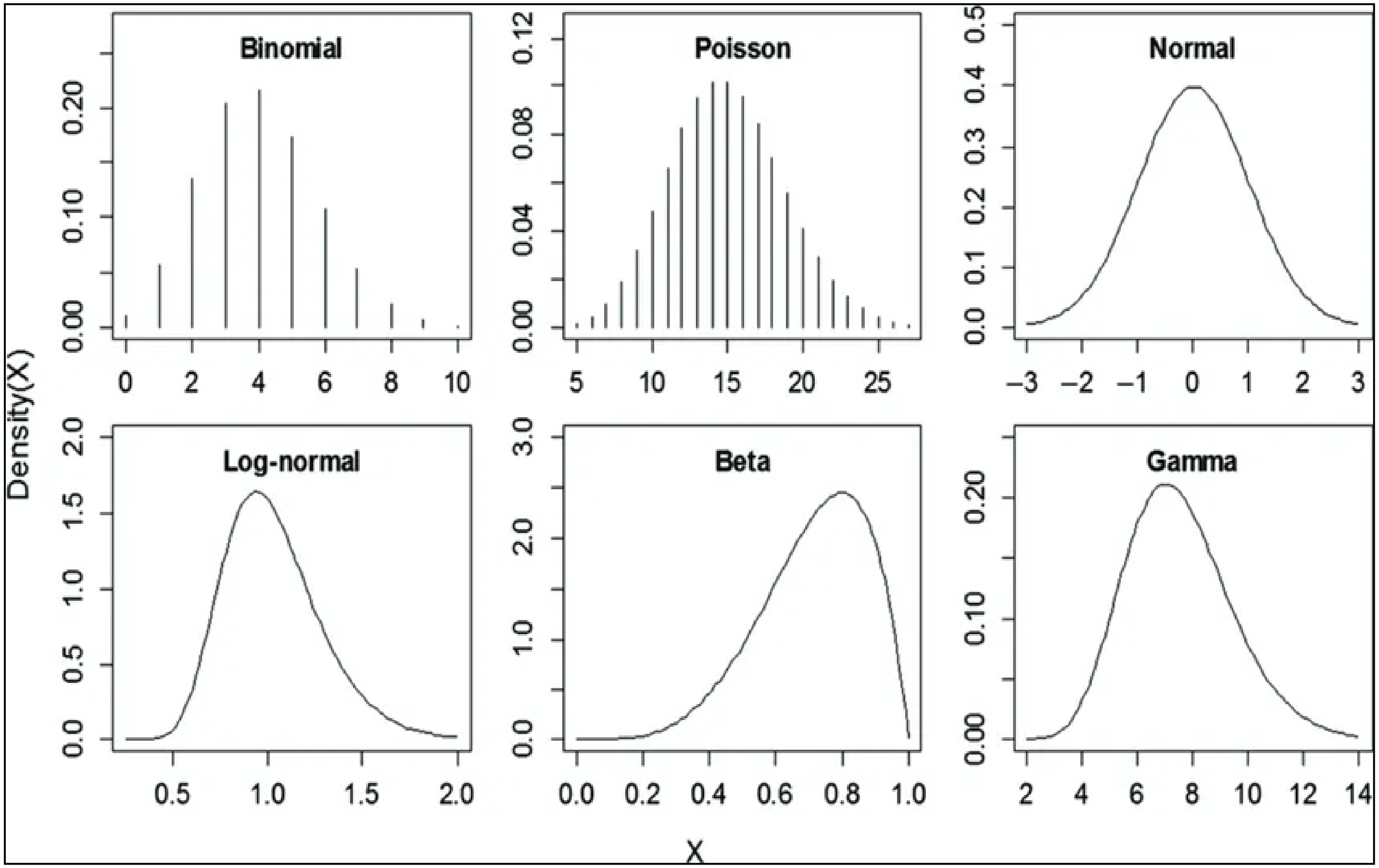
Bayesian network

- **Probability Basics**
  - **Probability Distribution**

- It describes how the probabilities of possible outcomes are spread over a set of events.
- Probability distributions can be:
  - **Continuous**:
  - **Discrete:**

Bayesian network

- **Probability Basics**
  - **Probability Distribution**



Bayesian network

❖ Probability Basics

○ Probability Distribution

- Example:

let's examine a discrete probability distribution representing the number of heads when flipping a fair coin twice.

- The PMF for this scenario is denoted as $f(x)=P(X=x)$, where X represents the number of heads in two flips.
- The specific probabilities for obtaining 0, 1, or 2 heads are as follows:

Bayesian network

- **Probability Basics**
  - **Probability distribution**
- For X=0, there is only one way to get 0 heads (both flips are tails): $\frac{1}{2}\times\frac{1}{2}=\mathbf{\frac{1}{4}}$
- For X=1, there are two ways to get 1 head (HT or TH): $2\times\frac{1}{2}\times\frac{1}{2}=\mathbf{\frac{1}{2}}$
- For X=2, there is only one way to get 2 heads (both flips are heads): $\frac{1}{2}\times\frac{1}{2}=\mathbf{\frac{1}{4}}$.
- So, the probability distribution is: $(0)=\frac{1}{4}$ $f(1)=\frac{1}{2}$ $(2)=\frac{1}{4}$.
- As expected, the sum of these probabilities is $\frac{1}{4}+\frac{1}{2}+\frac{1}{4}=1$, which ensures that all possible outcomes are covered.

Bayesian network

- **Probability Basics**
  - **Total Probability Law**

- The Law of Total Probability is a theorem that allows you to find the probability of an event by considering all possible ways that the event can occur.
- It is often expressed as follows: If B_1, B_2,…, B_n form a partition of the sample space S, then for any event A, the law of total probability states that:

$$P(A) = \sum_{i=1}^{n} P(A|B_i) * P(B_i)$$

  - P(A) is the marginal probability of event A.
  - P(A|Bi) is the conditional probability of event A given that event Bi has occurred.
  - P(Bi) is the probability of event Bi occurring.
  - n is the number of different scenarios or conditions.

Bayesian network

- **Probability Basics**
- **Chain rule**
- It helps find the joint probability of two variables, X and Y. It's expressed as

$$P(X, Y) = P(X|Y) * P(Y).$$

- Extending this to three random variables:

$$P(X, Y, Z) = P(X|Y, Z) * P(Y|Z) * P(Z).$$

- For n random variables:

$$P(X_1, X_2, \ldots, X_n) = P(X1|X_2, X_3, \ldots, X_n) * P(X_2|X_3, \ldots, X_n) * \ldots * P(X_{n-1}|X_n) * P(X_n).$$

Bayesian network

Conditional probability

Definition:

- the probability of an event occurring given that another event has already occurred.
- It is denoted as P(A|B), where A is the event of interest, and B is the conditioning event.
- The formula for conditional probability is:

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

  - **P(A|B):** Conditional probability of event A given event B.
  - **P(A ∩ B):** Probability of both events A and B occurring (joint probability).
  - **P(B):** Probability of event B occurring.

Bayesian network

- **Conditional probability**
  - **Definition: Example**
- Let's imagine a bag with 4 red marbles and 2 green marbles. If we want to find the conditional probability of picking a green marble on the second draw given that the first marble drawn was red, we can use the formula:
- $P(\text{Green on 2nd draw} \mid \text{Red on 1st draw}) = \dfrac{P(\text{Red on 1st draw and Green on 2nd draw})}{P(\text{Red on 1st draw})}$
- Here:
- P(Red on 1st draw) is the probability of drawing a red marble first, which is $\frac{4}{6}$ since there are 4 red marbles out of a total of 6 marbles initially.

- .

Bayesian network

- **Conditional probability**
  - **Definition: Example**

- P(Red on 1st draw and Green on 2nd draw) is the probability of drawing a red marble first and a green marble second.

- After drawing a red marble, there are now 5 marbles left, and 2 of them are green. So, this probability is $\frac{4}{6} \times \frac{2}{5}$.
- Now, we can calculate the conditional probability:
- $P(\text{Green on 2nd draw} \mid \text{Red on 1st draw}) = \frac{\frac{4}{6}\times\frac{2}{5}}{\frac{4}{6}} = \frac{\frac{4}{15}}{\frac{4}{6}} = \frac{6}{15} = 0.4 = 40\%$

- .

Bayesian network

❖ Conditional probability

○ Independence Events

- It is the occurrence of one event does not affect the likelihood or outcome of another.

P(A and B) = P(A) * P(B)).

Example:

- when rolling a die, the outcome of one roll is independent of the outcome of any other roll.
- The probability of getting a 4 on one roll doesn't change based on the results of previous or subsequent rolls.
- If you flip a coin twice and roll a die, these events are typically independent. The probability of getting "heads" twice and rolling a "4" would be:
- P(H and H and 4) = P(H) * P(H) * P(4) = (1/2) * (1/2) * (1/6) = 1/24

Bayesian network

- **Conditional probability**
  - **Dependence Events**

- In probability, dependence events signify the connection between two or more events.

- When the occurrence of one event influences the likelihood or outcome of another, these events are classified as dependent.

- Mathematically, if A and B are dependent events, the conditional probability of A given B is expressed as:

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

Bayesian network

- **Conditional probability**
  - **Dependence Events**
- Taking the example of drawing marbles.
- Drawing a red marble on the first draw affects the likelihood of drawing a green marble on the second draw.
- These events are dependent as the first draw changes the composition of the remaining marbles in the bag.
- One event strongly influences another's probability due to the non-replacement of marbles after each draw.

Bayesian network

- **Conditional probability**
  - **Conditional Probability of Specific Event Given a Multiple Events**

- The conditional probability of specific event given multiple events can be expressed using the concept of conditional probability.

- Consider events $A_1, A_2, \ldots, A_n$ where n is the number of events. The conditional probability of event A_k given that all the previous events $A_1, A_2, \ldots, A_{k-1}$ have occurred is given by:

$$P(A_k \mid A_1 \cap A_2 \cap \cdots \cap A_{k-1}) = \frac{P(A_1 \cap A_2 \cap \cdots \cap A_k)}{P(A_1 \cap A_2 \cap \cdots \cap A_{k-1})}$$

Bayesian network

Conditional probability

Conditional Probability of Multiple Events Given a Specific Event

- The conditional probability of many events given an event can be expressed using the concept of conditional probability.

- If you have events A_1, A_2, …, A_n and you want to find the conditional probability of these events given another event B, you can use the formula:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n \mid B) = \frac{P(A_1 \cap A_2 \cap \cdots \cap A_n \cap B)}{P(B)}$$

Bayesian network

- **Conditional probability**

- **Manipulating Joint Probabilities with Dependent Events**
- The joint probability of two events, can be expressed as:

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

- For more than two dependent events (A, B, and C):

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

- In a general form for n events:

$$P(A1 \cap A2 \cap \ldots \cap An) = P(A1|A2 \cap A3 \cap \ldots \cap An) \times P(A2|A3 \cap A4 \cap \ldots \cap An) \times$$

$$\ldots$$

$$\ldots \times P(An-1|An) \times P(An)$$

Bayesian network

Bayes' Theorem

- Bayes' theorem is also known as *Bayes' rule*, *Bayes' law*, or *Bayesian reasoning*, which determines the probability of an event based on prior knowledge of the conditions that might be relevant to the event.

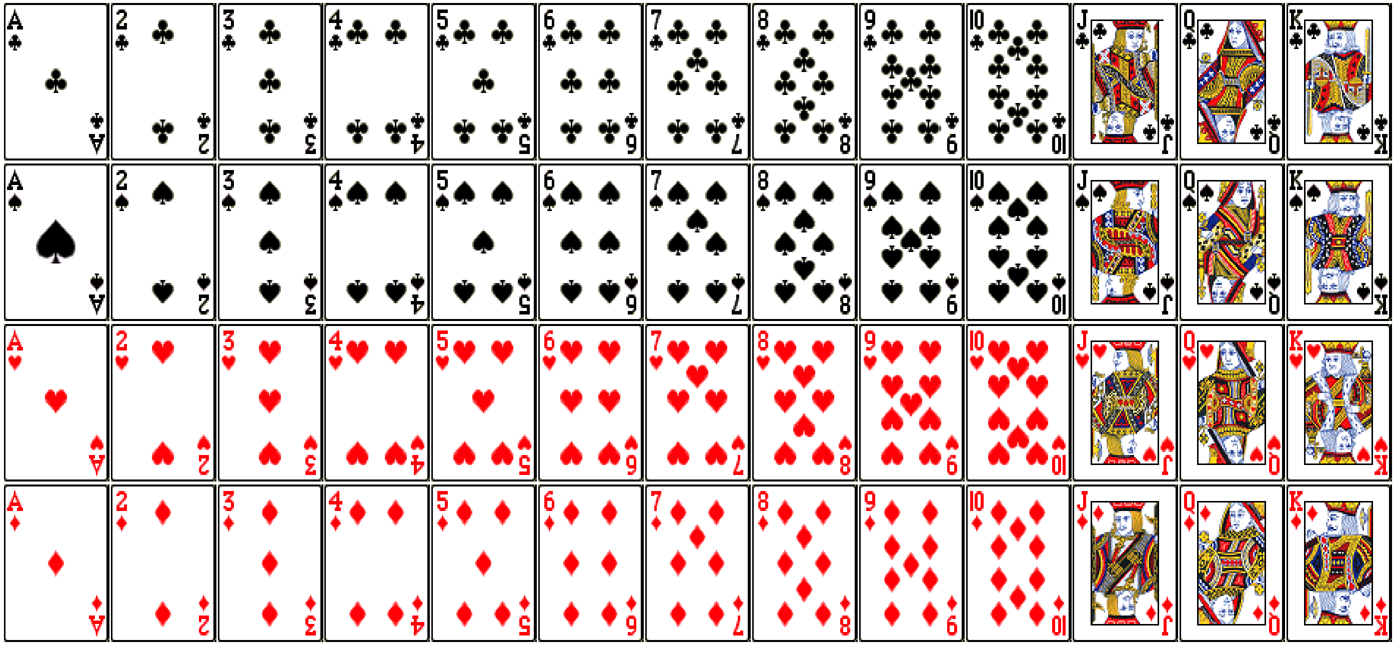
- **Formula for Bayes' Theorem:**

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$$

- P(A|B) is known as **posterior**, which we need to calculate, and it will be read as Probability of hypothesis (event) *A* when we have occurred an evidence *B*.
- P(B|A) is called the **likelihood**, in which we consider that hypothesis is true, then we calculate the probability of evidence.
- P(A) is called the **prior probability**, probability of hypothesis before considering the evidence
- P(B) is called **marginal probability**, pure probability of an evidence

BAYESIAN NETWORK

- **Bayes' Theorem**
  - **Illustrative Example: Standard 52-card deck**



Bayesian network

- **Bayes' Theorem**
  - **Illustrative Example: Standard 52-card deck**
- **Questions:**

1. Suppose you're holding a standard deck of cards and you're interested in finding the probability of drawing an Ace (event A) when you know that the card drawn is a red card (event B).

2. You draw a single card from a standard deck of playing cards. Now, calculate the posterior probability of P(King|Face), which represents the probability that the drawn face card is specifically a king.

Bayesian network

❖ Bayes' Theorem

Illustrative Example: Standard 52-card deck

- **Solution 1:** Here's how you can apply Bayes' theorem:

1. Define the events:

- Event A: Drawing an Ace.
- Event B: Drawing a red card (hearts or diamonds).

2. Find the individual probabilities:

- P(A): The probability of drawing an Ace. There are four Aces in a deck, so P(A) = 4/52 or 1/13.
- P(B): The probability of drawing a red card. There are 26 red cards in a deck (13 hearts and 13 diamonds), so P(B) = 26/52 or 1/2.

3. Determine the conditional probabilities:

- $P(A|B)$: The probability of drawing an Ace given that it's a red card.
- $P(B|A)$: The probability of drawing a red card given that it's an Ace.

Bayesian network

- **Bayes' Theorem**
  - **Illustrative Example: Standard 52-card deck**
- **Solution 1:**
- Apply Bayes' theorem to find $P(A|B)$
- $P(A|B)$ = the probability of drawing an Ace given that it's a red card.
- Bayes' theorem states:

$$P(A|B) = [P(B|A) * P(A)] / P(B)$$

- $P(B|A)$: The probability of drawing a red card given that it's an Ace is 2/4 .

$$P(A|B) = [(2/4) * (1/13)] / (1/2) = 1/13$$

- The probability of drawing an Ace given that it's a red card is **1/13**.

Bayesian network

- **Bayes' Theorem**
  - **Illustrative Example: Standard 52-card deck**
- **Solution 2:**

$$P(king|face) = \frac{P(face|king) * P(king)}{P(face)}$$

- P(king): probability that the card is King= 4/52= 1/13
- P(face): probability that a card is a face card= 3/13
- P(Face | King): probability of face card when we assume it is a king = 1

$$P(king|face) = \frac{1*\left(\frac{1}{13}\right)}{\left(\frac{3}{13}\right)} = 1/3,$$

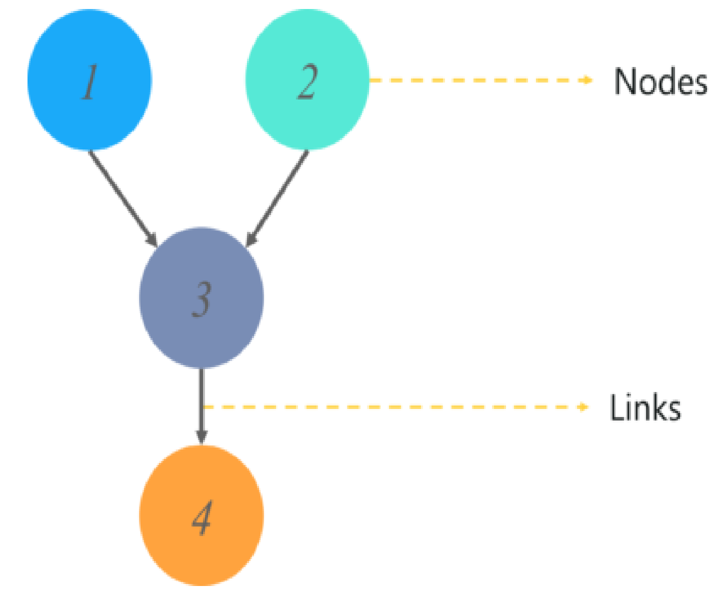
- probability that a face card is a king card = 1/3

Bayesian network

❖ Bayesian Network

- A Bayesian network (BN) or belief networks, is a graphical model that represents probabilistic relationships among a set of variables.

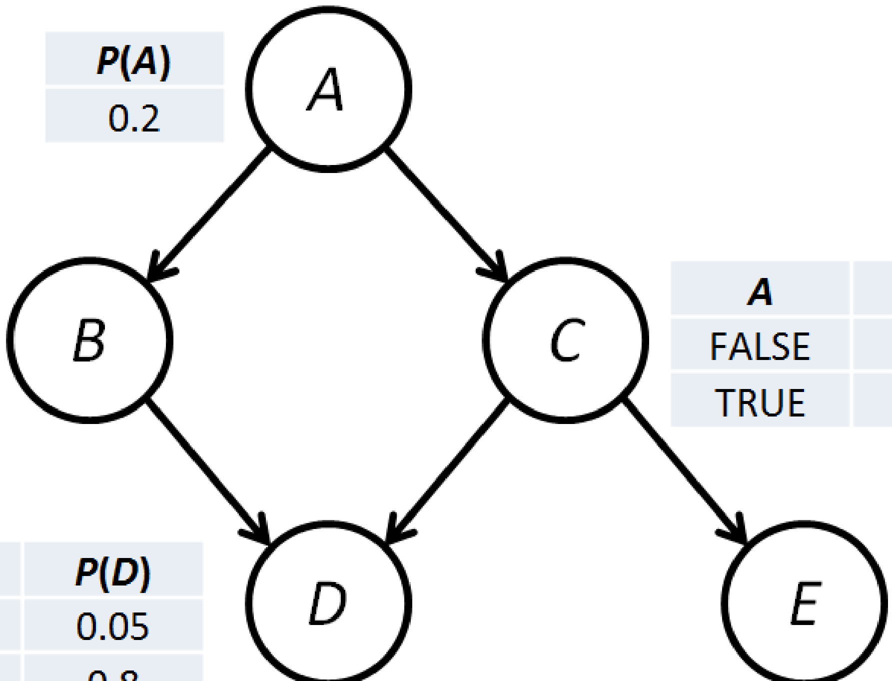
.



Bayesian network

❖ **Bayesian Network**

○



P(A)
0.2

A	P(B)
FALSE	0.2
TRUE	0.8

A	P(C)
FALSE	0.05
TRUE	0.2

B	C	P(D)
FALSE	FALSE	0.05
FALSE	TRUE	0.8
TRUE	FALSE	0.8
TRUE	TRUE	0.8

C	P(E)
FALSE	0.6
TRUE	0.8

Bayesian network

Bayesian Network

- Bayesian networks are widely used in various fields including
  - artificial intelligence,
  - machine learning,
  - medical diagnosis,
  - expert systems.
- They provide a flexible framework for modeling uncertainty and reasoning under uncertainty.

Bayesian network

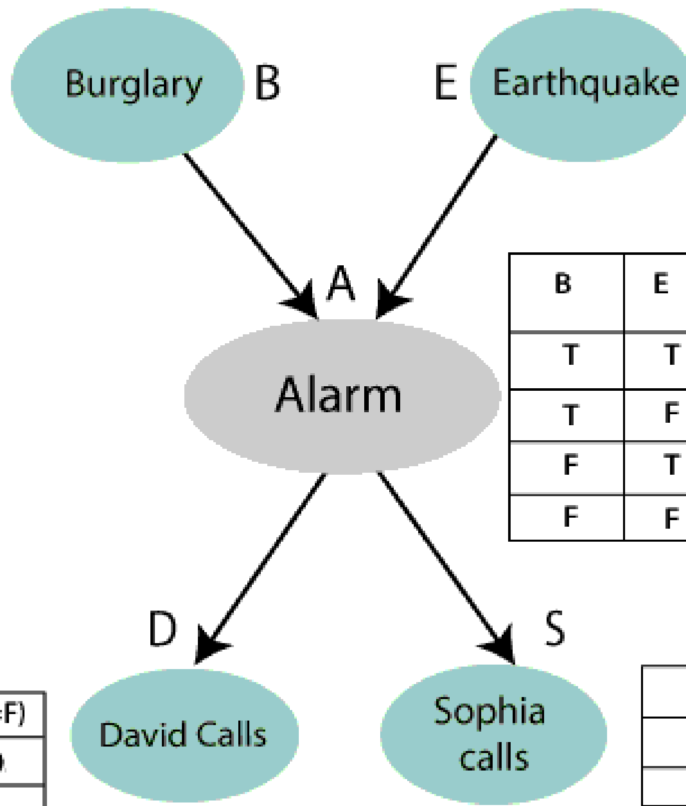
- **Bayesian Network**
- **Inference in Bayesian Networks**
- Is the process of making predictions about the values of certain variables in the network given observed evidence.
- There are two main types of inference in Bayesian networks:
  - Marginal Inference;
  - Conditional Inference.

Bayesian network

- **Bayesian Network**
  - **Inference in Bayesian Networks**
- Bayesian Network Inference involves the following steps:
  - **Observations or Evidence:**
  - **Propagation:**
  - **Inference:**

BAYESIAN NETWORK

- **Illustrative example: burglary alarm**



T	0.001
F	0.999

T	0.002
F	0.998

B	E	P(A=T)	P(A=F)
T	T	0.95	0.05
T	F	0.94	0.06
F	T	0.29	0.71
F	F	0.001	0.999

A	P (D=T)	P (D=F)
T	0.90	0.10
F	0.05	0.95

A	P (S=T)	P (S=F)
T	0.70	0.30
F	0.01	0.99

❖ Illustrative example: burglary alarm

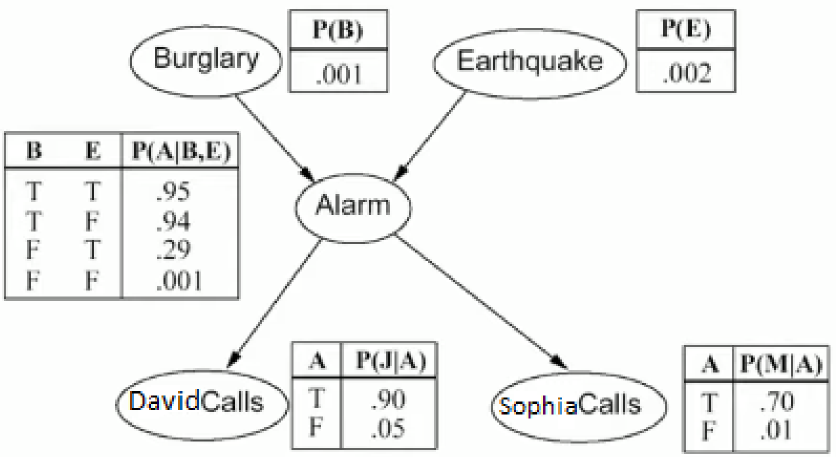
- **Example Query 1**: *What is the probability that the alarm has sounded but neither a burglary nor earthquake has occurred, and both David and Sophia call?*

Bayesian network

❖ Illustrative example: burglary alarm

- ***Example* Query 1:**

1. What is the probability that the alarm has sounded but neither a burglary nor an earthquake has occurred, and both David and Sophia call?



Solution:

$$P(D \wedge S \wedge a \wedge \neg b \wedge \neg e) = P(D \mid a)\, P(S \mid a)\, P(a \mid \neg b, \neg e)\, P(\neg b)\, P(\neg e)$$

$$= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998$$

$$= 0.00062$$

Bayesian network

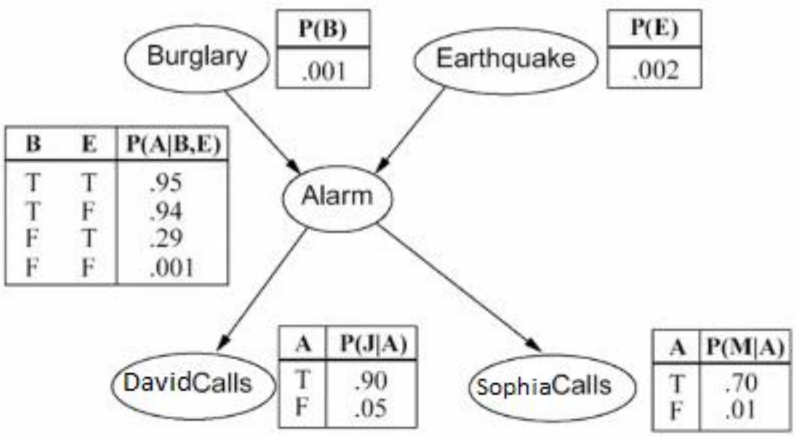
❖ Illustrative example: burglary alarm

- **Example Query 2**: *what is the probability that David call?*

❖ Illustrative example: burglary alarm

- **Example Query 2:**

2. What is the probability that David call?



Solution:

$P(j) = P(j \mid a)\,P(a) + P(j \mid \neg a)\,P(\neg a)$

$= P(j|a)\{P(a|b,e)*P(b,e)+P(a|\neg b,e)*P(\neg b,e)+P(a|b,\neg e)*P(b,\neg e)+P(a|\neg b,\neg e)*P(\neg b,\neg e)\}$

$+ P(j|\neg a)\{P(\neg a|b,e)*P(b,e)+P(\neg a|\neg b,e)*P(\neg b,e)+P(\neg a|b,\neg e)*P(b,\neg e)+P(\neg a|\neg b,\neg e)*$

$P(\neg b,\neg e)\}$

$= 0.90 * 0.00252 + 0.05 * 0.9974 = 0.0521$

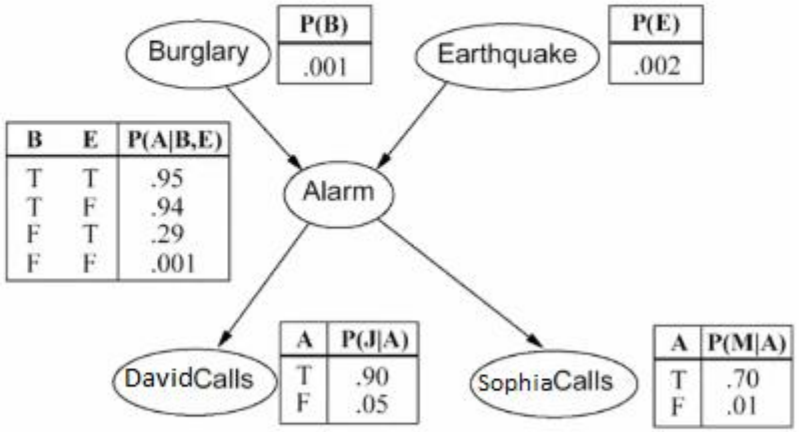
Illustrative example: burglary alarm

- **Example Query 3**: *what is the probability David call Sophia don’t call , the alarm is going off but there is no body in the house and there is no earthquake*

Bayesian network: burglary alarm

- **Example Query 3:**

3. what is the probability David call Sophia don't call , the alarm is going off but there is nobody in the house and there is no earthquake



Solution:

$$Pr(J, \neg M, A, \neg B, \neg E) =$$

$$Pr(J \mid A) Pr(\neg M \mid A) Pr(A \mid \neg B \wedge \neg E) Pr(\neg B) Pr(\neg E) =$$

$$0.9 \times 0.3 \times 0.001 \times 0.999 \times 0.998 \approx 0.0002$$