Lesson 6

Probabilistic Graphical Network Organizations— Bayesian and Markov Networks

Probability

- Probability means the chance of observing a dependent variable value with respect to some independent variable. .

"Big Data Analytics ", Ch.08 L06: Graph Analytics...SparkGraphX Platform,
Raj Kamal and Preeti Saxena, © McGraw-Hill Higher Edu. India

Probability Distribution Function (PDF)

- Means that distribution of P values as a function of all possible independent values, variables, situations, distances or variables .
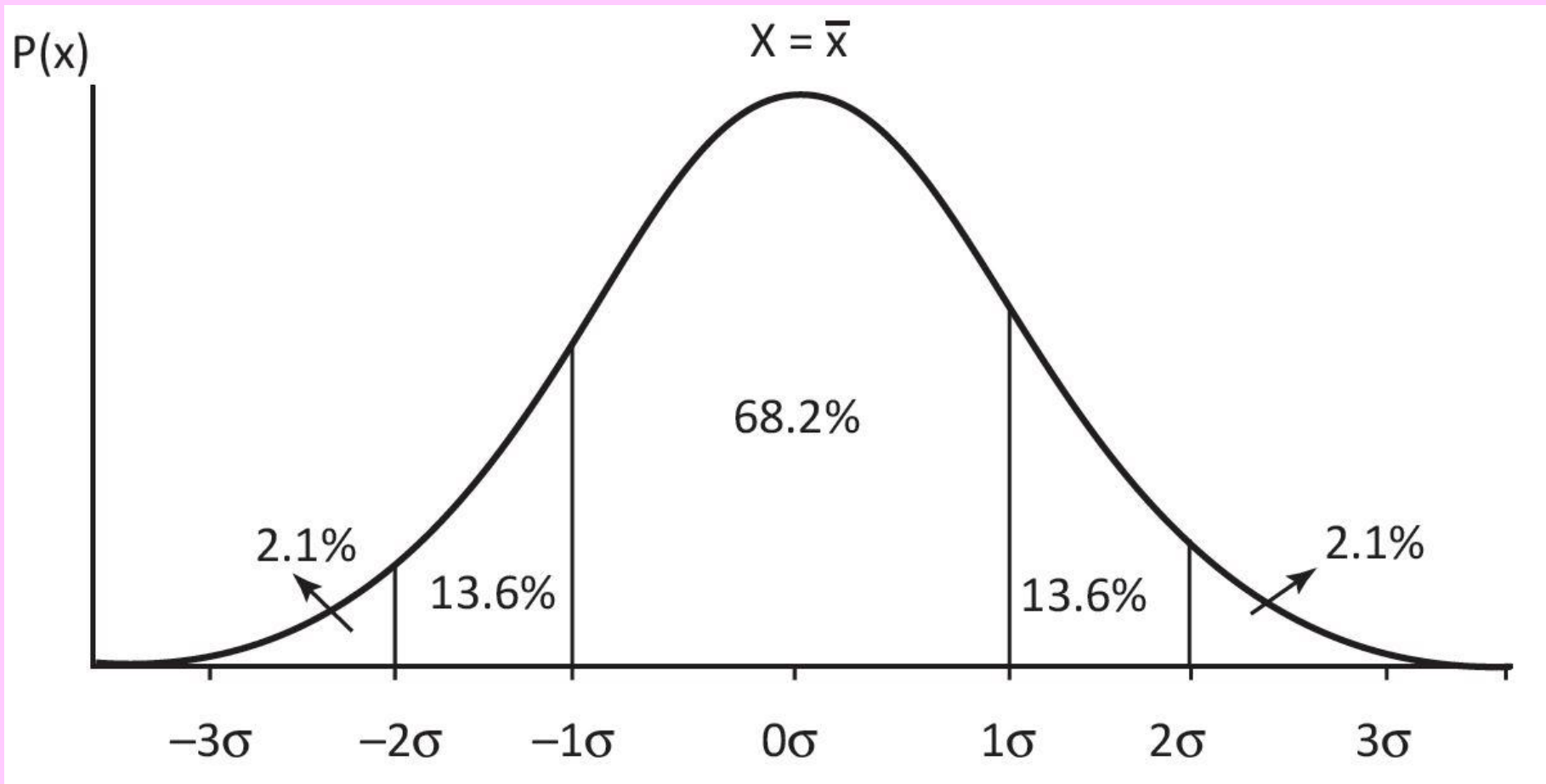
Discrete Probability Distribution Function (PDF)

Discrete Values of P: Winning, Drawn, Loosing P = (0.22, 0.78, 0)

Probability States

- Probability states mean that states of P values as a function of all possible independent values, situations or variables

- Three Probability states of Chess Grandmaster: Winning, Drawn, Loosing P = (0.22, 0.78, 0) then the probability function has three states.

Figure 6.3 Continuous Probability distribution function as a function of x assuming normal distribution around x− and standard deviation = σ

"Big Data Analytics ", Ch.08 L06: Graph Analytics...SparkGraphX Platform, Raj Kamal and Preeti Saxena, © McGraw-Hill Higher Edu. India

Probabilities Distribution

- The probabilities distribute in the entities (vertices)

- PDF P(x) at V_x distributes at the neighbouring vertices of a parent V

- Neighbouring means associated, influenced or effected vertices of the parent

Bayesian Network Graph (BNG)

- Refers to a graph where each node represents a random variable in a DAG

- The variable has s probabilistic distribution over the connected nodes

- No cyclic path traversals occur in a BNG during querying or computations

BNG Features

1. DAG with no cyclic path traversals

2. Enables a compact representation which gives probabilistic relationships among a set of variables

3. Enables the computations of joint probability distributions over the probability state variables

BNG Features

4. Each node has a set of conditional probabilities which specifies quantitatively the influences (effects) of the parent

5. Conditional Probability-Distribution (CPD) of the property values at vertices from parent to neighbours

BNG Conditional Probability Distribution

6. Enables the computations of joint probability distributions over the probability state variables

7. Table of graph nodes, node properties and probabilities, called Condition Probabilities Table (CPT)
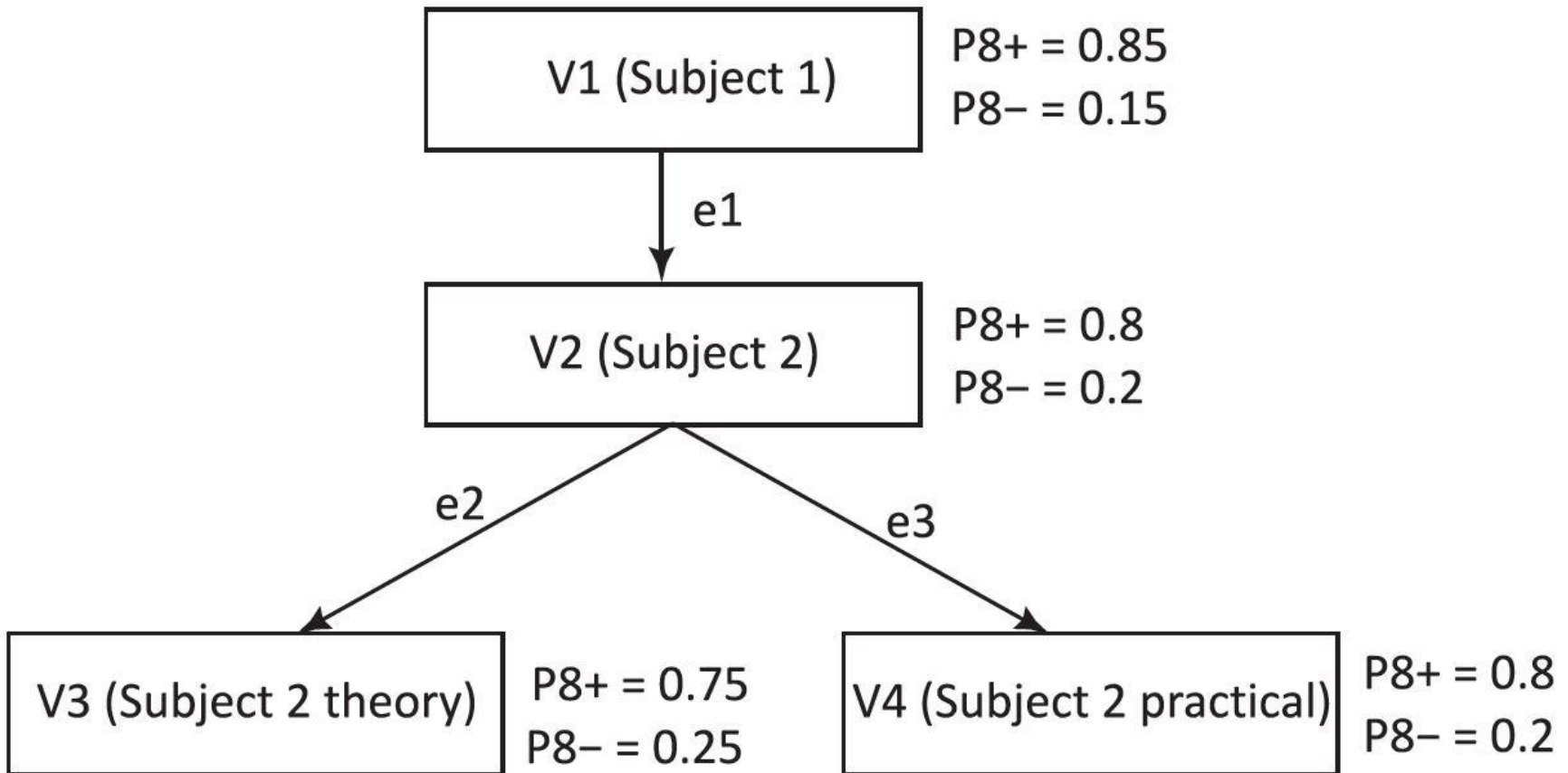
BNG Features

8. An edge between two nodes means that these two nodes have conditional probabilistic dependency

9. A missing edge between two nodes means conditional independence of the node from the parent node

"Big Data Analytics ", Ch.08 L06: Graph Analytics...SparkGraphX Platform, Raj Kamal and Preeti Saxena, © McGraw-Hill Higher Edu. India

Example 8.6

- Assume Probability P8+ for GPA 8.0 or above at V1 (Parent)

- P8− for GPA below 8.0 at V1

- BNG distributes probabilities from Parent V to neighbouring Vs

-

Figure 8.6 Bayesian network graph for a student obtaining GPAs with probabilistic distributions in the GPAs

"Big Data Analytics ", Ch.08 L06: Graph Analytics...SparkGraphX Platform, Raj Kamal and Preeti Saxena, © McGraw-Hill Higher Edu. India

Example 8.6

- Explains CPT and CPD

Condition Probability Distribution to V2 From V1

- Example 8.6
- $P(V2|1) = cp2100$ [V2 8− and V1 8−]
- $P(V2|1) = cp2101$ [V2 8− and V1 8+]
- $P(V2|1) = cp2110$ [V2 8+ and V1 8−]
- $P(V2|1) = cp2111$ [V2 8+ and V1 8+]

Condition Probability Distribution to V3 From V2

- Example 8.6
- P(V3|2) = cp3200 [V3 8− and V2 8−]
- P(V3|2) = cp3201 [V3 8− and V2 8+]
- P(V3|2) = cp3210 [V3 8+ and V2 8−]
- P(V3|2) = cp3211 [V3 8+ and V2 8+]

Markov Network Graph (MNG)

- A network organization which is undirected and can have cycles in path traversals

- Assumes that all vertices are reachable from a starting vertex

Breadth First Traversal (search) [BFS]

- Used when the graph has cycles

- The visited vertices marked

- The marks at each visited vertex stored in an array of bits (Booleans)

- First traversal along the breadth used for search

"Big Data Analytics ", Ch.08 L06: Graph Analytics...SparkGraphX Platform,
Raj Kamal and Preeti Saxena, © McGraw-Hill Higher Edu. India

Summary

We learnt:

- Probability, P states, Discrete and Continuous PDF

- Bayesian Network Features

- DAG with no cycles in path traversals

- Conditional Probability Distribution and Table

Summary

We learnt:

- Markov Network

- Cyclic Paths

- Breadth First Traversal

End of Lesson 6 on Probabilistic Graphical Network Organizations— Bayesian and Markov Networks