# Bayesian Network Chain Rule

In a Bayesian Network, the chain rule of probability is applied based on the conditional dependencies defined by the directed acyclic graph (DAG). The general form of the chain rule states:

P(X₁, X₂, ..., Xₙ) = ∏ P(Xᵢ | X₁, X₂, ..., Xᵢ₋₁) for i=1 to n

However, in a Bayesian Network, we can simplify this using the Markov property, which states that each variable Xᵢ only depends on its parents in the network, denoted as Pa(Xᵢ). This gives the Bayesian Network Chain Rule:

P(X₁, X₂, ..., Xₙ) = ∏ P(Xᵢ | Pa(Xᵢ)) for i=1 to n

Example:  
Suppose we have a Bayesian network with the following structure:

- A → B  
- A → C  
- B → D  
- C → D

The joint probability distribution is:

P(A, B, C, D) = P(A) P(B | A) P(C | A) P(D | B, C)

This is an application of the chain rule, where each variable is conditioned only on its parents.

Conclusion:  
The Bayesian Network Chain Rule helps simplify complex joint probability distributions by breaking them down into smaller, more manageable conditional probabilities, making probabilistic inference more computationally efficient.