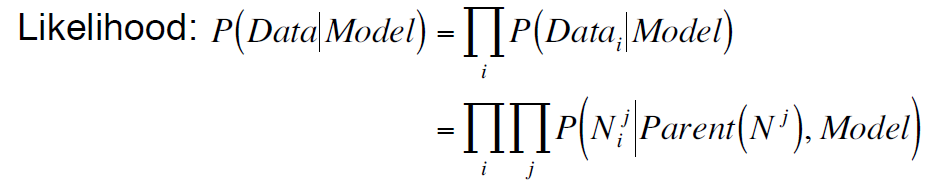
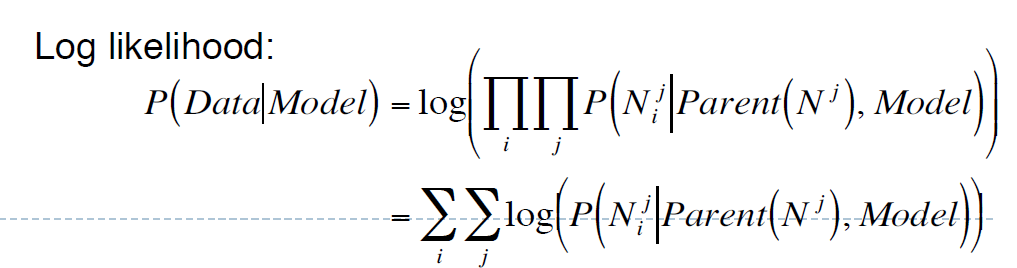
Task3

The Bayesian network is a compact representation of conditional dependency relation between random variables, consisting of a graph and conditional probability tables (CPT). And then the result of likelihood and log-likelihood are used to represent Bayesian network, and are easily to know which Bayesian network is better.

Likelihood formula:



Log-likelihood formula:



**Please explain how the likelihood and log-likelihood measure of the Bayesian Network differs as the number of training data set increases.**

If using more data from training data set to compute likelihood and log-likelihood, the result of likelihood and log-likelihood will be more accurate.

For example, if the number of training data set increase 1, so for the equation of likelihood of data set is product of probability of each row: P(Data-1)\*P(Data-2)\*…\*P(Data-n)\*P(Data-n+1), the result of the likelihood will decreases. For the log-likelihood, the equation is the summation of the log of probability of each row: log(P(Data-1))+ log(P(Data-2))+… log(P(Data-n))+ log(P(Data-n+1)), the result of log-likelihood of will increases.

**Please explain how the likelihood and log-likelihood measure of the Bayesian Network differs as the number of variables (nodes) increases.**

If the number of nodes increases, it means the structure of Bayesian Network is changed, so the result of likelihood and log-likelihood will be completely different, because the result of likelihood and log-likelihood are depended on the structure of Bayesian Network, so if the structure changes, the result of likelihood and log-likelihood also change.

For example, if there are 2 nodes A and B in Bayesian Network (A -> B, data: 0 0), and then add a new node C (A -> B -> C, data: 0 0 1). So for the equation of likelihood will change from P(∼A)∗P(∼B|∼A) to P(∼A)∗P(∼B|∼A)∗P(C|∼B), and result will decreases (CPT of node is from 0 to 1). For the log-likelihood, the equation will change from log(P(∼A))+log(P(∼B|∼A)) to log(P(∼A))+log(P(∼B|∼A))+log(P(C|∼B)), and result will increases.

**Please write a short discussion on how the likelihood and log-likelihood measure will differ when the possible values of each variable increases.**

If the possible values of each variable (node) increases, the structure of Bayesian network will not changes, but the CPT of each node will changes. As the results of likelihood and log-likelihood are depended on CPT, so if CPT of node is changed, the results of likelihood and log-likelihood will also change.

In this case, the possible values of each node will increases. For example, the data set of A is 0 1 1 0, so CPT of node A is P(A) = 2/4 = 0.5, and then the value of node A will changes from 0 to 2 (2 1 1 0), thus, the CPT of node A now is P(A) = 1+1+2/2+1+1+1 = 4/5 = 0.8. In addition, as to compute likelihood and log-likelihood need to using P(A), so results of likelihood and log-likelihood will change (the equation of likelihood and log-likelihood will be same).

In conclusion, the result of likelihood usually is a decimal number (such as 0.1 and 0.2) while the result of log-likelihood is an integer (such as -1 and -2), so the result of log-likelihood is easily to read and understand for user.