Multiple Testing Adjustment

Question

Is multiple testing adjustment (MTA) necessary for establishing significance of dependencies or associations in learned-BN?

I recently generated a learned BN (single equivalence class using SopLEQ method) from small dataset (47 instances) to associate viral DNA sequence to therapy outcomes. BN was then validated with 2 independent datasets. However, a reviewer pointed out that MTA needs to be performed.

Answer

The BayesiaLab's structural learning algorithms are based on the Minimum Description Length (MDL) score:

$$MDL = \alpha DL(B) + DL(D|B)$$

where:

- DL(B) is the Description Length of the Bayesian network B, i.e. the number of bits to represent the graph and the associated conditional probability distributions
- DL(D|B) is the Description Length of the Data given the Bayesian network B, i.e. the log-likelihood of the data given the network
- α is the Structural Coefficient that allows adjusting the relative importance of the structural complexity vs the data likelihood. This is equivalent to changing the size of the dataset N by $\frac{N}{\alpha}$

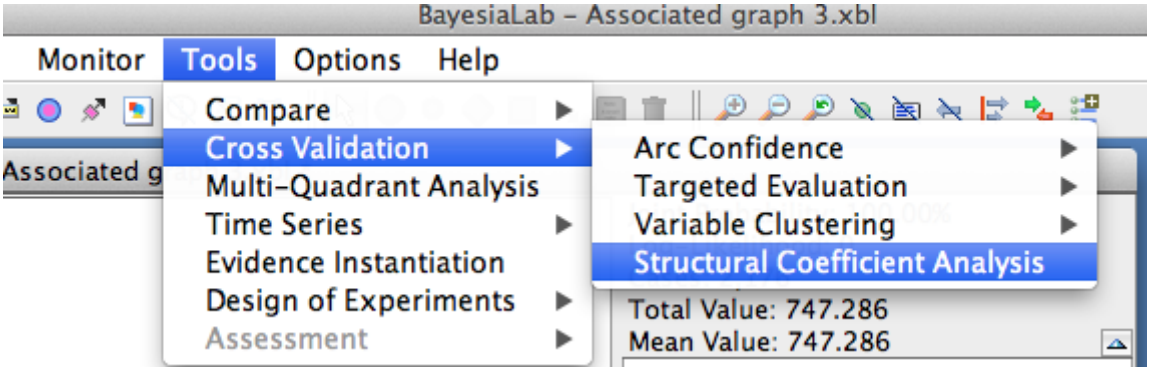
> ⚠ **Structural Coefficient**
>
> Be careful not choosing too low!
>
> Setting this coefficient to 0 leads to fully connected networks that are quickly unmanageable when the number of variable is higher than 10.

This score with $\alpha = 1$ is conservative and returns by default highly significant relations (classical statistical tests will return p-values = 0).

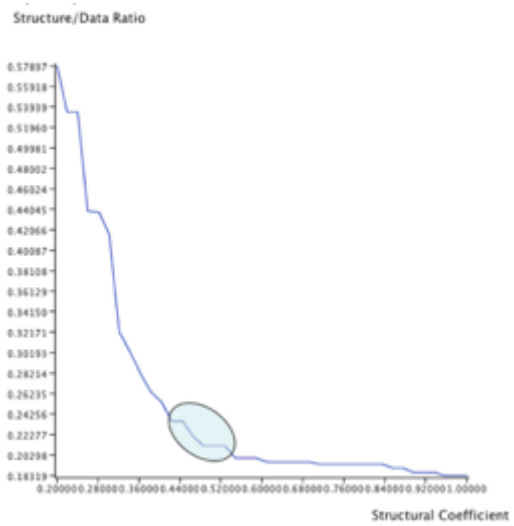
However, when data is scarce, we usually need to lower α. Choosing a value that is too low can lead to learning models with relationships that are not significant anymore. This value should then be chosen carefully.

Structural Coefficient Analysis

BayesiaLab comes with a tool that allows evaluating the Structure/Data ratio for a broad set of α values (**DL(B) / DL(D|B),** the two parts of the MDL score).

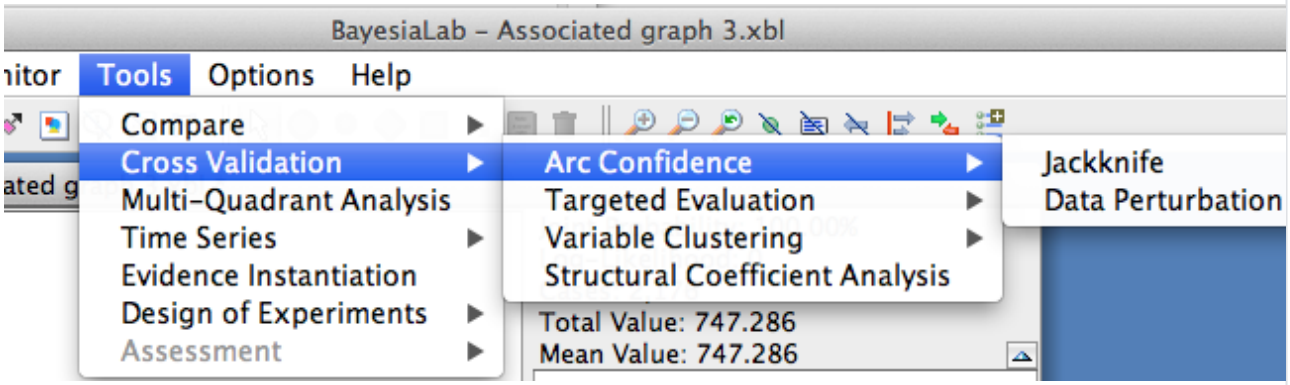


This tool generates a graph where the ratio is plotted for each α. Using the "elbow" method usually helps choosing the right value.

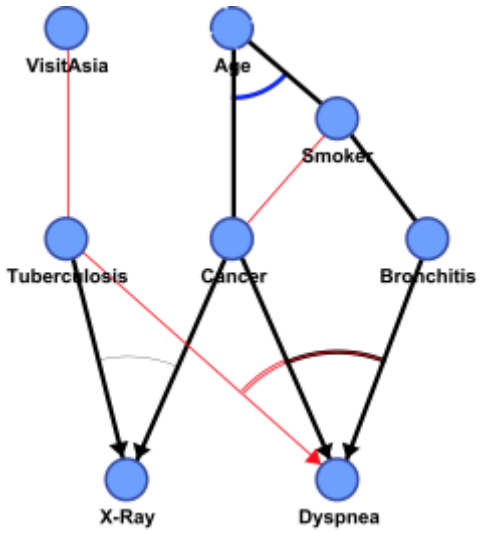


Arc Confidence

You can also use the cross validation tools for measuring the confidence of the arcs obtained on different subsets of data, or on perturbed data, with a given α



A synthetic graph is returned with black, blue and red links. The thickness of a link is directly proportional to the number of times it has been generated during the cross-validation.



- Black links are the links that were present in the original network
- Blue links are the links that have been generated during the cross-validation process but were not in the original network
- Red links are the links that were never generated during the cross-validation process but were present in the original network