Math 561 Assignment 5[*]

Due date: 14 Apr 2023.

Instructions Solve problems 1-4. Problem 5 is optional and you may earn up to 10 points extra credit on the HW if you solve it.

1. Exercise 13.1 from the book

List all the pairwise, local, and global conditional independence statements associated to the undirected graph G with V ={1,2,3,4,5,6} and undirected edge set E = {12, 13, 15, 23, 26, 34, 46, 56}.

2. Exercise 13.2.

Let G be the undirected path graph on five vertices. Give an example of distribution P that satisfies all the local Markov statements associated to G but not all the global Markov statements associated to G.

3. Exercise 13.3.

(1) Let G be a chordal graph.[1] Show that there is a directed graph H which realizes exactly the same global Markov statements as G.
(2) Conversely, if G is an undirected four-cycle, show that there is no directed graph which yields exactly the same global Markov statements as G.

4. An explicit computation for a maximum likelihood estimator from data

Consider the example of coronary heart disease from Lecture 20.
In class, we have covered example and set up the likelihood function. For this problem, compute the actual MLE using whatever method or software you choose - to maximize the log-likelihood function for the given observed data set. What might you conclude? Does the model fit the data?

As a guide, here are suggested steps to take:

- Write down explicitly the log-likelihood function (from lecture).
- Write down the score equations (all the partial derivatives).
- Find a software you like to sovle those equations to find (all local and the) global maximum of the log-likelihood function, given the data.
- Look up and study Proposition 9.3.10. Write down the result in your homework. *This is the dimension of the model.*
- Compute the model *codimension*, which is number of indeterminates minus the dimension. *This is the degrees of freedom for the model.*
- Perform the asymptotic tests of goodness of fit of this model. (If you are feeling adventurous, feel free to try an exact test!)
    - Perhaps you want to use the `loglin` functino in `R`; you can also check the degrees of freedom there. Here is an example from the R documentation:

```
# fit the model to data set called "HairEyeColor":
fm <- loglin(HairEyeColor, list(c(1, 2), c(1, 3), c(2, 3)))
```

---

[1]A graph is *chordal* if all cycles have a chord. Intuitive example: a chordal graph cannot see a 4-cycle that doesn't have a diagonal edge cutting through it.

```
## 5 iterations: deviation 0.04093795
```

```
# see the results:
fm
```

```
## $lrt
## [1] 6.761258
##
## $pearson
## [1] 6.868292
##
## $df
## [1] 9
##
## $margin
## $margin[[1]]
## [1] "Hair" "Eye"
##
## $margin[[2]]
## [1] "Hair" "Sex"
##
## $margin[[3]]
## [1] "Eye" "Sex"
```
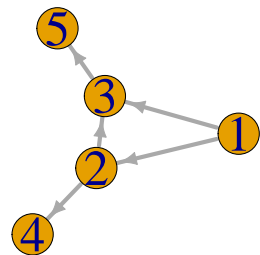
```
# comptue p-value:
1 - pchisq(fm$lrt, fm$df)
```

```
## [1] 0.66196
```

- Make sure you interpret your results.

5. Maximum likelihood for DAGs

Let G be the following DAG:

Assume that the 5 random variables are binary. Determine a formula for the maximum likelihood estimates of parameters of this model.

- Idea: maximize each local likelihood function independently of the rest of the network, and then combine the solutions to get an MLE solution.
- *(Note: The idea for this problem comes from Exercise 13.9, which asks to show that the maximum likelihood degree of any Bayesian network is 1. In particular, explain how to determine rational formulas for the maximum likelihood estimates of parameters in graphical models with directed acyclic graphs.)*