VE414 Lecture 25

Jing Liu

UM-SJTU Joint Institute

December 2, 2019

- Suppose we have no historical data on my opponent, i.e. the only data is on

$$F = \{0, 1, 2, 3\} \qquad \text{and} \qquad S = \{0, 1, 2\}$$

which are the No. of cards discarded and replaced in the 1st and 2nd round.

- Since the information is very limited, we should group hands into categories:
  1. High card (nothing special)
  2. One ace
  3. Two of consecutive value (ace is assumed to loop back to 2)
  4. Two of a suit (there are 4 suits ♦ ♣ ♥ ♠)
  5. One pair
  6. Flush (3 cards of the same suit)
  7. Straight (3 consecutive value)
  8. Three of a kind (3 of cards of the same value)
  9. Straight flush

  in increasing strength instead of trying to consider all the individual hands.

- The first thing to have in mind when constructing a Bayesian network model is that its purpose is to give estimates of probabilities for events that are not directly observable, in this case, the opponent's hand.

$$H = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

- The second thing to have in mind is that in order to obtain the probabilities, we should construct channels such that the observable variables may reveal something about the unobservable ones, in this case,

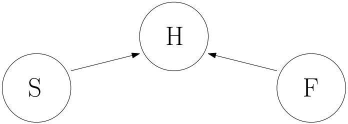
$$F \qquad \text{and} \qquad S$$

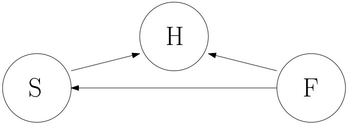
- Of course, we don't know the true causal relationships in practice, all models are simplifications of reality, thus all of them are wrong, but some are useful!

$$\mathrm{H} \longrightarrow \mathrm{F} \longrightarrow \mathrm{S} \qquad\qquad \mathrm{F} \longrightarrow \mathrm{H} \longleftarrow \mathrm{S}$$

Q: Both models are equally (in)valid without further information, however, we are making very different assumptions by using them, can you identify them?

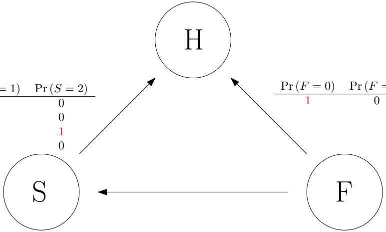
Q: Again, the followings are equally (in)valid, but is there any difference now?





Q: Given we have observed $F = 0$ and $S = 0$, what is the probability of $H = 1$?

F	S	$\Pr(H=1)$	$\Pr(H=2)$	$\Pr(H=3)$	$\Pr(H=4)$	$\Pr(H=5)$	$\Pr(H=6)$	$\Pr(H=7)$	$\Pr(H=8)$	$\Pr(H=9)$
0	0	1	0	0	0	0	0	0	0	0
0	1	0	1	0	0	0	0	0	0	0
0	2	0	0	1	0	0	0	0	0	0
1	0	0	0	0	1	0	0	0	0	0
1	1	0	0	0	0	1	0	0	0	0
1	2	0	0	0	0	0	1	0	0	0
2	0	0	0	0	0	0	0	1	0	0
2	1	0	0	0	0	0	0	0	1	0
2	2	1	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0
3	1	0	0	1	0	0	0	0	0	0
3	2	0	0	0	1	0	0	0	0	0

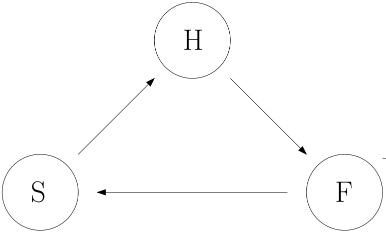


F	$\Pr(S=0)$	$\Pr(S=1)$	$\Pr(S=2)$
0	1	0	0
1	0	1	0
2	0	0	1
3	1	0	0

$\Pr(F=0)$	$\Pr(F=1)$	$\Pr(F=2)$	$\Pr(F=3)$
1	0	0	0

- Although all models are wrong, some are wrong in a different sense!

S	$\Pr(H=1)$	$\Pr(H=2)$	$\Pr(H=3)$	$\Pr(H=4)$	$\Pr(H=5)$	$\Pr(H=6)$	$\Pr(H=7)$	$\Pr(H=8)$	$\Pr(H=9)$
0	1	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0



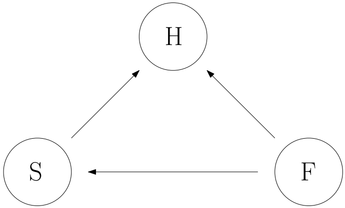
H	$\Pr(F=0)$	$\Pr(F=1)$	$\Pr(F=2)$	$\Pr(F=3)$
1	1	0	0	0
2	0	1	0	0
3	0	0	1	0
4	0	0	0	1
5	1	0	0	0
6	0	1	0	0
7	0	0	1	0
8	0	0	0	1

F	$\Pr(S=0)$	$\Pr(S=1)$	$\Pr(S=2)$
0	1	0	0
1	0	1	0
2	0	0	1
3	1	0	0

Q: Given we have observed $F=0$ and $S=0$, what is the probability of $H=1$?

- This is not a Bayesian network since it is a directed cyclic graph, the type of reasoning using probability in BNs cannot be used in CNs in general.

- Since all models are wrong, the quality of a model should be judged by data rather than intuition, but intuition is a key step in building a model.



- The above model seems to be very natural, but it is not clear at all how to properly specify the marginal and conditional probability for this network.
- For example, it is unclear what we should use for f_F, and it is just unclear if not less clear what we should use for $f_{H|F}$.
- So when specifying the structure, we shall also keep in mind whether there is a way to specify those probabilities for the given direction of implications.

Q: Is there any probability that we can work out for this poker game example?

- Note we are only interested in the final hand of my opponent

$$H = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

it is the hand after both rounds of discarding/replacing of cards, where

1. High card
2. One Ace
3. Two of consecutive value
4. Two of a suit
5. One pair
6. Flush
7. Straight
8. Three of a kind
9. Straight Flush

- There are two other unobservable hands of my opponent in a game,

$$FH = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$
$$SH = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

i.e. the hands before the 1st/2nd round of discarding/replacing, respectively.

- The most readily available probabilities are those for the first hand (FH).

- Through various combinatorial calculations, we can obtain

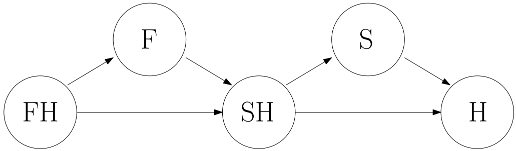
$$f_{FH}$$

- For example,

$$\Pr(\text{Straight}) = f_{FH}(6) = \frac{12 \cdot 4^3 - 4 \cdot 12}{22100}$$

Q: Without more assumptions, is there any other probabilities we can work out?

- If we use FH as the only node without a parent, and also include SH, then



seems to be a very nature Bayesian network for the poker game example.

- If we make some assumptions regarding how my opponent discards his cards according to common poker strategies, for example, in the first round
  - If High Card, then discard all three cards.
  - If One Ace, then keep the ace and discard the other two cards.
  - If 2 of consecutive value, 2 of a suit, or 2 of the same value, then discard the third card.
  - If 2 of a suit and 2 of consecutive value, then keep 2 of a suit.
  - If 2 of a suit and 2 of the same value or 2 of consecutive value and 2 of the same value, then keep the 2 of the same value.
  - If flush, straight, 3 of the same value, or straight flush, then keep it.

  then we can work out

$$f_{F|FH}$$

  making similar assumptions about the 2nd round, and obtain

$$f_{S|SH}$$

Q: How can we specify the remaining two $f_{SH|F,FH}$ and $f_{H|S,SH}$?

- When working with systems that evolve over time, we can introduce a

discrete time stamp

and have a model for each unit of time. Such a sub-model is known as a

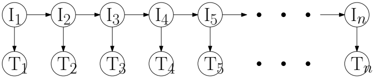
time slice.

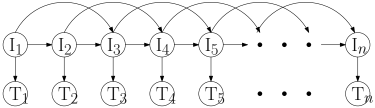
Infected Milk

Milk from a cow may be infected. To detect whether the milk is infected, you have a test, which may give either a positive or a negative test result. The test is not perfect. It may give a positive result on clean milk as well as a negative result on infected milk. From one day to another, the state of the milk can change. Cows with infected milk will heal over time, and a clean cow has a risk of having infected milk the next day. Suppose that the farmer performs the test each day. After a week, he has not only the current test result but also the six previous test results.

Q: What is a sensible model to use?

- The time slices are connected through temporal links to form a full model.





- If the structures of the time slices are identical, and if the temporal links are the same, we say that the model is a

  repetitive temporal model.

- If the conditional probabilities are also identical, we call the model a

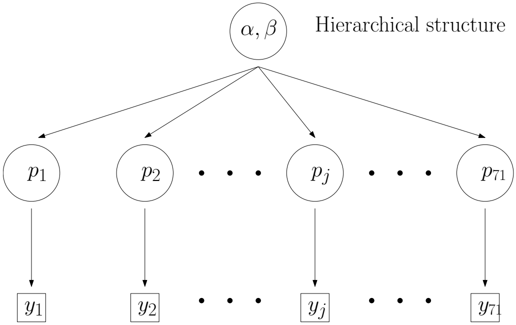
  dynamic Bayesian network model.

Q: Have you seen a dynamic Bayesian network before?

- HMM is a Bayesian network, specifically, it is a dynamic Bayesian network model, where the memoryless property is assumed.

$$\Pr\left(I_{t+1} = j \mid I_t, I_{t-1}, \ldots, I_1\right) = \Pr\left(I_{t+1} = j \mid I_t\right)$$

Q: How to deal with a continuous random variable?



- BNs with both discrete and continuous random variables are known as

  hybrid Bayesian networks

  where squares are usually used for discrete random variables by convention.

Q: How can we estimate the PMFs or parameters of PDF of a node given data?

Bayesian Analysis and Bayesian Network

- There are not the same!!!
- The defining feature of Bayesian analysis is the estimation principle

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

- On the contrary, Bayesian Networks are often used by frequentists.
- Originally a Bayesian network is just a complex, principled way of proposing a probabilistic model on a set of variables, without necessarily involving data.
- In terms of Bayesian analysis given some data, BN leads us to

  the posterior joint distribution

  by putting one component or one layer at time, once there then everything in early part of this course is applicable as well as frequentist's approaches.