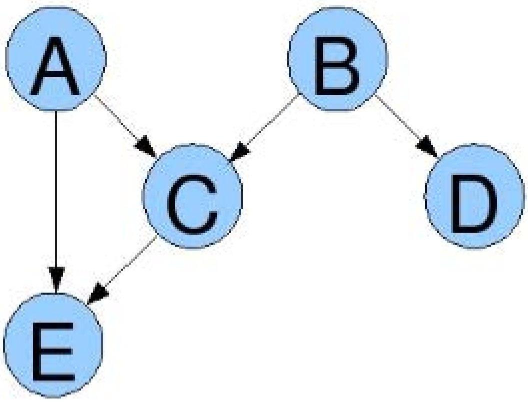
Bayesian Network Independencies

Latest Submission Grade 100%

1. Independencies in a graph.

1 / 1 point

Which pairs of variables are independent in the graphical model below, given that none of them have been observed? You may select 1 or more options.



- [x] A, B

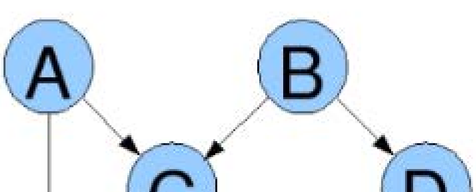
Correct
There are no active trails between A and B, so they are independent.

- [ ] A, E
- [ ] D, E
- [ ] A, C
- [ ] None - there are no pairs of independent variables.

2. *Independencies in a graph. (An asterisk marks a question that is more challenging. Congratulations if you get it right!)

1 / 1 point

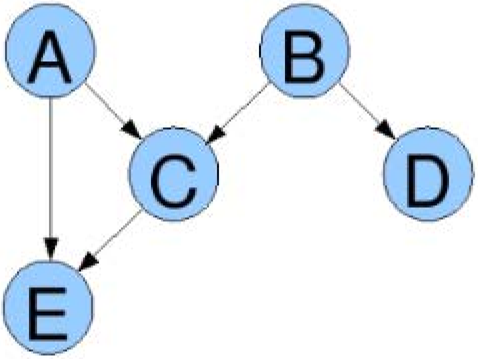
Now assume that the value of E is known. (E is observed. A, B, C, and D are not observed.) Which pairs of variables (not including E) are independent in the same graphical model, given E? You may select 1 or more options.



2. ***Independencies in a graph.** (An asterisk marks a question that is more challenging. Congratulations if you get it right!)

1 / 1 point

Now assume that the value of E is known. (E is observed. A, B, C, and D are not observed.) Which pairs of variables (not including E) are independent in the same graphical model, given E? You may select 1 or more options.



- [x] None - given E, there are no pairs of variables that are independent.

> Correct
> Observing E activates the V-structures around C and E, giving rise to active trails between every pair of variables in the network.

- [ ] A, B
- [ ] A, C
- [ ] A, D
- [ ] B, D
- [ ] D, C
- [ ] B, C

3. **I-maps.** I-maps can also be defined directly on graphs as follows. Let $I(G)$ be the set of independencies encoded by a graph G. Then G_1 is an I-map for G_2 if $I(G_1) \subseteq I(G_2)$.

1 / 1 point

Which of the following statements about I-maps are true? You may select 1 or more options.

- [x] A graph K is an I-map for a graph G if and only if all of the independencies encoded by K are also encoded by G.

> Correct
> K is an I-map for G if K does not make independence assumptions that are not true in G. An easy way to remember this is that the complete graph, which has no independencies, is an I-map of all distributions.

- [ ] A graph K is an I-map for a graph G if and only if K and G are identical, i.e., they have exactly the same nodes and edges.

3. **I-maps.** I-maps can also be defined directly on graphs as follows. Let $I(G)$ be the set of independencies encoded by a graph G. Then G_1 is an I-map for G_2 if $I(G_1) \subseteq I(G_2)$.

1 / 1 point

Which of the following statements about I-maps are true? You may select 1 or more options.

- [x] A graph K is an I-map for a graph G if and only if all of the independencies encoded by K are also encoded by G.

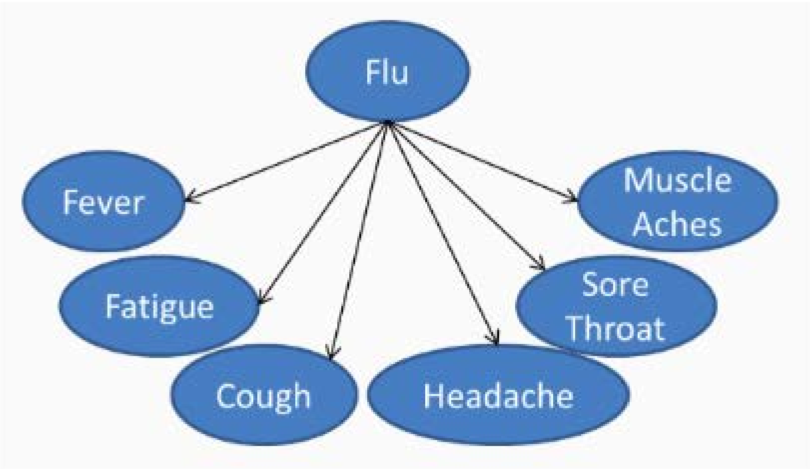
Correct
K is an I-map for G if K does not make independence assumptions that are not true in G. An easy way to remember this is that the complete graph, which has no independencies, is an I-map of all distributions.

- [ ] A graph K is an I-map for a graph G if and only if K and G are identical, i.e., they have exactly the same nodes and edges.
- [ ] An I-map is a function f that maps a graph G to itself, i.e., $f(G) = G$.
- [ ] The graph K that is the same as the graph G, except that all of the edges are oriented in the opposite direction as the corresponding edges in G, is always an I-map for G, regardless of the structure of G.
- [ ] I-maps are Apple's answer to Google Maps.

4. ***Naive Bayes.**

1 / 1 point

Consider the following Naive Bayes model for flu diagnosis:



Assume a population size of 10,000. Which of the following statements are true in this model? You may select 1 or more options.

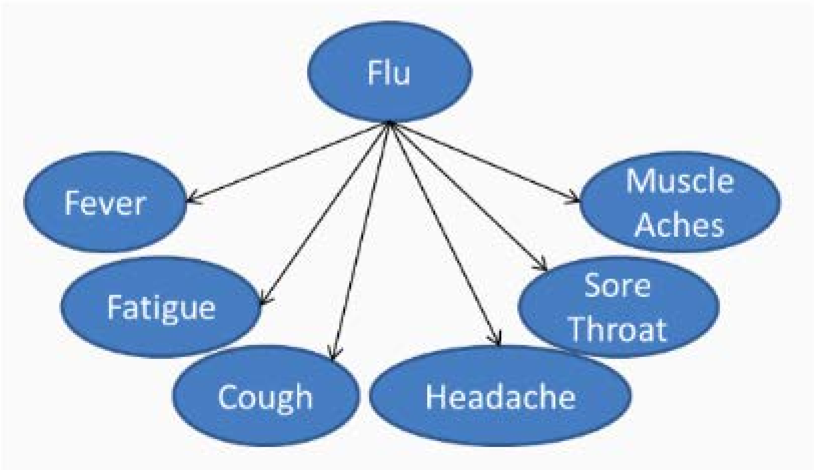
- [x] Say we observe that 1000 people have the flu, out of which 500 people have a headache (and possibly other symptoms) and 500 have a fever (and possibly other symptoms).

  We would expect that approximately 250 people with the flu also have both a headache and fever.

4. ***Naive Bayes.**

1 / 1 point

Consider the following Naive Bayes model for flu diagnosis:



Assume a population size of 10,000. Which of the following statements are true in this model? You may select 1 or more options.

- [x] Say we observe that 1000 people have the flu, out of which 500 people have a headache (and possibly other symptoms) and 500 have a fever (and possibly other symptoms).

  We would expect that approximately 250 people with the flu also have both a headache and fever.

Correct

Given that someone has the flu, whether he has a headache is independent of whether he has a fever. We can thus calculate:

$$\begin{aligned} P(Headache = 1, Fever = 1 | Flu = 1) &= P(Headache = 1 | Flu = 1) \times P(Fever = 1 | Flu = 1) \\ &\approx 0.5 * 0.5 \\ &= 0.25. \end{aligned}$$

Since 1000 people have the flu, we can estimate that 250 of these people will have both a headache and fever.

Note that this is only an estimate: we can assert with high confidence that $P(Headache = 1, Fever = 1 | Flu = 1)$ is near to 0.25, but in general it will not be exactly 0.25. Moreover, even if it is exactly 0.25, the number of people with the flu, a headache and a fever need not be exactly 250 all the time. Think of this as analogous to flipping a fair coin: even though the probability of seeing a heads is exactly 0.5, in any given sequence of coin flips we need not see exactly half of the coins turning up heads.

- [ ] Say we observe that 1000 people have the flu, out of which 500 people have a headache (and possibly other symptoms) and 500 people have a fever (and possibly other symptoms).

  Without more information, we cannot estimate how many people with the flu also have both a headache and fever.

- [x] Say we observe that 1000 people have a headache (and possibly other symptoms), out of which 500 people have the flu (and possibly other symptoms), and 500 people have a fever (and possibly other symptoms).

  Without more information, we cannot estimate how many people with a headache also have both the flu and a fever.

analogous to flipping a fair coin: even though the probability of seeing a heads is exactly 0.5, in any given sequence of coin flips we need not see exactly half of the coins turning up heads.

☐ Say we observe that 1000 people have the flu, out of which 500 people have a headache (and possibly other symptoms) and 500 people have a fever (and possibly other symptoms).

Without more information, we cannot estimate how many people with the flu also have both a headache and fever.

☑ Say we observe that 1000 people have a headache (and possibly other symptoms), out of which 500 people have the flu (and possibly other symptoms), and 500 people have a fever (and possibly other symptoms).

Without more information, we cannot estimate how many people with a headache also have both the flu and a fever.

Correct

Even after observing the Headache variable, there is still an active trail from Flu to Fever. Thus, the probability of someone with a headache also having a flu is dependent on the probability of his having a fever as well. For example, if someone has a flu, he could be more likely to have a fever, irrespective of whether he has a headache or not.

We therefore cannot estimate $P(Flu = 1, Fever = 1 | Headache = 1)$ from the conditional marginal probabilities $P(Flu = 1 | Headache = 1)$ and $P(Fever = 1 | Headache = 1)$.

☑ Say we observe that 500 people have a headache (and possibly other symptoms) and 500 people have a fever (and possibly other symptoms).

Without more information, we cannot estimate how many people have both a headache and fever.

Correct

Without having observed the Flu variable, there is an active trail from Headache and Fever. Thus, the probability of someone having a headache (without observing flu status) is not independent of the probability of the same person having a fever. For example, if someone has a headache, he might be more likely to have the flu, which would correspondingly increase the probability that he has a fever as well.

We therefore cannot estimate $P(Headache = 1, Fever = 1)$ from the marginal probabilities $P(Headache = 1)$ and $P(Fever = 1)$.

5. I-maps. 1 / 1 point

Suppose $(A \perp B) \in \mathcal{I}(P)$, and G is an I-map of P, where G is a Bayesian network and P is a probability distribution. Is it necessarily true that $(A \perp B) \in \mathcal{I}(G)$?

○ Yes

◉ No

Correct

Since G is an I-map of P, all independencies in G are also in P. However, this doesn't mean that all independencies in P are also in G. An easy way to remember this is that the complete graph, which has no independencies, is an I-map of all distributions.