



# Probabilistic Graphical Models

Daphne Koller, Kevin Murphy  
Winter 2011-2012

[Home](#)
[Quizzes](#)
[Theory Problems](#)
[Assignments](#)
[Assignment Questions](#)
[Video Lectures](#)
[Discussion Forums](#)
[Octave Installation](#)
[Lecture Slides](#)
[Course Schedule](#)
[Course Logistics](#)
[Course Information](#)
[Course Staff](#)

## Feedback — PA2 Quiz

You achieved a score of **22.00** out of **22.00**

### Question 1

James and Rene come to a genetic counselor because they are deciding whether to have another child or adopt. They want to know the probability that their un-born child will have cystic fibrosis.

Consider the Bayesian network for cystic fibrosis. Order the probabilities of their un-born child having cystic fibrosis in the following situations from smallest to largest: (1) No phenotypes are observed, (2) Jason has cystic fibrosis, (3) Sandra has cystic fibrosis.

Your Answer	Score	Explanation
<input checked="" type="radio"/> (1), (3), (2)	2.00	Since Benjamin's phenotype and genotype are not observed in all of these situations, the probability that he will have cystic fibrosis (CF) is equivalent to the probability that James and Rene's unborn child will have CF. Observing that Benjamin's cousin has CF makes Benjamin more likely to have CF because CF is a genetic disease. Observing that Benjamin's brother has CF makes Benjamin more likely to have CF than when observing that Benjamin's cousin has CF because Benjamin's brother is a more closely-related relative than his cousin is.
Total	2.00	

### Question 2


James never knew his father Ira because Ira passed away in an accident when James was a few months old. Now James comes to the genetic counselor wanting to know if Ira had cystic fibrosis. The genetic counselor wants your help in determining the probability that Ira had cystic fibrosis. Consider the Bayesian network for cystic fibrosis. Order the probabilities of Ira having had cystic fibrosis in the following situations from smallest to largest: (1) No phenotypes are observed, (2) Benjamin has cystic fibrosis, (3) Benjamin and Robin have cystic fibrosis.

Your Answer	Score	Explanation
<input checked="" type="radio"/> (1), (3), (2)	2.00	Observing that Ira's grandson has cystic fibrosis (CF) makes Ira more likely to have CF because CF is a genetic disease. Observing that Ira's wife also has CF partially explains away why Ira has CF.
Total	2.00	

### Question 3


Recall that for a trait with 2 alleles, the CPD for genotype given parents' genotypes has 27 entries. How many entries does the CPD for phenotype given parents' genotypes have?

many entries would be in the CPD if the trait had 3 alleles instead of 2?

Your Answer	Score	Explanation
216	 2.00	There are 6 possible genotypes for each parent and for the child, so the size of the CPD is $6 \times 6 \times 6 = 216$
Total	2.00	


#### Question 4

You will now gain some intuition for why decoupling a Bayesian network can be worthwhile. Consider a **non-decoupled** Bayesian network for cystic fibrosis with **3 alleles** over the pedigree that was used in section 2.4 and 3.3. How many CPD entries are there in total, across all the CPDs in the network?

Your Answer	Score	Explanation
1212	 2.00	There are 12 entries in each phenotype given genotype factor (of which there are 9), 6 entries in each genotype given allele frequency factor (of which there are 4), and 216 entries in each child genotype given parents' genotypes factor (of which there are 5).
Total	2.00	

#### Question 5

Now consider the **decoupled** Bayesian network for cystic fibrosis with **3 alleles** over the pedigree that was used in section 2.4 and 3.3. How many CPD entries are there in total, across all the CPDs in the network?


Your Answer	Score	Explanation
456	 2.00	There are 18 entries in each phenotype given genotype factor (there are 9 such factors), 3 entries in each copy of gene given allele frequency factor (there are 8 such factors), and 27 entries in each child copy of gene given parent's copies of gene factor (there are 10 such factors).
Total	2.00	

#### Question 6

Consider the **decoupled** Bayesian network for cystic fibrosis with three alleles that you constructed in section 3.3.

James and Rene are debating whether to have another child or adopt a child. They are concerned that, if they have a child, the child will have cystic fibrosis because both of them have one F allele observed (their other gene copy is not observed), even though neither of them have cystic fibrosis. You want to give them advice, but they refuse to tell you whether anyone else in their family has cystic fibrosis. What is the probability that their unborn child will have cystic fibrosis?



**Round your answer to 2 decimal places.** If the probability is between 0 and 1, put a 0 in front of the decimal point

Your Answer	Score	Explanation
0.47	 2.00	Knowing that James and Rene each have at least one F allele makes the child more likely to have cystic fibrosis, even though neither of them have CF.
Total	2.00	

### Question 7

Consider a Bayesian network for spinal muscular atrophy (SMA), in which there are multiple genes and 2 phenotypes.



Let  $n$  be the number of genes involved in SMA and  $m$  be the maximum number of alleles per gene. How many parameters are necessary if we use a table CPD for the probabilities for phenotype given copies of the genes from 1 parent?

Your Answer	Score	Explanation
 $O(m^{2n})$	 2.00	There are two alleles per gene, so there are $O(m^2)$ allele combinations per gene. Therefore, there are $O(m^{2n})$ parameters for $n$ genes.
Total	2.00	

### Question 8








Consider the Bayesian network for spinal muscular atrophy (SMA), in which there are multiple genes and two phenotypes.

Let  $n$  be the number of genes involved in SMA and  $m$  be the maximum number of alleles per gene. How many parameters are necessary if we use a sigmoid CPD for the probabilities for phenotype given copies of the genes from both parents?

Your Answer	Score	Explanation
 $O(mn)$	 2.00	Each gene has up to $m$ alleles, and there is an indicator for each allele for each copy of the gene. Therefore, if there were one gene, there would be $O(2m) = O(m)$ parameters. Since there are $n$ genes, there are $O(mn)$ possible parameters.
Total	2.00	

**Question 9**

Consider genes A and B that might be involved in spinal muscular atrophy. Assume that A has 2 alleles  $A_1$  and  $A_2$ , and B has 2 alleles,  $B_1$  and  $B_2$ . Which of the following relationships between A and B can a sigmoid CPD capture?

Your Answer	Score	Explanation
<input checked="" type="checkbox"/> Allele $A_1$ makes a person more likely to have SMA, while allele $B_1$ independently makes a person less likely to have SMA.	 0.29	A sigmoid CPD can capture this by making the weights for the indicators for allele $A_1$ positive while making the weights for the indicators for allele $B_1$ negative.
<input checked="" type="checkbox"/> Neither gene A nor gene B contribute to SMA.	 0.29	A sigmoid CPD can capture this by giving alleles for copies of gene A as well as alleles for copies of gene B weights with value zero.
<input type="checkbox"/> Allele $A_1$ and allele $B_1$ make a person equally more likely to have SMA, but when both are present the effect on SMA is the same as when only one is present.	 0.29	This OR relationship cannot be captured by a sigmoid CPD because interaction terms between the alleles are not present.
<input checked="" type="checkbox"/> Alleles $A_1$ and $B_1$ each independently make a person likely to have SMA.	 0.29	Since their contributions are independent, a sigmoid CPD that weights the alleles each gene based on the extent of their contribution would capture this perfectly.
<input checked="" type="checkbox"/> Gene A contributes to SMA, but gene B does not contribute to SMA and thus does not affect the effects of gene A on SMA.	 0.29	A sigmoid CPD can capture this by giving the alleles for copies of gene A positive weights and the alleles for copies of gene B zero weights.
<input type="checkbox"/> When the alleles are $A_1$ and $B_2$ or $A_2$ and $B_1$ the person has SMA; otherwise the person does not have SMA.	 0.29	This XOR relationship means that the effect of the allele for gene A depends on which allele for gene B is present; since the sigmoid CPD does not have interactive terms, it will not be able to capture this.
<input type="checkbox"/> Allele $A_1$ and allele $B_1$ make a person more likely to have SMA when both of these alleles are present, but neither affect SMA otherwise.	 0.29	This AND relationship cannot be captured by a sigmoid CPD because interaction terms between the alleles are not present.
Total	2.00	

**Question 10**

Consider the Bayesian network for spinal muscular atrophy that we provided in spinalMuscularAtrophyBayesNet.net.

Now say that Ira and Robin come to the genetic counselor because they are debating whether to have a biological child or adopt and are concerned that their child might have spinal muscular atrophy. They have some genetic information because sequencing is still far too expensive to be affordable for everyone, their information is limited to only a few genes and to only 1 chromosome in each pair of chromosomes.

Order the probabilities of their un-born child having spinal muscular atrophy in the following situations from smallest to largest: (1) No genetic information or phenotypes are observed, (2) Ira and Robin each have at least 1 M allele, (3) Ira

and Robin each have at least 1 M allele and at least 1 B allele.

Your Answer	Score	Explanation
<div><div><div></div></div><div>(1), (2), (3)</div></div>	<div><div></div><div>2.00</div></div>	Since James is unobserved, the probability that he will have spinal muscular atrophy (SMA) is equivalent to the probability that Ira and Robin's unborn child will have SMA. Observing that Ira and Robin each have an allele that is involved in causing SMA makes James more likely to have SMA than if no variables were observed. Observing that Ira and Robin each have alleles for 2 genes that are involved in causing SMA makes James even more likely to have SMA than if only 1 allele for 1 gene were observed.
Total	2.00	

Question 11

Consider the Bayesian network for spinal muscular atrophy that we provided in spinalMuscularAtrophyBayesNet.net

No longer interested in finding out whether his father had cystic fibrosis, James comes to the genetic counselor with another question: Did his father have spinal muscular atrophy? The genetic counselor now wants your help in figuring this out. This time, however, James has other information for you: both he and Robin have spinal muscular atrophy.

What is the probability that Ira had spinal muscular atrophy?

Round your answer to 2 decimal places. If the probability is between 0 and 1, put a 0 in front of the decimal point

0.35

Your Answer	Score	Explanation
0.35	<div><div></div><div>2.00</div></div>	Since Ira's wife has spinal muscular atrophy (SMA), this helps explain why his child has SMA, so Ira is more likely to have SMA than he would be if no phenotypes were observed but is less likely to have SMA than he would be if only James were observed to have SMA.
Total	2.00	