Written Assignment 4

Deadline: November 27th, 2019

Instruction: You may discuss these problems with classmates, but please complete the write-ups individually. (This applies to BOTH undergraduates and graduate students.) Remember the collaboration guidelines set forth in class: you may meet to discuss problems with classmates, but you may not take any written notes (or electronic notes, or photos, etc.) away from the meeting. Your answers must be **typewritten**, except for figures or diagrams, which may be hand-drawn. Please submit your answers (pdf format only) on **Canvas**.

Q1. Probability (15 points)

Below is a table listing the probabilities of three binary random variables. Compute each marginal or conditional probability below.

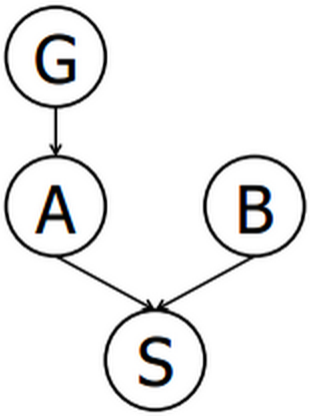
- $P(X_0 = 1, X_1 = 0, X_2 = 1)$
- $P(X_0 = 0, X_1 = 1)$
- $P(X_2 = 0)$
- $P(X_1 = 0 \mid X_0 = 1)$
- $P(X_0 = 1, X_1 = 0 \mid X_2 = 1)$
- $P(X_0 = 1 \mid X_1 = 0, X_2 = 1)$

X_0	X_1	X_2	$P(X_0, X_1, X_2)$
0	0	0	0.160
1	0	0	0.100
0	1	0	0.120
1	1	0	0.040
0	0	1	0.180
1	0	1	0.200
0	1	1	0.120
1	1	1	0.080

Q2. Bayes Nets and Probability (21 points)

Suppose that a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. The Bayes' Net and corresponding probability tables for this situation are shown below.

G	$P(G)$
g	0.10
$\neg g$	0.90



B	$P(B)$
b	0.40
$\neg b$	0.60

G	A	$P(A \mid G)$
g	a	1.00
g	$\neg a$	0.00
$\neg g$	a	0.10
$\neg g$	$\neg a$	0.90

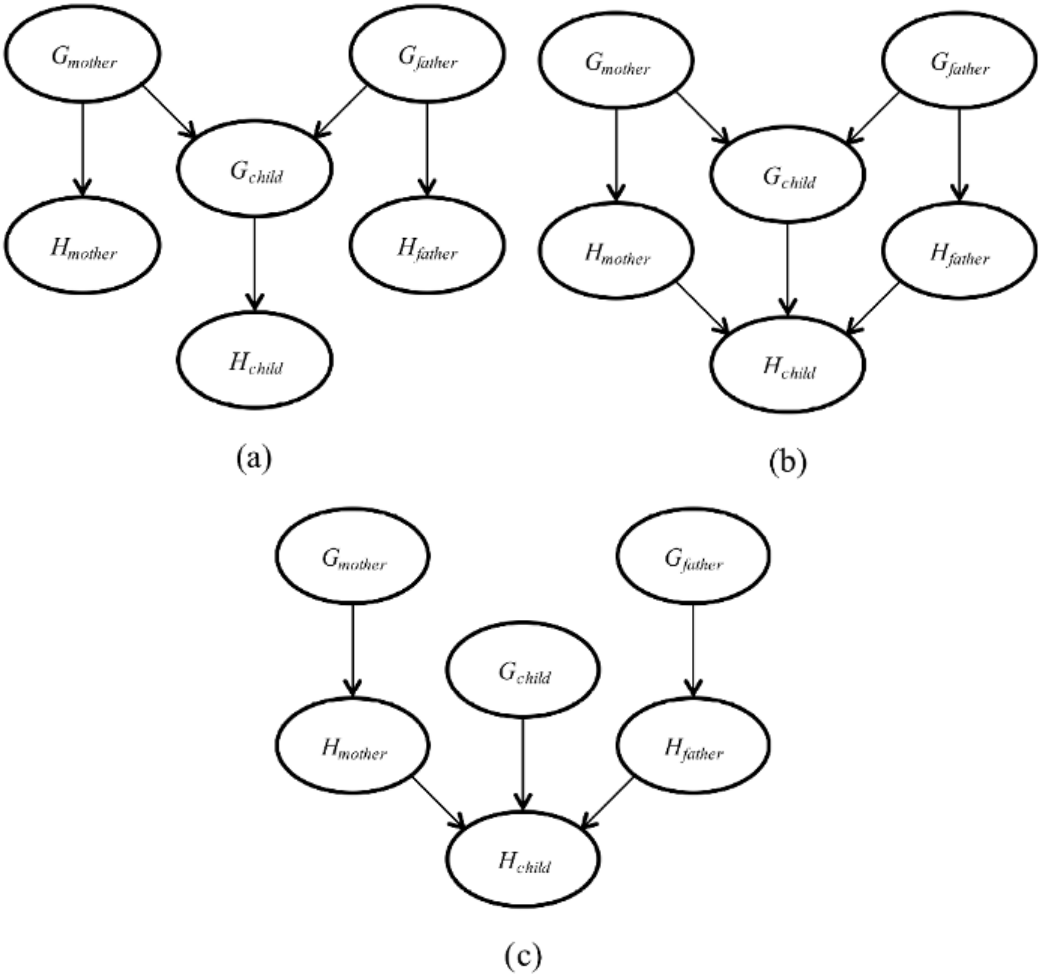
A	B	S	$P(S \mid A, B)$
a	b	s	1.00
a	b	$\neg s$	0.00
a	$\neg b$	s	0.90
a	$\neg b$	$\neg s$	0.10
$\neg a$	b	s	0.80
$\neg a$	b	$\neg s$	0.20
$\neg a$	$\neg b$	s	0.10
$\neg a$	$\neg b$	$\neg s$	0.90

- Compute $P(g, a, b, s)$
- What is the probability that a patient has disease A?
- What is the probability that a patient has disease A given that they have disease B?
- What is the probability that a patient has disease A given that they have symptom S and disease B?
- What is the probability that a patient has the disease carrying gene variation G given that they have disease A?
- What is the probability that a patient has the disease carrying gene variation G given that they have disease B?

Q3. Bayes Nets: Independence (15 points)

Let H_x be a random variable denoting the handedness of an individual x, with possible values l or r. A common hypothesis is that left- or right-handedness is inherited by a simple mechanism; that is, perhaps there is a gene G_x, also with values l or r, and perhaps actual handedness turns out mostly the same (with some probability s) as the gene an individual possesses. Furthermore, perhaps the gene itself is equally likely to be inherited from either of an individual's parents, with a small nonzero probability m of a random mutation flipping the handedness.

The following three images are possible models involving the genes G and handednesses H.



- Which of the three network above claim that:

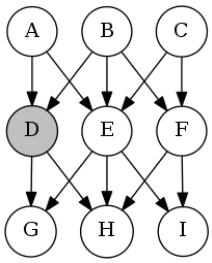
$$P(G_{father}, G_{mother}, G_{child}) = P(G_{father})P(G_{mother})P(G_{child})$$

- Which of the three networks make independence claims that are consistent with the hypothesis about the inheritance of handedness?
- Which of the three networks is the best description of the hypothesis?

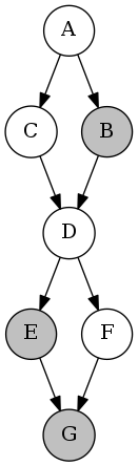
Q4. Bayes Nets: D-Separation (21 points)

You are given several graphical models below, and each model is associated with an independence (or conditional independence) assertion. Please specify if the assertion is true or false.

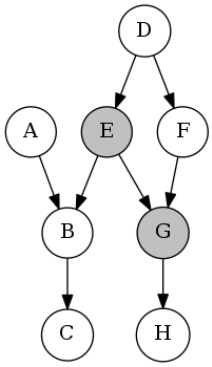
Q4.1. It is guaranteed that G is independent of H given D.



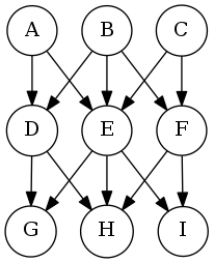
Q4.2. It is guaranteed that A is independent of D given E, B, G.



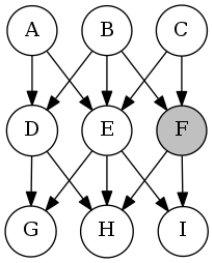
Q4.3. It is guaranteed that H is independent of B given G, E.



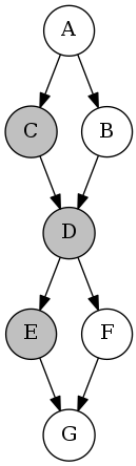
Q4.4. It is guaranteed that A is independent of C.



Q4.5. It is guaranteed that D is independent of C given F.

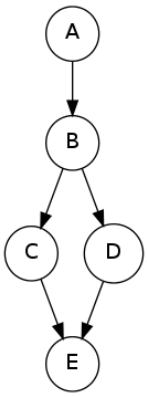


Q4.6. It is guaranteed that G is independent of B given C, E, D.



Q5. Bayes Nets: Variable Elimination (13 points)

Assume the following Bayes Net and corresponding CPTs. We are given the query $P(C \mid e = 1)$, and we will complete the tables for each factor generated during the elimination process.



After introducing evidence, we have the following probability tables.

A	$P(A)$
0	0.100
1	0.900

B	A	$P(B\|A)$
0	0	0.500
1	0	0.500
0	1	0.400
1	1	0.600

C	B	$P(C\|B)$
0	0	0.400
1	0	0.600
0	1	0.300
1	1	0.700

D	B	$P(D\|B)$
0	0	0.600
1	0	0.400
0	1	0.900
1	1	0.100

C	D	$P(e = 1\|C, D)$
0	0	0.600
1	0	0.200
0	1	0.600
1	1	0.200

Q5.1. (10 pts) Three steps are required for elimination, with the resulting factors listed below:

- Step 1: Eliminate A. We get the factor $f_1(B) = \sum_a P(a)P(B \mid a)$
- Step 2: Eliminate B. We get the factor $f_2(C, D) = \sum_b P(C \mid b)P(D \mid b)f_1(b)$
- Step 3: Eliminate D. We get the factor $f_3(C, e = 1) = \sum_d P(e = 1 \mid C, d)f_2(C, d)$

Fill in the missing quantities. (some of the quantities are computed for you)

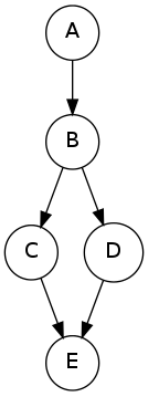
- $f_1(B = 0) =$
- $f_1(B = 1) =$
- $f_2(C = 0, D = 0) =$
- $f_2(C = 1, D = 0) =$
- $f_2(C = 0, D = 1) = 0.083$
- $f_2(C = 1, D = 1) = 0.14$
- $f_3(C = 0, e = 1) =$
- $f_3(C = 1, e = 1) = 0.132$

Q5.2. (3 pts) After getting the factor $f_3(C, e = 1)$, a final re-normalization step needs to be carried out to obtain the conditional probability $P(C \mid e = 1)$. Fill in the final conditional probabilities below.

- $P(C = 0 \mid e = 1) =$
- $P(C = 1 \mid e = 1) =$

Q6. Bayes Nets: Sampling (15 points)

Q6.1. Rejection Sampling We will work with a Bayes' net of the following structure.



In this question, we will perform rejection sampling to estimate $P(C = 1 \mid B = 1, E = 1)$. Perform one round of rejection sampling, using the random samples given in the table below. Variables are sampled in the order A, B, C, D, E.

0.320	0.037	0.303	0.318	0.032	0.969	0.018	0.058	0.908	0.249

A	$P(A)$
0	0.200
1	0.800

B	A	$P(B\|A)$
0	0	0.800
1	0	0.200
0	1	0.400
1	1	0.600

C	B	$P(C\|B)$
0	0	0.600
1	0	0.400
0	1	0.400
1	1	0.600

D	B	$P(D\|B)$
0	0	0.800
1	0	0.200
0	1	0.600
1	1	0.400

E	C	D	$P(E\|C,D)$
0	0	0	0.800
1	0	0	0.200
0	1	0	0.600
1	1	0	0.400
0	0	1	0.400
1	0	1	0.600
0	1	1	0.400
1	1	1	0.600

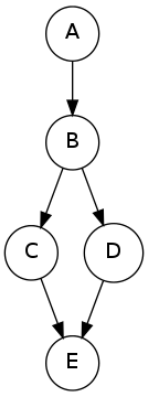
Note that the sampling attempt should stop as soon as you discover that the sample will be rejected. In that case mark the assignment of that variable and write "none" for the rest of the variables. When generating random samples, use as many values as needed from the table below, which we generated independently and uniformly at random from $[0, 1)$. Use

numbers from left to right. To sample a binary variable W with probability $P(W = 0) = p$ and $P(W = 1) = 1 - p$ using a value a from the table, choose $W = 0$ if $a < p$ and $W = 1$ if $a \geq p$.

Choose the value (0 or 1) that each variable gets assigned to:

- A: B: C: D: E:
- Which variable will get rejected?

Q6.2. Likelihood Weighting We will work with a Bayes' net of the following structure.



In this question, we will perform likelihood weighting to estimate $P(C = 1 \mid B = 1, E = 1)$. Generate a sample and its weight, using the random samples given in the table below. Variables are sampled in the order A, B, C, D, E.

When generating random samples, use as many values as needed from the table below, which we generated independently and uniformly at random from $[0, 1)$. Use numbers from left to right. To sample a binary variable W with probability $P(W = 0) = p$ and $P(W = 1) = 1 - p$ using a value a from the table, choose $W = 0$ if $a < p$ and $W = 1$ if $a \geq p$.

0.249	0.052	0.299	0.773	0.715	0.550	0.703	0.105	0.236	0.153

A	$P(A)$
0	0.200
1	0.800

B	A	$P(B\|A)$
0	0	0.400
1	0	0.600
0	1	0.200
1	1	0.800

C	B	$P(C\|B)$
0	0	0.600
1	0	0.400
0	1	0.600
1	1	0.400

D	B	$P(D\|B)$
0	0	0.800
1	0	0.200
0	1	0.600
1	1	0.400

E	C	D	$P(E\|C,D)$
0	0	0	0.200
1	0	0	0.800
0	1	0	0.600
1	1	0	0.400
0	0	1	0.800
1	0	1	0.200
0	1	1	0.800
1	1	1	0.200

Select the assignments to the variables you sampled.

- A: B: C: D: E:
- What is the weight for the sample you obtained above?