CMPSCI 240: Reasoning Under Uncertainty
Third Midterm Exam

April 15, 2015.

Name: _______________________________ ID: _______________________

Instructions:

• Answer the questions directly on the exam pages.

• Show all your work for each question. Giving more detail including comments and explanations can help with assignment of partial credit.

• If the answer to a question is a number, you may give your answer using arithmetic operations, such as addition, multiplication, and factorial (e.g., “9 \times 35! + 2” or “0.5 \times 0.3/(0.2 \times 0.5 + 0.9 \times 0.1)” is fine).

• If you need extra space, use the back of a page.

• No books, notes, calculators or other electronic devices are allowed. Any cheating will result in a grade of 0.

• If you have questions during the exam, raise your hand.

• The formulas for some standard random variables can be found on the last page.

<table>
<thead>
<tr>
<th>Question</th>
<th>Value</th>
<th>Points Earned</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>8+2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>58+2</td>
<td></td>
</tr>
</tbody>
</table>
**Question 1. (10 points)** Indicate whether each of the following five statements is TRUE or FALSE. No justification is required.

1.1 (2 points): If $X$ and $Y$ are independent random variables then $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$.

  **Answer:** True.

1.2 (2 points): If events $A$ and $B$ satisfy $P(A|B) = P(A)$ then they also satisfy $P(B|A) = P(B)$.

  **Answer:** True.

1.3 (2 points): For any two events $A$ and $B$, $P(A \cap B) \leq P(A) \times P(B)$.

  **Answer:** False.

1.4 (2 points): Expected number of throws of a 6-sided dice before we’ve seen all sides is $6 + 1/6$.

  **Answer:** False.

1.5 (2 points): For any two events $A$ and $B$, $P(A|B^c) = 1 - P(A|B)$.

  **Answer:** False.

**Question 2. (10 points)** The economic cycle can be modeled with a three-state Markov chain as shown below. The three states are shrinking (S), flat (F), and growing (G). Suppose the time steps in the Markov chain correspond to years. The transition diagram is as follows.

2.1 (2 points): Suppose the economy is currently shrinking. What is the probability the economy will be shrinking next year and will become flat the year after that?

  **Answer:** $P(S \to S \to F) = 1/2 \times 1/2 = 1/4$.

2.2 (2 points): Suppose the economy is currently shrinking. What is the probability the economy will be growing in three years time?

  **Answer:** $P(S \to S \to F \to G) + P(S \to F \to G \to G) = 1/2 \times 1/2 \times 2/3 + 1/2 \times 2/3 \times 1/2 = 1/3$.

2.3 (2 points): Circle the two words that apply to the above Markov chain:

  aperiodic periodic reducible irreducible

2.4 (2 points): What is the matrix of transition probabilities corresponding to this Markov chain?
Answer:
\[
\begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 \\
\frac{1}{3} & 0 & \frac{2}{3} \\
0 & \frac{1}{2} & \frac{1}{2}
\end{pmatrix}
\]

2.5 (2 points): Extra Credit. What are the probabilities in the steady state of the Markov chain?

Answer: Let \( v = (a, b, c) \) be the steady-state distribution. Then
\[
(a, b, c) = (a, b, c) \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 \\
\frac{1}{3} & 0 & \frac{2}{3} \\
0 & \frac{1}{2} & \frac{1}{2}
\end{pmatrix} = (a/2 + b/3, a/2 + c/2, 2b/3 + c/2)
\]
By comparing the first coordinate we learn \( b/3 = a/2 \) and hence
\[
(a, 3a/2, c) = (a, a/2 + c/2, a + c/2)
\]
By comparing the last coordinate we learn \( a = c/2 \). Hence the steady state distribution is
\[
(a, 3a/2, 2a)
\]
and since we need \( a + 3a/2 + 2a = 9a/2 = 1 \) we get
\[
(2/9, 1/3, 4/9)
\]

Question 3. (10 points) You’re about to try a new type of ice-cream and have two hypotheses:

\( H_1 = \) you will like the ice-cream and \( H_2 = \) you won’t like the ice-cream

Suppose that you’ve tried 40 different types of ice-cream in your life and so far you’ve found 20 types that you liked and 20 that you didn’t like. You also remember that:

1. of the types you liked included asparagus
2. of the types you liked included banana
0. of the types you didn’t like included asparagus
20. of the types you didn’t like included banana

Let \( A \) and \( B \) be the events the new type of ice-cream includes asparagus and banana respectively. You may assume that \( P(A \cap B | H_1) = P(A | H_1)P(B | H_1) \) and \( P(A \cap B | H_2) = P(A | H_2)P(B | H_2) \).

3.1 (1 point): What values should you use for the priors:
\[
P(H_1) = 1/2 \quad P(H_2) = 1/2
\]
3.2 (2 points): What values should you use for the likelihoods:

\[ P(A|H_1) = \frac{1}{20} \quad P(B|H_1) = \frac{1}{10} \quad P(A|H_2) = 0 \quad P(B|H_2) = 1 \]

3.3 (2 points): Which is the maximum a posteriori (MAP) hypothesis if the new type of ice-cream includes both asparagus and bananas. Show your working.

**Answer:** We choose \( H_1 \) as the MAP hypothesis since

\[ P(A|H_1)P(B|H_1)P(H_1) = \frac{1}{20} \times \frac{1}{10} \times \frac{1}{2} = \frac{1}{400} \]

is bigger than \( P(A|H_2)P(B|H_2)P(H_2) = 0 \times 1 \times \frac{1}{2} = 0 \).

3.4 (2 points): Given that \( A \) and \( B \) conditionally independent on \( H_1 \) is it necessarily true that \( A^c \) and \( B \) are conditionally independent on \( H_1 \)? Justify your answer.

**Answer:** Yes. Since \( A \) and \( B \) are independent conditioned on \( H_1 \) we have \( P(A|B,H_1) = P(A|H_1) \). Therefore \( P(A^c|B,H_1) = 1 - P(A|B,H_1) = 1 - P(A|H_1) = P(A^c|H_1) \) i.e., \( A^c \) and \( B \) are independent conditioned on \( H_1 \).

3.5 (2 points): Which is the maximum a posteriori (MAP) hypothesis if the new type of ice-cream includes bananas but doesn’t include asparagus. Show your working.

**Answer:** We choose \( H_2 \) as the MAP hypothesis since

\[ P(A^c|H_1)P(B|H_1)P(H_1) = \frac{19}{20} \times \frac{1}{10} \times \frac{1}{2} = \frac{19}{400} \]

is less than \( P(A^c|H_2)P(B|H_2)P(H_2) = 1 \times 1 \times \frac{1}{2} = \frac{1}{2} \).

3.6 (1 points): If you wanted to improve your ability to predict the types of ice-cream you like, what additional information would be useful. (A variety of possible suggestions will be accepted.)

**Answer:** We could use more “features” (i.e., consider the presence or absence of additional ingredients including combinations of ingredients), and/or collect more training data (i.e., try more varieties of ice cream to get better estimates of the likelihoods).

**Question 4. (10 points)** Suppose that five friends, Albert, Bernard, Cheryl, Duane, and Erik, are discussing their birthdays. For this question you should assume that each of their birthdays are equally likely to be any of the 365 days in the year. You should ignore leap years.

4.1 (2 points): What’s the probability that the five friends have five different birthdays?

**Answer:** \( \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \frac{361}{365} \).

4.2 (2 points): What’s the probability all five friends have the same birthday?

**Answer:** \( \left( \frac{1}{365} \right)^4 \).
4.3 (2 points): What’s the probability Cheryl’s birthday is different from Albert’s birthday and Bernard’s birthday? Hint: Albert and Bernard may or may not share a birthday with each other.

**Answer:** Let $D_1$ be the event that Cheryl’s birthday is different from Albert’s birthday and Bernard’s birthday. Let $D_2$ be the event that Albert’s birthday and Bernard’s birthday are different. Then

$$P(D_1) = P(D_1|D_2)P(D_2) + P(D_1|D_2^c)P(D_2^c) = \frac{363}{365} \times \frac{364}{365} + \frac{364}{365} \times \frac{1}{365} = \left(\frac{364}{365}\right)^2$$

In the rest of the question, you should assume that the five friends have five different birthdays.

4.4 (2 points): What are the values for the following probabilities. Hint: Consider the possible orderings of Albert’s, Bernard’s and Cheryl’s birthdays.

**Answer:**

$$P(H_1) = \frac{1}{2} \quad P(H_2) = \frac{1}{2} \quad P(D|H_1) = \frac{2}{3} \quad P(D|H_2) = \frac{1}{3}$$

To determine the values for $P(D|H_1)$ and $P(D|H_2)$ the easiest approach is to consider the possible orderings of the birthdays, e.g., $ABC, ACB, \ldots$. Then $H_1$ corresponds to $ABC, ACB, CAB$ and since $A$ is before $C$ is exactly 2 of these orderings $P(D|H_1) = \frac{2}{3}$. Similarly, $H_2$ corresponds to $BAC, BCA, CBA$ and since $A$ is before $C$ is exactly 1 of these orderings $P(D|H_2) = \frac{1}{3}$.

4.5 (2 points): Conditioned on event $D$, which of $H_1$ and $H_2$ is more likely? Show your working.

**Answer:** $H_1$ is more likely because $P(D|H_1)P(H_1) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$ is greater than $P(D|H_2)P(H_2) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$.

**Question 5.** (10 points) This is based on a true story from World War II. Suppose that Germany has numbered their tanks and has printed these numbers on the sides of the tanks. The numbers are $1, 2, 3, \ldots, t$ where $t$ is unknown. The Allied forces want to learn the value of $t$, i.e., the number of tanks that Germany has built. They have narrowed it down to two hypotheses:

- $H_1$ = “there are $t = 100$ tanks”
- $H_2$ = “there are $t = 200$ tanks”

and the priors for these hypotheses are $P(H_1) = \frac{1}{10}$ and $P(H_2) = \frac{9}{10}$. Each day an Allied spy reports the number of a tank that they have observed. You should assume that the tank is equally likely to be any of the tanks that have been built. Today the spy reports that a tank was sighted with the number 50. Call this event $D_1$.

5.1 (2 points): What values would you pick for the following likelihoods.

$$P(D_1|H_1) = \frac{1}{100} \quad P(D_1|H_2) = \frac{1}{200}$$
5.2 (2 points): What is the value of the following probability:

Answer: \[ P(D_1) = P(D_1|H_1)P(H_1) + P(D_1|H_2)P(H_2) = \frac{1}{100} \times \frac{1}{10} + \frac{1}{200} \times \frac{9}{10} = \frac{11}{2000} \]

5.3 (2 points): Which is the maximum a posteriori (MAP) hypothesis given \( D_1 \)? Show your working.

Answer: \( H_2 \) is the MAP hypothesis because

\[ P(D_1|H_1) = \frac{1}{100} \times \frac{1}{10} = \frac{1}{1000} \]

is less than \( P(D_1|H_2) = \frac{1}{200} \times \frac{9}{10} = \frac{9}{2000} \).

5.4 (2 points): Tomorrow you receive a report of a tank with number 75. Call this event \( D_2 \). Which is the MAP hypothesis given \( D_1 \) and \( D_2 \)? Show your working.

Answer: \( H_2 \) is the MAP hypothesis because

\[ P(D_1|H_1)P(D_2|H_1) = \frac{1}{100} \times \frac{1}{100} \times \frac{1}{10} = \frac{1}{100000} \]

is less than \( P(D_1|H_2)P(D_2|H_2) = \frac{1}{200} \times \frac{1}{200} \times \frac{9}{10} = \frac{9}{400000} \).

5.5 (2 points): If \( H_1 \) is true, how many reports (including the report from today and tomorrow) from the spy will it take to convince you that \( H_1 \) is true?

Answer: After \( k \) days,

\[ P(D_1 \cap D_2 \cap \ldots \cap D_k|H_1)P(H_1) = \left(\frac{1}{100}\right)^k \times \frac{1}{10} \]

and

\[ P(D_1 \cap D_2 \cap \ldots \cap D_k|H_2)P(H_2) = \left(\frac{1}{200}\right)^k \times \frac{9}{10} \]

Therefore

\[ \frac{P(D_1 \cap D_2 \cap \ldots \cap D_k|H_1)P(H_1)}{P(D_1 \cap D_2 \cap \ldots \cap D_k|H_2)P(H_2)} = \frac{1}{9/2^k} \]

so after 4 days we choose \( H_1 \) since \( P(D_1 \cap D_2 \cap \ldots \cap D_4|H_1)P(H_1) \geq P(D_1 \cap D_2 \cap \ldots \cap D_4|H_2)P(H_2) \).

Question 6. (10 points) Every year there is an 80% chance that you are exposed to the flu virus. If you are exposed to the flu virus, there is a 50% chance you’ll develop multiple flu symptoms. If you are not exposed there’s a 0% chance you’ll develop multiple flu symptoms. If you develop multiple flu symptoms, there’s a 75% chance that you have a high temperature. Assume there’s a 10% chance you’ll have a high temperature even if you don’t develop multiple flu symptoms. Let \( V = 1 \) if you are exposed to the virus and \( V = 0 \) otherwise. Let \( S = 1 \) if you develop multiple flu symptoms and \( S = 0 \) otherwise. Let \( T = 1 \) if you have a high temperature and \( T = 0 \) otherwise. You should assume that \( V \) and \( T \) are independent conditioned on \( S \).
6.1 (3 points): Enter the values for the following probabilities:

\[
P(V = 1, S = 0, T = 1) = \frac{4}{5} \times \frac{1}{2} \times \frac{1}{10} = \frac{1}{25} = 0.04
\]
\[
P(V = 1, S = 1, T = 1) = \frac{4}{5} \times \frac{1}{2} \times \frac{3}{4} = \frac{3}{10} = 0.3
\]
\[
P(V = 1, T = 1) = 0.04 + 0.3 = 0.34
\]

6.2 (3 points): Enter the values for the following probabilities:

\[
P(S = 0) = P(V = 0, S = 0) + P(V = 1, S = 0) = \frac{1}{5} \times 1 + \frac{4}{5} \times \frac{1}{2} = \frac{3}{5}
\]
\[
P(S = 1) = 1 - 3/5 = 2/5
\]
\[
P(T = 1) = P(S = 0)P(T = 1|S = 0) + P(S = 1)P(T = 1|S = 1) = \frac{3}{5} \times \frac{1}{10} + \frac{2}{5} \times \frac{3}{4} = \frac{9}{25}
\]

6.3 (2 points): Are S and T independent? Justify your answer numerically.

Answer: They are not independent since \(P(T = 1) = \frac{9}{25}\) but \(P(T = 1|S = 1) = \frac{3}{4}\).

6.4 (2 points): Draw the Bayesian network for the random variables V, S, and T.

Answer: A possible Bayesian network is:

```
V -- S --> T
```

but other networks are also valid.
Standard Random Variables

- **Bernoulli Random Variable** with parameter $p \in [0, 1]$

  $$P(X = k) = \begin{cases} 
  1 - p & \text{if } k = 0 \\
  p & \text{if } k = 1
  \end{cases}, \quad E(X) = p, \quad \text{var}(X) = p(1 - p)$$

- **Binomial Random Variable** with parameters $p \in [0, 1]$ and $N \in \{1, 2, 3, \ldots\}$

  For $k \in \{0, 1, 2, \ldots, N\}$:
  $$P(X = k) = \binom{N}{k} p^k (1-p)^{N-k}, \quad E(X) = Np, \quad \text{var}(X) = Np(1-p)$$

- **Geometric Random Variable** with parameter $p \in [0, 1]$

  For $k \in \{1, 2, 3, \ldots\}$:
  $$P(X = k) = (1 - p)^{k-1} \cdot p, \quad E(X) = \frac{1}{p}, \quad \text{var}(X) = (1 - p)/p^2$$

- **Poisson Random Variable** with parameter $\lambda > 0$:

  For $k \in \{0, 1, 2, \ldots\}$:
  $$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad E(X) = \lambda, \quad \text{var}(X) = \lambda$$

- **Discrete Uniform Random Variable** with parameters $a, b \in \mathbb{Z}$ and $a < b$

  For $k \in \{a, a+1, \ldots b\}$:
  $$P(X = k) = \frac{1}{b - a + 1}, \quad E(X) = \frac{a + b}{2}, \quad \text{var}(X) = \frac{(b - a + 1)^2 - 1}{12}$$

Bayes Formula

- If $A_1, \ldots, A_n$ partition $\Omega$ then for any event $B$:

  $$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)}$$