HOMEWORK 6
PROBABILITY & STATISTICS *

10-606 MATHEMATICAL FOUNDATIONS FOR MACHINE LEARNING

START HERE: Instructions

- **Collaboration Policy**: Please read the collaboration policy in the syllabus.
- **Late Submission Policy:** See the late submission policy in the syllabus.
- **Submitting your work:** You will use Gradescope to submit answers to all questions.
  - **Written:** For written problems such as short answer, multiple choice, derivations, proofs, or plots, please use the provided template. Submissions can be handwritten onto the template, but should be labeled and clearly legible. If your writing is not legible, you will not be awarded marks. Alternatively, submissions can be written in LaTeX. Each derivation/proof should be completed in the boxes provided. To receive full credit, you are responsible for ensuring that your submission contains exactly the same number of pages and the same alignment as our PDF template.
  - **Latex Template:** https://www.overleaf.com/read/fkbjfjcjbtfg#393528

Question	Points
Probability	15
Random Variables	7
Gaussian Random Variables	15
Total:	37

*Compiled on Monday 30th September, 2024 at 01:30

Instructions for Specific Problem Types

For "Select One" questions, please fill in the appropriate bubble completely:

Select One: Who taught this course?

- ● Matt Gormley
- ○ Marie Curie
- ○ Noam Chomsky

If you need to change your answer, you may cross out the previous answer and bubble in the new answer:

Select One: Who taught this course?

- ● Henry Chai
- ○ Marie Curie
- ⊗ Noam Chomsky

For "Select all that apply" questions, please fill in all appropriate squares completely:

Select all that apply: Which are scientists?

- ■ Stephen Hawking
- ■ Albert Einstein
- ■ Isaac Newton
- □ I don't know

Again, if you need to change your answer, you may cross out the previous answer(s) and bubble in the new answer(s):

Select all that apply: Which are scientists?

- ■ Stephen Hawking
- ■ Albert Einstein
- ■ Isaac Newton
- □ I don't know

For questions where you must fill in a blank, please make sure your final answer is fully included in the given space. You may cross out answers or parts of answers, but the final answer must still be within the given space.

Fill in the blank: What is the course number?

10-606 10-6067

1 Probability (15 points)

1. (2 points) Suppose we flip a fair coin 20 times, and record the sequence of heads (H) and tails (T) that is observed. Which of the following subsequences is most likely to appear within the full sequence?

   - ◉ HHHH
   - ◯ HHHHH
   - ◯ HTHHH
   - ◯ HTHTH
   - ◯ HTHTHT

2. (2 points) Suppose we have a six-sided die (numbered 1 to 6) that is not evenly weighted (i.e. it is not fair). Define A to be the event of rolling a number less than 4. Define B to be the event of rolling an even number. Assume we know that $P(A) > 0.5$. Which of the following is definitely true? **Select all that apply.**

   - ◯ ☐ $P(B) < 0.5$
   - ◯ ☐ $P(A \wedge B) < 0.5$
   - ◯ ☐ $P(A) > P(B)$
   - ◉ ☐ $P(A \vee B) > P(A)$
   - ◯ ☐ None of the above

3. (2 points) Assume there are three events, A, B, and C. Also, assume that $P(A \vee B) = 0.5$, and $P(C) = 0.5$. Which of the following is true? **Select all that apply.**

   - ◯ ☐ A and B are disjoint
   - ◯ ☐ A, B, and C are disjoint
   - ◉ ☐ $P(C) > P(A \wedge B)$
   - ◉ ☐ $P(A \wedge B \wedge C) < 0.5$
   - ◯ ☐ None of the above

4. (4 points) Consider the following experiment: you stand on a street corner in Pittsburgh and survey people. You ask two questions: "Do you like computer science?" and "Do you like the Steelers?" You find that 67% of passers by respond "yes" to the first question, and 90% of passers by respond "yes" to the second question. Furthermore, 65% of passers by respond "yes" to both questions.

   What is the probability of finding a person who responds "no" to both questions?

   $P(A \wedge B) = P(A) + P(B) - P(A \wedge B)$
   $= 0.92$
   $P(\text{no to both}) = 1 - P(A \wedge B)$
   $= 0.08$ or 8%.

5. (5 points) Suppose A and B are independent events. Prove that A^c and B^c are also independent.

$P(A \wedge B) = P(A) \cdot P(B)$

$P(A^c \wedge B^c) = P(A^c) \cdot P(B^c)$

$P(A^c) = 1 - P(A)$

$P(B^c) = 1 - P(B)$

$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

$P(A \vee B) = P(A) + P(B) - P(A) \cdot P(B)$

$P(A^c \wedge B^c) = 1 - P(A \vee B)$

$= 1 - (P(A) + P(B) - P(A) \cdot P(B))$

$P(A^c \wedge B^c) = 1 - P(A) - P(B) + P(A) \cdot P(B)$

$P(A^c) \cdot P(B^c) = 1 - P(A) - P(B) + P(A) \cdot P(B)$

$P(A^c \wedge B^c) = P(A^c) \cdot P(B^c)$

Hence, A^c and B^c are also independent.

2 Random Variables (7 points)

Consider two random variables $X \in \{1, 2, 3, 4\}$ and $Y \in \{1, 2\}$. The marginal distribution of X is $P(X) = [0.1, 0.4, 0.3, 0.2]$

The conditional distribution of $Y = 1$ for the four values of X is given by $P(Y = 1 \mid X) = [0.7, 0.7, 0.1, 0.1]$

(Note that both of the above equations represent probability tables, as described in lecture. The free variable in both cases is X.)

1. (2 points) What is $P(X = 2 \mid Y = 1)$?

   0.7

2. (2 points) Under the same assumptions as above, what is $P(X = 1 \mid Y = 2)$?

   0.05

3. (1 point) In the probability table $P(X, Y, Z)$, suppose that $X \in \{1, 2\}$, $Y \in \{1, 2, 3\}$, and $Z \in \{1, 2, 3\}$. How many entries does this table have?

   18

4. (2 points) Consider that we flip 5 fair coins. What is probability of at least 4 heads?

   $\frac{3}{16}$

3 Gaussian Random Variables (15 points)

Consider a mixture of two Gaussian distributions:

$$0.4\mathcal{N}\left(\begin{pmatrix}10\\2\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right) + 0.6\mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}8.4 & 2\\2 & 1.7\end{pmatrix}\right).$$

1. (4 points) Compute the marginal distributions for each dimension.

for x: $\mu_x = 10$, $\sigma^2_x = 1$ — for y: $\mu_y = 2$, $\sigma^2_y = 1$

$N\left(\begin{pmatrix}10\\2\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right)$

for x: $\mu_x = 0$, $\sigma^2_x = 8.4$ — for y: $\mu_y = 0$, $\sigma^2_y = 1.7$

$N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}8.4 & 2\\2 & 1.7\end{pmatrix}\right)$

first dimension:
$\mu_x = 4$
$\sigma^2_x = 29.44$
marginal distribution:
$x \sim N(4, 29.44)$
second dimension:
$\mu_y = 0.8$ $\sigma^2_y = 2.38$
marginal distribution:
$y \sim N(0.8, 2.38)$

2. (6 points) Compute the mean, mode and median for each marginal distribution.

Mean:
for x: 4
for y: 0.8

$\mu_x = 0.4 \cdot 10 + 0.6 \cdot 0 = 4$

$\mu_y = 0.4 \cdot 2 + 0.6 \cdot 0 = 0.8$

Mode: Gaussian distribution
for x = 4
for y = 0.8

Median → gaussian is symmetric
So median = mean.

for x = 4
for y = 0.8

3. (5 points) Compute the mean and mode for the two-dimensional (joint) distribution.

$\mu_1 = \begin{pmatrix}10\\2\end{pmatrix}$, weight = 0.4

$\mu_2 = \begin{pmatrix}0\\0\end{pmatrix}$; weight = 0.6

$\mu_{joint} = 0.4\begin{bmatrix}10\\2\end{bmatrix} + 0.6\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}4\\0.8\end{bmatrix}$

mode is closer to $\begin{bmatrix}0\\0\end{bmatrix}$ since the weight is higher.

4 Collaboration Questions

After you have completed all other components of this assignment, report your answers to these questions regarding the collaboration policy. Details of the policy can be found in the syllabus.

1. Did you receive any help whatsoever from anyone in solving this assignment? If so, include full details.
2. Did you give any help whatsoever to anyone in solving this assignment? If so, include full details.
3. Did you find or come across code that implements any part of this assignment? If so, include full details.

Your Answer

1. No help received.
2. No help given.
3. No code found or implemented.