

---

# CMPSCI 240: Reasoning Under Uncertainty

## Third Midterm Exam

April 10, 2013.

---

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.

Question	Value	Points Earned
1	10	
2	11	
3	10	
4	12	
5	12	
6 (Extra Credit)	5	
Total	50+5	

**Question 1.** (10 points) Indicate whether each of the following four statements is TRUE or FALSE. No justification is required.

**1.1** (2 points): If  $A, B$ , and  $C$  are events with  $P(A) \leq 0.1, P(B) \leq 0.1$ , and  $P(C) \leq 0.1$  then  $P(A \cup B \cup C) \leq 0.3$ .

**1.2** (2 points): Suppose you throw 2 balls and each ball is equally likely to land in any one of 10 bins, then the probability that the balls land in different bins is  $9/10$ .

**1.3** (2 points): If there exists  $a$  and  $b$  such that  $P(X = a, Y = b) = P(X = a)P(Y = b)$  then  $X$  and  $Y$  are independent random variables.

**1.4** (2 points): If  $X, Y$ , and  $Z$  are any three random variables then  $E(X + Y + Z) = E(X) + E(Y) + E(Z)$ .

**1.5** (2 points): What's the expected number of times you need to toss a fair coin until you see both a heads and a tails at least once? Circle the correct answer.

- (a) 1
- (b) 2
- (c) 3
- (d) 4

**Question 2.** (11 points) When you get home, you have two hypotheses:

$$H_1 = \text{"your housemate is home"} \quad \text{and} \quad H_2 = \text{"your housemate is not home."}$$

Let  $D_1$  be the observation that your housemate's car is parked outside. Let  $D_2$  be the observation that the house lights are on.

**2.1** (1 points): *Over the last 100 times you've returned home, your housemate has already been home 60 times. What values would you choose for the priors?*

$$P(H_1) = \quad \quad \quad P(H_2) =$$

**2.2** (2 points): *Of the 60 times your housemate has already been home, his car was parked outside 40 times and the lights were on 20 times. Of the 40 times your housemate wasn't already home, his car was parked outside 10 times and the house lights were on 20 times. Given this data, what probabilities would you choose for the likelihoods?*

$$P(D_1|H_1) = \quad \quad \quad P(D_2|H_1) = \quad \quad \quad P(D_1|H_2) = \quad \quad \quad P(D_2|H_2) =$$

**2.3** (3 points): *If you observe  $D_1$  and  $D_2$  which hypothesis would a Naive Bayes Classifier pick? Justify your answer.*

**2.4** (3 points): Suppose you observe  $D_1$  and  $D_2^c$ , i.e., your housemate's car is parked outside but the house lights are off. What's values would you pick for the following probabilities:

$$P(D_2^c|H_1) = \qquad P(D_2^c|H_2) =$$

and which hypothesis would a Naive Bayes Classifier pick? Justify your answer.

**2.5** (2 points): Suppose that out of the 60 times your housemate has already been home, there were 0 times that both his car was parked and the house lights were on. Out of the 40 times your housemate wasn't home, there was 1 time that both his car was parked and the house lights were on. Which hypothesis would you now pick if you observe  $D_1$  and  $D_2$ ? If you had more historical data do you think your decision would change and why?

**Question 3.** (10 points) You're at a casino and are given a dice. Unfortunately you're not sure it's a fair dice. You have two hypotheses:

$H_1$  = "it's a fair dice with equal probability of landing 1, 2, 3, 4, 5, or 6"

$H_2$  = "it's an unfair dice with equal probability of landing 1, 2, 3, 4, or 5 but never lands as a 6"

Your prior probabilities for these hypotheses are  $P(H_1) = 2/3$  and  $P(H_2) = 1/3$ . Let  $D$  be the event that when you roll the dice it lands as a "1".

**3.1** (2 points): What's the value of the likelihoods:

$$P(D|H_1) = \qquad P(D|H_2) =$$

**3.2** (2 points): Given that event  $D$  is observed, which is the maximum a posteriori (MAP) hypothesis? Show your working.

**3.3** (2 points): What's the value of  $P(H_1|D)$ ?

**3.4** (2 points): Let  $C_k$  be the event that, when you roll the dice  $k$  times, it never lands as a "6". What's the value of the likelihoods:

$$P(C_k|H_1) = \qquad P(C_k|H_2) =$$

**3.5** (2 points): Find the smallest value of  $k$  such that if  $C_k$  is observed, the maximum a posteriori hypothesis is  $H_2$ ? Show your working.

**Question 4.** (12 points) Suppose you're listening to the radio and a new song comes on. Let  $X = 1$  if you like the song and let  $X = 0$  if you don't like the song. Let  $Y = 1$  if you turn the radio off and let  $Y = 0$  if keep the radio on.

	$Y = 0$	$Y = 1$
$X = 0$	$5/12$	$3/12$
$X = 1$	$7/36$	$5/36$

So, for example  $P(X = 0, Y = 1) = 3/12$ .

**4.1** (2 points): Enter the values for the following probabilities:

$$P(X = 0) = \quad P(X = 1) = \quad P(Y = 0) = \quad P(Y = 1) =$$

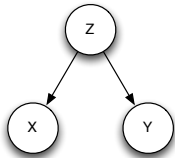
**4.2** (2 points): Are  $X$  and  $Y$  independent? Justify your answer.

**4.3** (2 points): Let  $Z = 1$  if the song is by Bon Jovi and let  $Z = 0$  if the song is by another band. If  $P(Z = 1) = 1/3$ ,  $P(X = 1|Z = 1) = 1/2$ , and  $P(Y = 1|Z = 1) = 1/2$ , what are the values of:

$$P(X = 0|Z = 1) = \quad P(Y = 0|Z = 1) =$$

**4.4** (2 points): What's the value of  $P(Z = 1|Y = 1)$ ?

**4.5** (2 points): What's the value of  $P(X = 0, Y = 1, Z = 1)$  if  $X$  and  $Y$  are independent conditioned on  $Z$ , i.e., the relationship between  $X, Y$ , and  $Z$  can be represented by the following Bayesian network:



**4.6** (2 points): What's the value of  $P(X = 0, Y = 1|Z = 1)$ ?

**Question 5.** (12 points) Suppose you repeatedly throw an *unbiased* six-sided dice. Define the following two random variables:

$X$  = number of throws taken when the dice lands as a “4” for the first time

$Y$  = number of throws taken when you’ve seen all of the six possible values.

For example, if the results of the throws were 2,4,3,3,5,6,1,4,2,... then  $X = 2$  and  $Y = 7$ .

**5.1** (2 points): What’s the value of  $P(X = 1)$ ?

**5.2** (2 points): What’s the value of  $E(X)$ ?

**5.3** (2 points): What’s the value of  $P(Y = 6)$ ?

**5.4** (2 points): What’s the value of  $E(Y)$ ?

**5.5** (2 points): Are  $X$  and  $Y$  independent random variables? Justify your answer.

**5.6** (2 points): What’s the value of  $P(X = Y)$ ? Justify your answer. Hint:  $P(X = Y) = \sum_{k=1}^{\infty} P(X = k, Y = k)$  but there’s a much easier way to solve this question.

**Question 6.** (5 points) *Extra Credit:* This question has two unrelated parts.

**6.1** (2 points): From the information in question 4, compute the following values:

$$P(X = 1|Z = 0) = \qquad P(Y = 1|Z = 0) =$$

**6.2** (3 points): Suppose you toss a fair coin  $n$  times and let  $X_i = 1$  if the  $i$ th coin toss is heads and  $X_i = 0$  otherwise. Let  $R$  be the event that there's a run of at least  $2 \log_2 n$  heads, i.e., there exists  $j$  such that  $X_j = X_{j+1} = \dots = X_{j+2 \log_2 n - 1} = 1$ . Show that  $P(R)$  becomes very small when  $n$  becomes large. Hint: Let  $A_j$  be the event that  $X_j = X_{j+1} = \dots = X_{j+2 \log_2 n - 1} = 1$ .



### 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  $N \in \{1, 2, 3, \dots\}$ :**

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

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

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

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

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

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

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

### Bayes Formula

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

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