CompSci 527 Midterm Exam Sample

The exam will be closed-book, closed-notes, and you will not be allowed to have anything other than the exam and a pen/pencil and an eraser on your desk. The amount of space provided under each question is *not* an indication of the length of the answer. Materials covered are all of chapter 2 and sections 3.1, 3.3, 3.4, 3.5, 4.1, 4.2, 4.3, 6.1, 6.2 of the book, class notes on basic definitions for probabilities, the first two homework assignments, and this sample.

1. The following table shows the joint probability $p(x, y)$ for random variables X and Y. The variable X can take values in $\{1, 2, 3\}$ and the variable Y can take values in $\{0, 1\}$. Fill the empty cells in the tables below. Represent values as irreducible fractions of integers ($2/3$ but not $4/6$) or as decimals (0.4). The decimal representation of $2/3$ is $0.\bar{6}$. If in doubt about your numbers, show your calculations *clearly* for partial credit.

$p(x, y)$	$x = 1$	2	3
$y = 0$	0.1	0.1	0.2
1	0.3	0.0	

$p(y)$	
$y = 0$	
1	

	$x = 1$	2	3
$p(x)$			

$\mathsf{E}[X] = \boxed{}$ $\sigma_X^2 = \boxed{}$

$\mathsf{E}[Y] = \boxed{}$ $\sigma_Y^2 = \boxed{}$

$p(x\|y)$	$x = 1$	2	3
$y = 0$			
1			

$p(y\|x)$	$x = 1$	2	3
$y = 0$			
1			

2. In the table below, write the joint probability $p(x, y)$ of two *independent* random variables X and Y that have the same marginal distributions $p(x)$ and $p(y)$ as in the previous problem.

$p(x, y)$	$x = 1$	2	3
$y = 0$			
1			

3. A Bernoulli random variable X with parameter λ has universe $\{0, 1\}$ and distribution $p(0) = \mathbb{P}[X = 0] = 1 - \lambda$ and $p(1) = \mathbb{P}[X = 1] = \lambda$. Show that the mean and variance of X are $m_X = \lambda$ and $\sigma_X^2 = \lambda(1 - \lambda)$.

4. The Dirichlet distribution over $K = 2$ continuous values λ_1, λ_2 and with parameters $\alpha_1 = \alpha_2 = 2$ has density

$$p(\lambda_1, \lambda_2) = c\lambda_1\lambda_2 \quad \text{wherever} \quad \lambda_1 + \lambda_2 = 1 \quad \text{and} \quad \lambda_1 \geq 0, \ \lambda_2 \geq 0 \,,$$

and is zero elsewhere. What is the value of c? Show your reasoning. [Hint: Compute c from first principles rather than trying to remember a formula. What part of the (λ_1, λ_2) plane satisfies the constraints in the definition of p?]

5. The *triangle distribution with parameter* θ is a continuous probability distribution defined as follows:

$$p(z|\theta) = \begin{cases} 1 + z - \theta & \text{for } \theta - 1 \le z \le \theta \\ 1 - z + \theta & \text{for } \theta \le z \le \theta + 1 \\ 0 & \text{elsewhere} \end{cases}$$
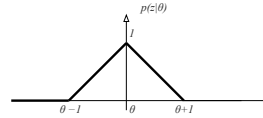
and is shown in the figure. You are given a training set with two samples

$$z_1 = 0.3 \quad \text{and} \quad z_2 = 0.5$$

drawn out of the triangle distribution $p(z|\theta)$, where the unknown parameter θ is a random variable that has uniform distribution between 0 and 0.3.

(a) Find the Maximum Likelihood (ML) estimate $\hat{\theta}_{ML}$ of θ and explain your answer. [Hint: A drawing and symmetry considerations may help. Brute force is not recommended.]

(b) Find the Maximum a Posteriori (MAP) estimate $\hat{\theta}_{MAP}$ of θ and explain your answer. [Hint: Do you even have a choice?]

6. A company makes pencils, and 20 percent of its pencils are broken. A quality check reports all the good pencils as such, but reports 10 percent of the broken ones as good. Use the following event symbols in your formulas (you may not need all of them):

$$B \quad : \quad \text{a pencil is broken}$$
$$G \quad : \quad \text{a pencil is good}$$
$$P \quad : \quad \text{a pencil passes the check}$$
$$F \quad : \quad \text{a pencil fails the check}$$

(a) What is the probability that a pencil that passes the check is good? Show your reasoning. You may leave your answer as a ratio of two (but no more) decimals (example: $0.42/0.65$).

(b) If the company sells all pencils that pass the check, what is the probability that it sells a bad one? Show your reasoning. You may leave your answer as a ratio of two (but no more) decimals (example: $0.42/0.65$).

7. A world quantity w with standard uniform prior

$$p(w) = \begin{cases} 1 & \text{for } 0 \le w \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

is observed through scalar measurements with a standard normal likelihood

$$p(x|w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-w)^2}{2}} \ .$$

(a) Write an explicit formula for the Bayes regressor

$$\hat{w} = f(x) \ .$$

"Explicit" here entails that your final formula cannot contain the symbols $\int$, p, or w. You may of course use these symbols in your derivation. [Hint: Once you have a (correct!) formula for the joint density of x and w, consider the three cases $x < 0$ and $0 \le x \le 1$ and $x > 1$ separately. A drawing may help.]

(b) Sketch your regressor function $f(x)$, and label the axes accurately.

8. You are asked to design an inference system to track the puck in an ice hockey game from a video recording of the game. Assume that you only need to track the puck when it is visible. Suppose that it is straightforward to model the imaging process as a function $\mathbf{y} = g(\mathbf{w})$ that computes the ideal image $\mathbf{y}$ of the puck from the 3D coordinates $\mathbf{w}$ of the puck's centroid in the world. The imaging process is noisy, so the real image is

$$\mathbf{x} = \mathbf{y} + \mathbf{n}$$

where $\mathbf{n}$ is random noise of known probability distribution.

(a) Is your inference system a classifier or a regressor? Why?

(b) Would you design a generative or a discriminative inference system? Justify your answer.

(c) What else do you need to know or estimate to complete your design, in addition to the function g and the probability distribution of $\mathbf{n}$? Why?

(d) What does a training set for this problem look like? Describe a training sample, and state what annotations (human-entered information) are needed for it. [No need to state how training samples are collected.]