MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Media Arts and Sciences Section

Course 4.902: Pattern Recognition and Analysis — Assigned: 2/2/93

Problem Set No. 1 — Due: 2/9/93

READING ASSIGNMENT: Therrien, Chapters 1 and 2.

This problem set consists of four warm-up exercises and four problems. Remember that you are *not* allowed to look at old problem sets. Collaboration is allowed and encouraged.

Exercise 1: (2 points)

A random experiment consists of flipping a fair coin ten times. How many possible outcomes are there? (Here, an *outcome* is a particular sequence of ten coin flips.)

Exercise 2: (3 points)

Let Ω be the set of all possible elementary outcomes of a random experiment, and let the events $A \subset \Omega$ and $B \subset \Omega$ be disjoint (that is, let A and B have no members in common). The events A and B are ______.

a) definitely independent
b) definitely not independent
c) possibly but not definitely independent

Exercise 3: (5 points)

Suppose that three fair, independent dice are rolled, and the resulting numbers summed. What is the probability that the sum is 9? What is the probability that the sum is 10?

Exercise 4: (10 points)

Let $\mathbf{x}$ be a random vector with distribution function $P_{\mathbf{x}}(\hat{\mathbf{x}})$ and density function $p_{\mathbf{x}}(\hat{\mathbf{x}})$. Assume that $P_{\mathbf{x}}(\hat{\mathbf{x}})$ is continuous and differentiable everywhere in every dimension (that is, assume $p_{\mathbf{x}}(\hat{\mathbf{x}})$ contains no impulses).

a) Is it possible that $p_{\mathbf{x}}(\hat{\mathbf{x}}) > 1$ for some $\hat{\mathbf{x}}$?
b) Is it possible that $P_{\mathbf{x}}(\hat{\mathbf{x}}) > 1$ for some $\hat{\mathbf{x}}$?
c) Let $\hat{\mathbf{x}}$ be some particular vector in the domain of $P_{\mathbf{x}}$. What is the probability that $\mathbf{x} = \hat{\mathbf{x}}$?
d) Suppose $\mathbf{y}$ is a 2-dimensional random vector which assumes the values $(1, 2)$ and $(0, 1)$, each with probability $1/2$. What is the value of the distribution function $P_{\mathbf{y}}(\hat{\mathbf{y}})$ at the point $\hat{\mathbf{y}} = (1/2, 4)$?

Problem 1: (20 points)

Let x and y be discrete random variables.

a) Show that $E(x + y) = E(x) + E(y)$.

b) If x and y are independent, show that $E(xy) = E(x)E(y)$; that is, x and y are uncorrelated.

c) Is it possible for x and y to be dependent but uncorrelated? If so, give an example; if not, show why not.

d) If x and y are independent, show that $Var(x + y) = Var(x) + Var(y)$.

Problem 2: (20 points)

A box contains two dice, one fair and the other loaded, so that for the first one $p(1) = p(2) = \cdots = p(6) = 1/6$ and for the second one $p(1) = \cdots = p(5) = 1/15$; $p(6) = 2/3$. (Here, $p(n)$ denotes the probability of the die showing the number n.) We choose one die from the box at random (with probability of 1/2 for each die), and roll it.

a) If the die is rolled once and the outcome is the number 6, what is the probability that the loaded die has been selected?

b) If the same die is rolled again and the outcome is some number other than 6, what is the probability that the loaded die has been selected, taking both rolls into account?

Problem 3: (20 points)

A generous friend gives you a present in the following way. She shows you two sealed envelopes, marked x and y. You are told that both contain money, and that one contains exactly twice as much as the other (you are not told which contains the larger amount). You may select either of the two and keep its contents.

a) You select x, open it, and find $\hat{x}$ dollars. What is the expected value of the amount in envelope y, in terms of $\hat{x}$? Make whatever reasonable assumptions you wish, but state them clearly and justify them (in a sentence or two). Interpret your answer (again, in a sentence or two).

b) You now learn that your friend prepared the two envelopes in the following way. (Assume in this part of the problem that dollar amounts can vary continuously, instead of having to be quantized to the nearest cent.)

   First, she used a computer to generate a continuous random number uniformly distributed between 20 and 50. She then put that many dollars in one envelope, and twice that many dollars in the other envelope. Finally, she flipped a fair coin to decide which envelope should be marked x.

   Based on this description of how the envelopes were prepared, what is $p_x(\hat{x})$, the marginal probability density function for x? What is $p_{y|x}(\hat{y}|\hat{x})$, the conditional probability density function for y, given that $x = \hat{x}$? Finally, what is $E(y|\hat{x})$, the conditional expected value for y given that $x = \hat{x}$?

Problem 4: (20 points)

Two of the faces of a fair die are colored green; the remaining four are blue. Suppose that the die is rolled twelve times, independently and fairly each time, with the result of each roll recorded as either G (green) or B (blue). Of the 4096 possible sequences of 12 rolls, three are given below.

i) $B\ G\ B\ G\ G\ B\ G\ B\ B\ G\ G\ B$
ii) $B\ B\ B\ B\ B\ B\ B\ B\ B\ B\ B\ B$
iii) $B\ G\ B\ B\ B\ B\ G\ B\ B\ G\ B\ G$

a) In your opinion, which of these three sequences seems most believable as an observed particular outcome of the experiment? That is, if I told you that I already did the experiment and that only one of the three sequences really happened (I made up the other two), which one would you say really happened? Explain your answer in a sentence or two. (There is no right or wrong answer to this part of the question, but some answers are better than others, depending on the explanation you give.)

b) Calculate the probability of each of the three sequences. Which sequence has the greatest probability? Do you think that your answer to part (a) should have been the same? Why or why not? Explain in a sentence or two.

c) Again considering twelve rolls of the die as the basic experiment, find an expression for the probability that there will be exactly k occurrences of B, without regard to ordering. For what value of k is this expression maximum? (Remember, k is an integer.) Does your answer make you more confident about your answer to part (a)?

Puzzle: (0 points, but fun if you haven't seen it before)
This problem is completely optional, and will not count in any way to your grade.

A fellow named Monty Hall shows you three boxes. The boxes are covered, so that you can't see what's inside. He tells you that one of the boxes contains a five dollar bill, and that the other two are empty. He knows which box contains the money, but you don't.

You are asked to choose one of the three boxes. After you choose it, Monty will open one of the two remaining boxes, always choosing one that is empty (he can do this because he knows where the money is). He will then offer you the chance to trade your box for the remaining unopened box.

Assume that you will play this game many times, so that maximizing the probability of finding the money each time will maximize your average total winnings. With this objective, are you better off keeping the box you initially choose, or accepting Monty's offer to trade it for the other unopened box? Or are your chances the same either way?