

Decision making and discrete random variables

Assigned reading: Review *Ross*, Chapters 2&3, read *Ross* Sections 4.1& 4.2, and read the *Notes on Decision Making for ECE 313 Fall 2002* (excerpted from notes of Prof. D.V. Sarwate).

Noncredit exercises: Chapter 4, Problems 2 and 7, and Theoretical Exercise 11.

Reminder: Hour Exam I: Monday, October 7, 7:00-8:00 p.m., Room 269 Everitt Lab This exam covers the material of problem sets 1-5 and the associated reading (*Ross* through Section 4.2, and the notes mentioned above), and lectures through September 30. You may bring one 8.5" by 11" sheet of notes to each hour exam. You may use both sides of the sheet, using font size 10 or larger (or similar handwriting size). The exam is closed book otherwise. Calculators, laptop computers, tables of integrals, etc. are not permitted.

1. Sensor fusion

Two motion detectors are used to detect the presence of a person in a room, as part of an energy saving temperature control system. The first sensor outputs a value X and the second sensor outputs a value Y . Both outputs have possible values $\{0, 1, 2\}$, with larger numbers tending to indicate that a person is present. Let H_0 denote the event (or hypothesis) that a person is absent and let H_1 denote the event that a person is present. The likelihood matrices for X and for Y , giving the probabilities for possible outputs given H_0 or H_1 , are as follows:

	$X = 0$	$X = 1$	$X = 2$
H_0	0.8	0.1	0.1
H_1	0.1	0.3	0.6

	$Y = 0$	$Y = 1$	$Y = 2$
H_0	0.7	0.2	0.1
H_1	0.1	0.1	0.8

For example, $P[Y = 2|H_1] = 0.8$. Suppose that, given one of the hypotheses is true, the sensors provide conditionally independent readings, so that $P[X = i, Y = j|H_k] = P[X = i|H_k]P[Y = j|H_k]$ for $0 \leq i, j \leq 2$ and $0 \leq k \leq 1$. Since both X and Y are observed, there are nine possible values for the observation (X, Y) .

(a) Find the likelihood matrix for the observation given the hypothesis. This matrix specifies $P[X = i, Y = j|H_k]$ for each hypothesis H_k and for each possible observation value (i, j) . (The matrix has two rows, the first row corresponding to H_0 and the second row corresponding to H_1 . The matrix has nine columns of numbers, each corresponding to the nine possible values of the observation (X, Y) : 00, 01, 02, 10, 11, 12, 20, 21, 22. The sum of the entries in each row is one.) Also, indicate for each of the nine possible observation values, the maximum likelihood (ML) decision of which hypothesis is true. This is done by simply circling the larger number in each column. Break ties in favor of H_1 .

(b) For the ML decision rule, compute the false alarm (conditional) probability and miss (conditional) probability defined by

$$p_{\text{false_alarm}} = P[\text{decide } H_1 \text{ is true} | H_0]$$

$$p_{\text{miss}} = P[\text{decide } H_0 \text{ is true} | H_1].$$

For example, $p_{\text{false_alarm}}$ is the sum of all the entries in the first row of the likelihood matrix that were *not* circled in part (a).

(c) Suppose it is decided on the basis of observations of occupancy patterns that $\pi_0 = P[H_0] = 0.8$ and $\pi_1 = P[H_1] = 0.2$. These probabilities π_0 and π_1 are called prior probabilities, because they are assumed to hold before the detector outputs are taken into account. Compute the *joint probability matrix*, which specifies $P[X = i, Y = j, H_k]$ for each hypothesis H_k and for each possible observation value (i, j) . (The 18 numbers in the matrix sum to one.) Also, indicate for each possible observation value, the maximum a posteriori probability (MAP) decision of which hypothesis is true by circling the larger element of each column of the joint probability matrix.

(d) For the MAP decision rule, compute $p_{\text{false_alarm}}$, p_{miss} , and the unconditional probability of error $p_e = \pi_0 p_{\text{false_alarm}} + \pi_1 p_{\text{miss}}$ (we called this p_{ave} on previous problem set), using the same prior probabilities as in part (c). (Hint: The conditional probability $p_{\text{false_alarm}}$ for the MAP decision rule is the sum of all the entries in the first row of the *likelihood matrix* that correspond to entries *not* circled in the first row of the *joint probability matrix*. The conditional probability p_{miss} is computed similarly, using the second rows.)

(e) Compute, for the same priors given in (c), the error probability p_e for the ML rule. Just use $p_e = \pi_0 p_{\text{false_alarm}} + \pi_1 p_{\text{miss}}$ with the $p_{\text{false_alarm}}$ and p_{miss} computed in part (b). Comment: Among all decision rules, the MAP rule minimizes p_e , so your answer to (e) should be greater than or equal to your answer to part (d). On the other hand, the MAP rule requires knowledge of π_0 and π_1 whereas the ML rule does not.)

2. Simple discrete random variables

Consider the following experiment. Two fair dice are rolled. To model this experiment, take the sample space S to be $S = \{(i, j) : 1 \leq i, j \leq 6\}$. This notation means that both i and j are between 1 and 6 inclusive. Suppose the 36 outcomes are equally likely. Let Y be the random variable denoting the maximum of the numbers on the dice. That is, if the outcome is (i, j) , then $Y = \max\{i, j\}$. For example, if the first die shows 4 and the second die shows 2, then the outcome is $(4, 2)$ and $Y = 4$. (a) Draw a 6×6 table, such that the entry in the i^{th} row and j^{th} column is $\max\{i, j\}$. This table explicitly shows Y as a function of the outcome (i, j) .

(b) What are the possible values of Y ?

(b) Find and sketch the probability mass function p for Y . This function is defined by $p(y) = P(Y = y)$. Make sure $p(y)$ is positive only if y is a possible value of Y , and that the sum of all the masses is one.

3. Computing a simple probability mass function

Problem 13 in Chapter 4 of *Ross*, page 172.

4. Maximum likelihood parameter estimation

A certain robot's computer vision system feeds a sequence of measurements to the robot's central processing unit, with each measurement in the set $\{0, 1, 2\}$. In total, ten measurements are taken. There is a parameter a to be estimated, which physically corresponds to the offset of the robot from a line painted on the floor. It is assumed that $-1 \leq a \leq 1$, and that each measurement has the probability mass function $(p(0), p(1), p(2)) = (\frac{1-a}{3}, \frac{1}{3}, \frac{1+a}{3})$. Thus, if $a > 0$ then the value 2 for a single measurement has greater probability than the value 1. Suppose that a is fixed while all ten observations are made, and suppose the observations are independent. Thus, for example, the probability the first two observations are 1 and the third observation is 2 is $p(1)^2 p(2)$.

(a) Express the probability that the sequence of measurements is $(2, 1, 0, 2, 0, 1, 2, 2, 0, 2)$ in terms of a .

(b) The *maximum likelihood estimate* of a , denoted by $\hat{a}_{ML}$ for the observations given in part (a) is the value of a that maximizes the answer to part (a) with respect to a . Find the numerical value of $\hat{a}_{ML}$ given the observation $(2, 1, 0, 2, 0, 1, 2, 2, 0, 2)$. (Hint: It might be slightly easier to maximize the log of your answer to part (a) with respect to a , which is equivalent.)

(c) Find the numerical value of $\hat{a}_{ML}$ given the observation $(0, 2, 0, 2, 2, 2, 2, 1, 0, 1)$. Compare to your answer in part (b).

(d) Give the value of $\hat{a}_{ML}$ in case a total of n measurements is made, with n_0 0's observed, n_1 1's observed, and n_2 2's observed, where $n_0 + n_1 + n_2 = n$.

5. Bayes decision making for given cost

An expressway traffic sensor counts the number of axles that roll over it in a one minute period. For example, a car is counted twice, and a long-haul trailer truck is counted five times. Let X denote the number of counts observed in a one minute interval. Let H_0 denote the event that there was no accident within the last minute, and let H_1 denote the event that there was an accident within the last minute. Based on empirical evidence, the following assumptions are made:

$$P(X = n | H_0) = \begin{cases} 0.0001 & \text{if } 1 \leq n \leq 300 \\ 0.002 & \text{if } 301 \leq n \leq 700 \\ 0.001 & \text{if } 701 \leq n \leq 870 \end{cases}$$

and

$$P(X = n | H_1) = 0.0025 \quad \text{if } 1 \leq n \leq 400.$$

(a) Compute the likelihood ratio $\Lambda(n) = \frac{P(X=n|H_1)}{P(X=n|H_0)}$ for $1 \leq n \leq 870$.

(b) What is the maximum likelihood (ML) decision rule? What are $p_{\text{false_alarm}}$ and p_{miss} for the ML rule?

(c) Suppose that the probability of an accident in a one minute period is $\pi_1 = P(H_1) = 10^{-4}$. What is the maximum a posteriori probability (MAP) decision rule? What are $p_{\text{false_alarm}}$, p_{miss} , and p_e for the MAP decision rule?

(d) Suppose that the cost of making a correct decision is zero (so $C_{00} = C_{11} = 0$). Suppose that a miss is decided to be 1000 times more costly than a false alarm, because if there is a false alarm, a secondary sensor will be consulted before taking drastic action. So take $C_{01} = 1000$ and $C_{10} = 1$. Find the decision rule that minimizes the average cost (i.e. find the Bayes decision rule). For that rule, find $p_{\text{false_alarm}}$, p_{miss} , and the average cost C_{ave} . (Hint: The rule is an LRT with threshold $\tau = \frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})}$ and $C_{\text{ave}} = \sum_{i,j} C_{ij} P(\{\text{decide } H_i \text{ is true}\} \cap H_j)$.)