

Hour Exam I, Monday February 26, 7 p.m. — 8 p.m., 269 Everitt Lab
One 8.5"×11" sheet of notes allowed. Coverage of material: Problem Sets 2-5

Hypothesis testing

Reading: *Notes on Decision-Making for ECE 413 Spring 2007* distributed in class.
Supplemental reading *Chapter 5: Decision-Making under Uncertainty* can be found on the class web page.

Problems:

- Returning to Problem 2 of Problem Set 5, let H_i denote the hypothesis that the transmitted signal is s_i , $i = 1, 2, 3$. The matrix shown on page 1 of Problem Set 5 is thus the likelihood matrix for the decision problem.
 - What is the maximum-likelihood (ML) decision rule?
 - For the prior probabilities given in Problem Set 5, find the maximum *a posteriori* probability (MAP) decision rule (a.k.a. minimum-error-probability (MEP) or Bayesian decision rule) and the error probability of the MAP rule. What is the error probability of the ML decision rule?
 - Repeat part (b) for prior probabilities $\pi_1 = 0.1, \pi_2 = 0.3, \pi_3 = 0.6$. What error probability is achieved by the ML decision rule for these prior probabilities?
- ["Give me an A!" shouted the cheerleader...] H_e, H_g , and H_a respectively denote the hypotheses that a student is excellent, good, or average (there are no poor students). The number of grade points earned by the student in a course is a random variable $\mathcal{X}$ that takes on values 3, 6, 9, and 12 only. The Department Head, who does not know the student but has to write a letter of recommendation for him, knows that the likelihood matrix for this decision problem is as shown below:

	$\mathcal{X} = 3$	$\mathcal{X} = 6$	$\mathcal{X} = 9$	$\mathcal{X} = 12$
H_e	0.02	0.08	0.15	0.75
H_g	0.10	0.15	0.60	0.15
H_a	0.20	0.65	0.10	0.05

that is, an excellent student has 75% chance of doing well enough on the exam to get an A, 15% chance of a B, etc. The Department Head observes $\mathcal{X}$ by looking at the student's transcript, decides which of the hypotheses H_e, H_g , and H_a is true, and writes the recommendation letter accordingly.

- What is the Department Head's maximum-likelihood (ML) decision rule?
- What is the probability that an excellent student is mistakenly labeled as good? What is the probability that an excellent student is mistakenly labeled as average? What is the probability that an average student is misclassified?

- (c) If the prior probabilities are $P(H_e) = 0.2$, $P(H_g) = 0.55$, and $P(H_a) = 0.25$, what is the probability that the ML decision rule misclassifies students?
 - (d) What is the MAP decision rule corresponding to the probabilities in part (c) and what is the probability that the MAP decision rule misclassifies students?
 - (e) At the Lake Wobegon campus of the University, 95% of students are excellent and 5% are good (and thus they are all above average!) What is the MAP decision rule in this case? That is, what does the Bayesian Department Head decide about a student based on the four possible outcomes?
3. A coin is tossed repeatedly (independent trials) until a Head is observed for the first time. $\mathcal{X}$ denotes the number of trials needed to observe the first Head. The two hypotheses are
- $H_1: \mathcal{X} \sim \text{Geometric}(p_1)$
 - $H_0: \mathcal{X} \sim \text{Geometric}(p_0)$

where p_1 and p_0 are known numbers with $p_1 < p_0$.

- (a) The maximum-likelihood decision rule can be stated as
 “Decide that H_1 is the true hypothesis if $\mathcal{X} \geq \eta$ ”
 where $\geq$ is either $<$ or $>$, and η is the threshold against which $\mathcal{X}$ is compared. Determine whether $\geq$ is $<$ or $>$ and express η as a function of p_0 and p_1 .
 - (b) Let π_0 and $\pi_1 = 1 - \pi_0$ respectively denote the *a priori* probabilities of hypotheses H_0 and H_1 and assume that $0 < \pi_0 < 1$.
 For what values of π_0 (if any) does the MAP decision rule always choose hypothesis H_1 regardless of the value of the observation $\mathcal{X}$?
 For what values of π_0 (if any) does the MAP decision rule always choose hypothesis H_0 regardless of the value of the observation $\mathcal{X}$?
4. [“I’m leaving on a jet plane . . .”] Consider again Problem 6 of Problem Set 4, and now suppose that 15 of the 105 passengers who hold reservations are arriving in Chicago on a connecting flight. If the connecting flight is on time, all 15 show up for the flight to Champaign; else, none of the 15 shows up. Let $\mathcal{Y}$ denote the number of nonconnecting passengers who actually show up for the flight. Let H_0 denote the hypothesis that the connecting flight is late, and H_1 the hypothesis that the connecting flight is on time. It is reasonable to assume that the pmf of $\mathcal{Y}$ is the same regardless of which hypothesis is true, and hence we model $\mathcal{Y}$ as a binomial random variable with parameters (90, 0.9). On the other hand, $\mathcal{X}$, the total number of passengers showing up for the flight, equals $\mathcal{Y}$ if H_0 is true, while if H_1 is true, then $\mathcal{X} = 15 + \mathcal{Y}$, and thus the pmf of $\mathcal{X}$ *does* depend on which hypothesis is true.
- (a) Suppose that the gate agent observes that $\mathcal{X} = 86$. What is $P\{\mathcal{X} = 86\}$ when H_0 is the true hypothesis? What is $P\{\mathcal{X} = 86\}$ when H_1 is the true hypothesis? What is the value of the likelihood ratio when $\mathcal{X} = 86$? What is the agent’s maximum-likelihood decision as to whether the connecting flight is late?
 - (b) Repeat part (a) for the case when the gate agent observes that $\mathcal{X} = 96$.
 - (c) The gate agent knows that $\pi_0 = P\{H_0 \text{ is the true hypothesis}\} = \frac{7}{8}$. For each of the two observations considered in parts (a) and (b), what is the agent’s MAP decision as to whether the connecting flight is late?