University of Illinois **Fall 2010**

ECE 313: Problem Set 5
Bayes' Formula and binary hypothesis testing

Due:	Wednesday, September 29 at 4 p.m.
Reading:	*ECE 313 Notes* Sections 2.10 & 2.11
Reminder:	Hour Exam I on Monday October 11, 7:00 p.m. – 8:00 p.m. Sections C&D Room 141 Wohlers Hall, Sections X&E Room 114, David Kinley Hall One two-sided 8.5"×11" sheet of notes allowed, with font size no smaller than 10 pt or equivalent handwriting. Bring a picture ID. The exam will cover the reading assignments, lectures, and problems associated with problem sets 1-6.

1. [**Dissecting a vote**]
A panel of three judges has to make a yes-no decision. Each judge votes yes or no by secret ballot; each judge votes yes with some probability p, where $0 < p < 1$; the votes of the judges are mutually independent; the majority rules. Let M be the event that the decision of the panel is yes (i.e. a majority of the judges vote yes) and let A denote the event that the first judge votes yes.

   (a) Express $P(M)$ in terms of p.

   (b) Express $P(M|A)$ in terms of p.

   (c) Express $P(A|M)$ in terms of p. Sketch your answer as a function of p. What are the limits as $p \to 0$ or as $p \to 1$? Explain why these limits make sense.

2. [**Conditional distribution of half-way point**]
Consider a robot taking a random walk on the integer line. The robot starts at zero at time zero. After that, between any two consecutive integer times, the robot takes a unit length left step or right step, with each possibility having probability one half. Let F denote the event that the robot is at zero at time eight, and let X denote the location of the robot at time four.

   (a) Find $P(F)$.

   (b) Find the pmf of X.

   (c) Find $P(\{X = i\}F)$ for all integer values of i. (For what values of i is $P(\{X = i\}F) > 0$?)

   (d) Find the conditional pmf of X given that F is true. It is natural to use the notation $p_X(i|F)$ for this, and it is defined by $p_X(i|F) = P(X = i|F)$ for all integers i. Is the conditional pmf more spread out than the unconditional pmf p_X, or more concentrated?

3. [**Detection problem with the geometric distribution**]
The number of attempts, Y, required for a certain basketball player to make a 25 foot shot, is observed, in order to choose one of the following two hypotheses:

$$\begin{array}{ll} H_1 \text{ (outstanding player)}: & \text{Y has the geometric distribution with parameter } p = 0.5 \\ H_0 \text{ (average player)}: & \text{Y has the geometric distribution with parameter } p = 0.2. \end{array}$$

   (a) Describe the ML decision rule. Express it as directly in terms of Y as possible.

   (b) Find $p_{\text{false alarm}}$ and p_{miss} for the ML rule.

   (c) Describe the MAP decision rule under the assumption that H_0 is a priori twice as likely as H_1. Express the rule as directly in terms of Y as possible.

   (d) Find the average error probability, p_e, for both the ML rule and the MAP rule, using the same prior distribution given in part (c). For which rule is the average error probability smaller?

4. **[Testing hypotheses about a die]**
Consider a binary hypothesis testing problem based on observation of n independent rolls of a die. Let $X_1, \ldots, X_n$ denote the numbers rolled on the die. Let a denote a known constant that is slightly greater than one. The two hypotheses are:

   H_0 : the die is fair

   H_1 : for each roll of the die, i shows with probability $p_i = Ca^i$, where $C = \frac{1-a}{a-a^7}$, so that the probabilities sum to one.

   (a) Find simple expressions for $p_i(u_1, \ldots, u_n) = P(X_1 = u_1, \ldots, X_n = u_n|H_i)$ for $i = 0$ and $i = 1$, where $u_i \in \{1, 2, 3, 4, 5, 6\}$ for each i. Express your answers using the variables t_k, for $1 \leq k \leq 6$, where t_k is the number of the n rolls that show k. The vector $(t_1, \ldots, t_6)$ is called the type vector of $(u_1, \ldots, u_{100})$. Intuitively, the order of the observations shouldn't matter, so decision rules will naturally only depend on the type vector of the observation sequence.

   (b) Find a simple expression for the likelihood ratio, $\Lambda(u_1, \ldots, u_n) = \frac{p_1(u_1,\ldots,u_n)}{p_0(u_1,\ldots,u_n).}$ and describe, as simply as possible, the likelihood ratio test for H_1 vs. H_0 given the observations $(u_1, \ldots, u_n)$.

   (c) In particular, suppose that $n = 100$, $(t_1, \ldots, t_6) = (18, 12, 13, 19, 18, 20)$, and $a = 1.1$. Which hypothesis does the maximum likelihood decision rule select?