ECE531 Homework Assignment Number 2

Due by 8:50pm on Thursday 05-Feb-2009

Make sure your reasoning and work are clear to receive full credit for each problem.

1. 3 points. A city has two taxi companies distinguished by the color of their taxis: 85% of the taxis are Yellow and the rest are Blue. A taxi was involved in a hit and run accident and was witnessed by Mr. Green. Unfortunately, Mr. Green is mildly color blind and can only correctly identify the color 80% of the time. In the trial, Mr. Green testified that the color of the taxi was blue. Should you trust him?

   Solution: This problem can be interpreted as: Does Mr. Green's decision minimizes the Bayesian risk? Let's assume a uniform cost assignment (UCA) and use x_0 and $\mathcal{H}_0$ to represent "Yellow" and x_1 and $\mathcal{H}_1$ to represent "Blue". Also let $\pi_0 = 0.85$ denote the prior probability of a yellow taxi, $\pi_1 = 0.15$ denote the prior probability of a blue taxi, and $Y = y_1$ denote Mr. Green's observation of a blue taxi. Then from slide 26 in Lecture 2b, a deterministic Bayes decision rule can be written in terms of the posterior probabilities as

$$\delta^{B\pi}(y) = \arg\max_{i \in \{0,1\}} \sum_{x_j \in \mathcal{H}_i} \pi_j(y),$$

   where $\pi_j(y) = \text{Prob}(\text{state is } x_j \mid \text{observation is } y)$. Given Mr. Green's observation $Y = y_1$ (the taxi was blue), then if $\pi_1(y_1) > \pi_0(y_1)$ the posterior probability of the taxi being blue is greater than the posterior probability of the taxi being yellow and we can trust Mr.Green. We can write the posterior probabilities

$$\pi_1(y_1) = \frac{p_1(y_1)\pi_1}{p(y_1)}$$

$$\pi_0(y_1) = \frac{p_0(y_1)\pi_0}{p(y_1)}.$$

   Thus we only need to compare $\kappa_0 = p_0(y_1)\pi_0$ and $\kappa_1 = p_1(y_1)\pi_1$.

$$\kappa_0 = 0.2 * 0.85 = 0.17$$

$$\kappa_1 = 0.8 * 0.15 = 0.12$$

   We see that $\kappa_0 > \kappa_1$, hence the posterior probability of the taxi being yellow is greater than the posterior probability of the taxi being blue and we can not trust Mr.Green.

2. 8 points total. Consider the coin flipping problem where you have an unknown coin, either fair (HT) or double headed (HH), and you observe the outcome of n flips of this coin. Assume a uniform cost assignment. For notational consistency, let the state and hypothesis x_0 and $\mathcal{H}_0$ be the case when the coin is HT and x_1 and $\mathcal{H}_1$ be the case when the coin is HH.

(a) 3 points. Plot the conditional risk vectors (CRVs) of the deterministic decision rules for the cases of $n = 1, 2, 3, 4, \ldots$ coin flips. You might want to write a Matlab script to do this for $n > 2$.

Solution: See the Matlab script and plots below.

```matlab
%----------------------------------------------------------------
% ECE531 Spring 2009
% DRB 05-Feb-2009
% Solution to Homework 2 Problem 2 part a
%----------------------------------------------------------------
% USER PARAMETERS BELOW
%----------------------------------------------------------------
ntest = 1:4;          % values of n to test
%----------------------------------------------------------------

N = 2;             % number of hypotheses
M = 2;             % number of states
p0_H = 0.5;  % conditional probability of H given x0
p1_H = 1;    % conditional probability of H given x1
C = [0 1 ; 1 0]; % UCA

for n = ntest

    L = n+1;             % number of possible observations
    totD = M^L;       % total number of decision matrices
    B = makebinary(L,1);

    % form conditional probability matrix
    % columns are indexed by state
    % rows are indexed by observation
    P0 = zeros(L,1);
    P1 = zeros(L,1);
    for i = 0:(L-1),
        P0(i+1) = nchoosek(n,i) * p0_H^i * (1 - p0_H)^(n-i);
        P1(i+1) = nchoosek(n,i) * p1_H^i * (1 - p1_H)^(n-i);
    end
    P = [P0 P1];

    % compute CRVs for all possible deterministic decision matrices
    for i = 0:(totD-1),
        D =  [ B(:,i+1)' ; 1-B(:,i+1)' ];     % decision matrix
        % compute risk vectors
        for j=0:1,
            R(j+1,i+1) = C(:,j+1)'*D*P(:,j+1);
        end
    end

    figure
    plot(R(1,:),R(2,:),'p');
    xlabel('R0')
    ylabel('R1')
    title(['n=' num2str(n)]);
    axis square
    grid on

end
```

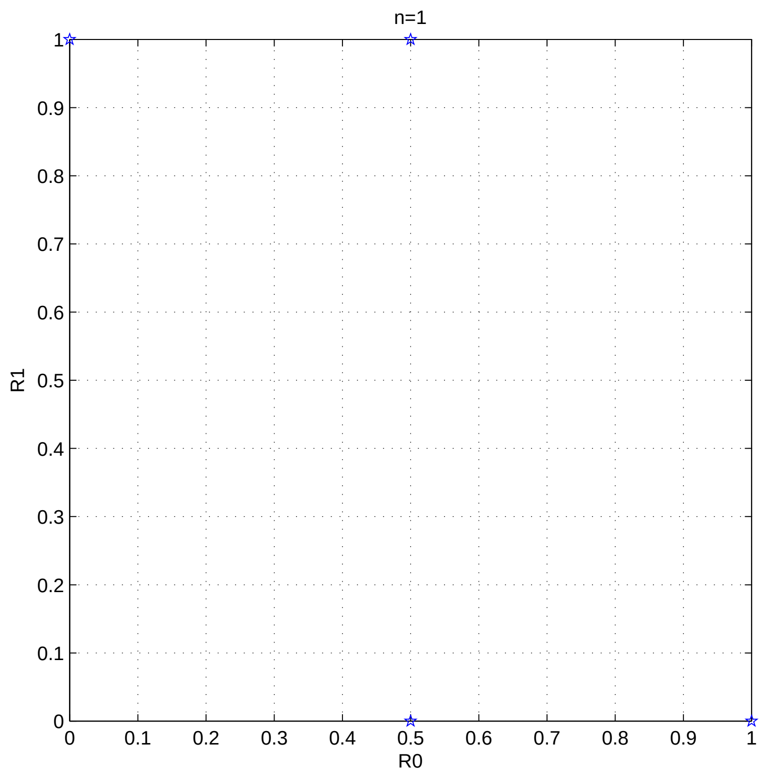
Here is the function "makebinary.m" that is called by the main code.

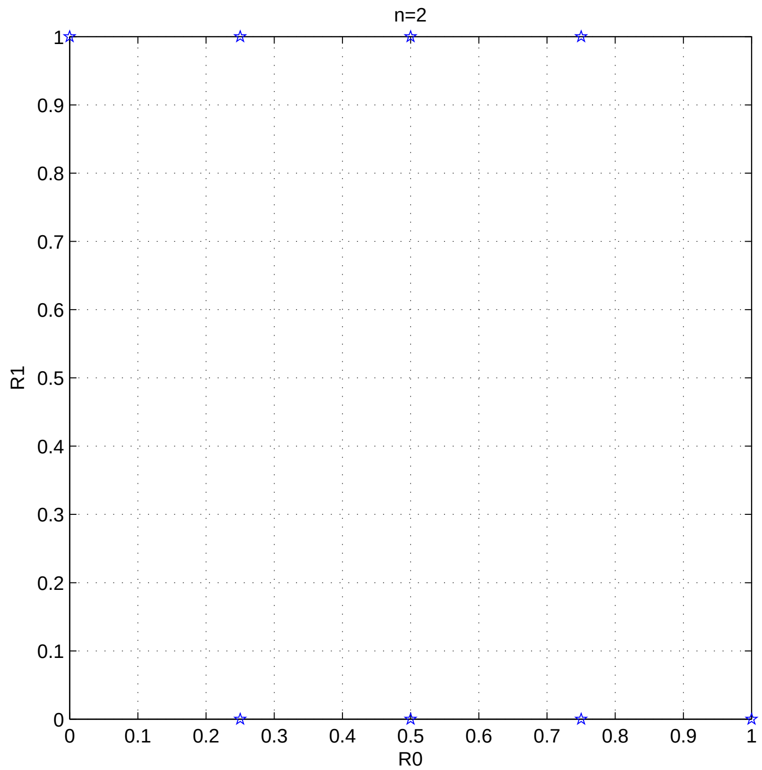
```matlab
function y=makebinary(K,unipolar)

```

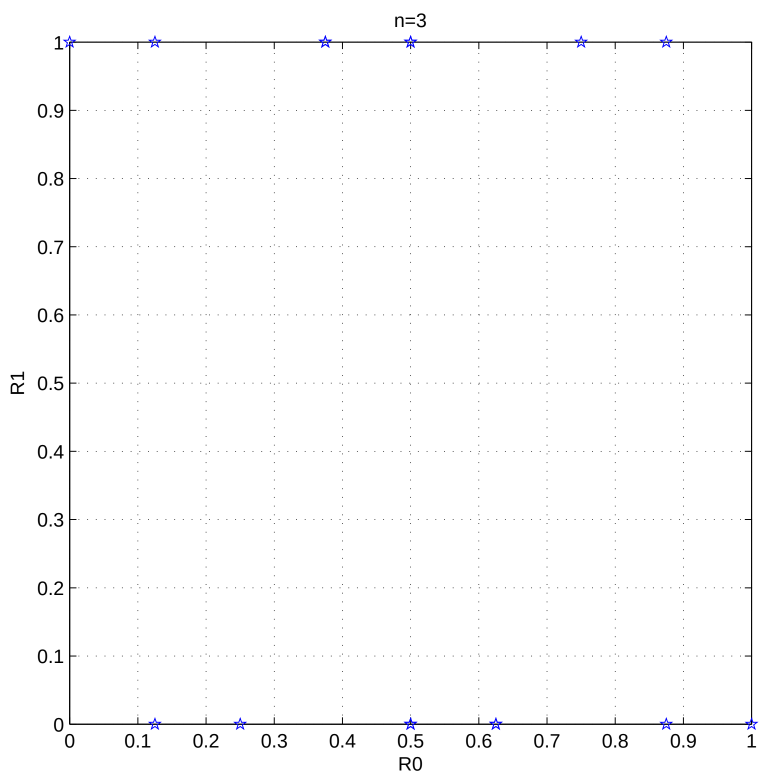
```
y=zeros(K,2^K);                    % all possible bit combos
for index=1:K,
   y(K-index+1,:)=(-1).^ceil([1:2^K]/(2^(index-1)));
end
if unipolar>0,
   y = (y+1)/2;
end
```

Here are the plots.

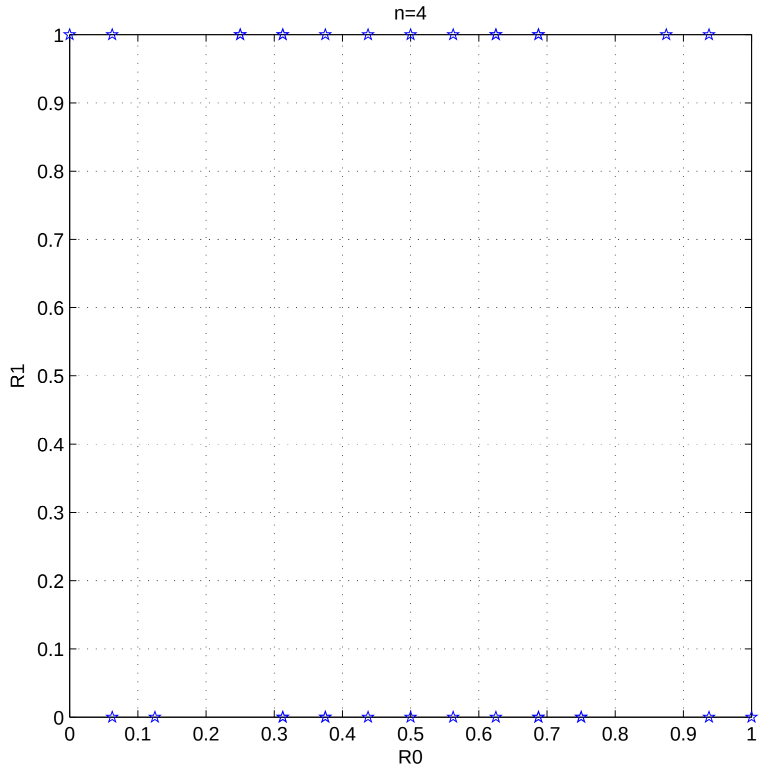




n=2
R1
R0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



n=3
R1
R0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



(b) 2 points. What can you say about the convex hull of the deterministic CRVs as n increases?

Solution: You can see the trend from the plots of how the CRVs of the deterministic decision rules move closer to the bottom left corner as n increases. The convex hull of achievable CRVs then fills more of the risk plane as n gets larger. Hence, you can achieve any arbitrarily small risk combination for sufficiently large n.

(c) 3 points. When $n = 2$, find the deterministic decision rule(s) that minimize the Bayes risk for the prior $\pi_0 = 0.6$ and $\pi_1 = 0.4$. Repeat this for the case when $n = 3$. Does the additional observation reduce the Bayes risk for this prior?

Solution: Notation: $\mathsf{HT} = x_0 \leftrightarrow H_0$ and $\mathsf{HH} = x_1 \leftrightarrow H_1$. When n=2, we can write the conditional probability matrix P as

$$P = \begin{bmatrix} 0.25 & 0 \\ 0.5 & 0 \\ 0.25 & 1 \end{bmatrix}$$

and the expected cost matrix G as

$$G = \begin{bmatrix} 0 & 0 & 0.4 \\ 0.15 & 0.3 & 0.15 \end{bmatrix}$$

Hence the optimal deterministic decision matrix that minimizes the Bayes cost is

$$D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and the resulting Bayes risk is $r = 0.15$.

When n=3, we can write the conditional probability matrix P as

$$P = \begin{bmatrix} 0.125 & 0 \\ 0.375 & 0 \\ 0.375 & 0 \\ 0.125 & 1 \end{bmatrix}$$

and the expected cost matrix G as

$$G = \begin{bmatrix} 0 & 0 & 0 & 0.4 \\ 0.075 & 0.225 & 0.225 & 0.075 \end{bmatrix}$$

Hence the optimal deterministic decision matrix that minimizes the Bayes cost is

$$D = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the resulting Bayes risk is $r = 0.075$. The risk is reduced by flipping the coin one more time, as you would expect.

3. 3 points. Poor textbook Chapter II. Problem 2 (a).

   Solution: The likelihood ratio is given by

$$L(y) = \frac{p_1(y)}{p_0(y)} = \frac{3}{2(y+1)}, \qquad 0 \le y \le 1.$$

   With uniform costs and equal priors, we can compute the optimum threshold on the likelihood function as $\tau = 1$. From our expression for $L(y)$, an equivalent condition to $L(y) \ge 1$ is $y \in [0, 0.5]$. Thus, for priors $\pi_0 = \pi_1 = 0.5$, the deterministic Bayes decision rule is given by

$$\delta^{B\pi}(y) = \begin{cases} 1 & \text{if } 0 \le y < 0.5 \\ 0/1 & \text{if } y = 0.5 \\ 0 & \text{if } 0.5 < y \le 1 \end{cases}.$$

   The corresponding minimum Bayes risk for this prior is

$$r(\delta^{B\pi}) = 0.5 \int_0^{0.5} \frac{2}{3}(y+1)dy + 0.5 \int_{0.5}^{1} dy = \frac{11}{24}.$$

4. 3 points. Poor textbook Chapter II, Problem 4 (a).

   Solution: Here the likelihood ratio is given by

$$L(y) = \frac{p_1(y)}{p_0(y)} = \sqrt{\frac{2}{\pi}} e^{y - \frac{y^2}{2}} \equiv \sqrt{\frac{2e}{\pi}} e^{-\frac{(y-1)^2}{2}}, \quad y \ge 0.$$

   Let

$$\tau = \frac{\pi_0}{1 - \pi_0}$$

   and note that $L(y) = \tau$ is equivalent to $y = \tau'$ where

$$\tau' = -2 \ln \sqrt{\frac{\pi}{2e}} \left(\frac{\pi_0}{1 - \pi_0} \right).$$

Hence, we can express the "critical region" of observations where we decide $\mathcal{H}_1$ as

$$\Gamma_1 = \left\{ y \geq 0 \,\middle|\, (y-1)^2 \leq \tau' \right\}.$$

There are three cases that depend on τ':

$$\Gamma_1 = \begin{cases} \emptyset & \text{if } \tau' < 0 \\ \left[1-\sqrt{\tau'}, 1+\sqrt{\tau'}\right] & \text{if } 0 \leq \tau' \leq 1 \\ \left[0, 1+\sqrt{\tau'}\right] & \text{if } \tau' > 1 \end{cases}.$$

For notational convenience, let's define the following constants:

$$\alpha := \frac{\sqrt{\frac{2e}{\pi}}}{1+\sqrt{\frac{2e}{\pi}}} \approx 0.56813$$

$$\beta := \frac{\sqrt{\frac{2}{\pi}}}{1+\sqrt{\frac{2}{\pi}}} \approx 0.44379$$

Then we can say that

- The condition $\tau' < 0$ is equivalent to $\alpha < \pi_0 \leq 1$.
- The condition $0 \leq \tau' \leq 1$ is equivalent to $\beta \leq \pi_0 \leq \alpha$.
- The condition $\tau' > 1$ is equivalent to $0 \leq \pi_0 < \beta$.

The minimum Bayes risk $V(\pi_0)$ can be calculated according to these three regions:

$$V(\pi_0) = \begin{cases} 1-\pi_0 & \text{if } \alpha < \pi_0 \leq 1 \\ \pi_0 \int_{1-\sqrt{\tau'}}^{1+\sqrt{\tau'}} e^{-y} dy + (1-\pi_0)\sqrt{\frac{2}{\pi}} \left[\int_0^{1-\sqrt{\tau'}} e^{-\frac{y^2}{2}} dy + \int_{1+\sqrt{\tau'}}^{\infty} e^{-\frac{y^2}{2}} dy \right] & \text{if } \beta \leq \pi_0 \leq \alpha \\ \pi_0 \int_0^{1+\sqrt{\tau'}} e^{-y} dy + (1-\pi_0)\sqrt{\frac{2}{\pi}} \int_{1+\sqrt{\tau'}}^{\infty} e^{-\frac{y^2}{2}} dy & \text{if } 0 \leq \pi_0 < \beta \end{cases}$$

Note that these integrals can be expressed as Q-functions (or erf/erfc functions) but cannot be evaluated in closed form for arbitrary priors.

5. 3 points. Poor textbook Chapter II, Problem 6 (a).

Solution: Here we have $p_0(y) = p_N(y+s)$ and $p_1(y) = p_N(y-s)$, where $p_N(x)$ is the pdf of the random variable N. This gives the likelihood function

$$L(y) = \frac{1+(y+s)^2}{1+(y-s)^2}.$$

With equal priors and uniform costs, the "critical region" where we decide $\mathcal{H}_1$ is $\Gamma_1 = \{L(y) \geq 1\} = \left\{1+(y+s)^2 \geq 1+(y-s)^2\right\} = \{2sy \geq -2sy\} = [0,\infty)$. Thus, the Bayes decision rule reduces to

$$\delta^{B\pi}(y) = \begin{cases} 1 & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases}$$

The minimum Bayes risk is then

$$r(\delta^{B\pi}) = \frac{1}{2}\int_0^{\infty} \frac{1}{\pi\left[1+(y+s)^2\right]} dy + \frac{1}{2}\int_{-\infty}^{0} \frac{1}{\pi\left[1+(y-s)^2\right]} dy = \frac{1}{2} - \frac{\tan^{-1}(s)}{\pi}.$$