ECE531 Homework Assignment Number 3

Solution

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

1. 4 points. Poor textbook Chapter II. Problem 2 (b).

   Solution: With uniform costs, the least-favorable prior will be interior to $(0,1)$, so we should use the equalizer rule. From part (a) of this problem, we know that the conditional risks are

$$R_0(\delta_{\pi_0}) = \int_0^{\tau'} \frac{2}{3}(y+1)dy = \frac{2\tau'}{3}\left(\frac{\tau'}{2}+1\right),$$

   and

$$R_1(\delta_{\pi_0}) = \int_{\tau'}^{1} dy = 1-\tau'$$

   since all Bayes decision rules decide $\mathcal{H}_1$ if $t \geq \tau'$, for some threshold $\tau' \in \mathbb{R}$, otherwise decide $\mathcal{H}_0$. Using the equalizer rule, a minimax threshold τ'_{lf} is the solution to the equation

$$\frac{2\tau'_{lf}}{3}\left(\frac{\tau'_{lf}}{2}+1\right) = 1-\tau'_{lf},$$

   which yields $\tau'_{lf} = (\sqrt{37}-5)/2$. Hence

$$\rho^{mm}(y) = \begin{cases} 1 & \text{if } 0 \leq y < (\sqrt{37}-5)/2 \\ 0/1 & \text{if } y = (\sqrt{37}-5)/2 \\ 0 & \text{if } (\sqrt{37}-5)/2 < y \leq 1. \end{cases}$$

   The minimax risk is the value of the equalized conditional risk; i.e., $R_0(\rho^{mm}) = R_1(\rho^{mm}) = 1-\tau'_{lf} = (7-\sqrt{37})/2$.

2. 4 points. Poor textbook Chapter II, Problem 4 (b).

   Solution: The minimax rule can be found by equating the conditional risks

$$R_0(\rho) = \int_{\Gamma_1} p_0(y)\,dy$$

   and

$$R_1(\rho) = \int_{\mathcal{Y}\backslash\Gamma_1} p_1(y)\,dy$$

   where $\Gamma_1 \subseteq \mathcal{Y}$ is the critical region of observations in which we decide $\mathcal{H}_1$. Recall from HW2 that there are three cases for the integration region depending on the prior. It should be clear that we can't equalize the risks in the first case. In the second case, i.e. $\beta \leq \pi_0 \leq \alpha$

corresponding to a threshold $0 \le \tau' \le 1$, the equalizer rule requires us to find the value of $\tau' \in [0,1]$ that solves

$$\begin{aligned}
\int_{1-\sqrt{\tau'}}^{1+\sqrt{\tau'}} e^{-y} dy &= \sqrt{\frac{2}{\pi}} \left[\int_0^{1-\sqrt{\tau'}} e^{-\frac{y^2}{2}} dy + \int_{1+\sqrt{\tau'}}^{\infty} e^{-\frac{y^2}{2}} dy \right] \\
\exp(-(1-\sqrt{\tau'})) - \exp(-(1+\sqrt{\tau'})) &= 2\left[0.5 - Q\left(1-\sqrt{\tau'}\right) + Q\left(1+\sqrt{\tau'}\right)\right]
\end{aligned}$$

There is no closed-form analytical solution here, but you can find a numerical solution using, for example, Matlab. Here is the Matlab command I used.

```
x = fsolve(@(x) exp(-(1-sqrt(x)))-exp(-(1+sqrt(x)))-2*(0.5-Q(1-sqrt(x))+Q(1+(sqrt(x)))),[0 1])
```

This yields a result $\tau' \approx 0.3291$. Since

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

allows us to solve for $\pi_{lf} \approx 0.5274$. Note that $\beta \le \pi_{lf} \le \alpha$, hence we can equalize the risks in the second case. We don't need to look at the third case.

The minimax decision rule is then simply the Bayes decision rule at the prior $\pi_0 = \pi_{lf}$. The minimax risk can easily be computed from either of the equal conditional risks; i.e.,

$$V(\pi_{lf}) = R_0(\rho^{mm}) = R_1(\rho^{mm}) = \exp(-(1-\sqrt{\tau'})) - \exp(-(1+\sqrt{\tau'})) \approx 0.4456.$$

3. 4 points. Poor textbook Chapter II, Problem 6 (b).

Solution: Because of the symmetry of this problem with uniform costs, we can guess that $\pi_0 = 0.5$ is the least-favorable prior. When $\pi_0 = 0.5$, the Bayes decision rule δ^B simply decides $\mathcal{H}_0$ if $y < 0$ and decides $\mathcal{H}_1$ if $y \ge 0$. To confirm that this is indeed the least favorable prior, we can check that this guess satisfies the equalizer rule:

$$R_0(\delta^B) = \int_0^{\infty} \frac{1}{\pi\left[1+(y+s)^2\right]} dy = \int_{-\infty}^{0} \frac{1}{\pi\left[1+(y-s)^2\right]} dy = R_1(\delta^B)$$

which it does. Hence the minimax decision rule is then simply the Bayes decision rule at the prior $\pi_0 = 0.5$ and the resulting risk is $r(\rho^{mm}) = \frac{1}{2} - \frac{\tan^{-1}(s)}{\pi}$ (the same answer that we had in HW2).

4. 4 points. Poor textbook Chapter II, Problem 9 (a).

Solution: Here the likelihood ratio is given by

$$L(y) = \frac{p_1(y)}{p_0(y)} = \begin{cases} \frac{2-|y|}{4(1-|y|)} & \text{if } |y| \le 1 \\ \infty & \text{if } 1 < |y| \le 2 \end{cases}$$

Hence $\{y : 1 < |y| \le 2\}$ is always included in the critical region Γ_1. The intuition is that if the true state is x_0, then we don't observe $y > 1$. Let's first focus on the observation region $|y| \le 1$. Note that when $|y| \le 1$, we can say

$$L(y) = \frac{2-|y|}{4(1-|y|)} \ge \frac{1}{2}.$$

Hence we consider the following regions of π_0.

Region (i) $0 \leq \pi_0 < \frac{1}{2}$. In this region, $\tau = \frac{\pi_0}{2(1-\pi_0)} < \frac{1}{2}$, hence $L(y)$ is always larger than the threshold τ and we always decide $\mathcal{H}_1$. Hence, we can write

$$\delta^B(y) \equiv 1, \;\; |y| \leq 1.$$

Region (ii) $1/2 \leq \pi_0 \leq 1$. In this region, $\tau = \frac{\pi_0}{2(1-\pi_0)} \geq \frac{1}{2}$. We don't always decide $\mathcal{H}_1$. A little bit of algebra leads to the decision rule

$$\delta^B(y) = \begin{cases} 1 & \text{if } \frac{4\pi_0-2}{3\pi_0-1} \leq |y| \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

So putting these results together with the prior result that we should always decide $\mathcal{H}_1$ if we get an observation $1 < |y| \leq 2$, we can write

- Region (i) $0 \leq \pi < \frac{1}{2}$ overall decision rule and risk:

$$\delta^B(y) \equiv 1, \;\; |y| \leq 2$$

$$V(\pi_0) = C_{10}\pi_0$$

- Region (ii) $1/2 \leq \pi_0 \leq 1$ overall decision rule and risk:

$$\delta^B(y) = \begin{cases} 1 & \text{if } \frac{4\pi_0-2}{3\pi_0-1} \leq |y| \leq 2 \\ 0 & \text{if } 0 \leq |y| < \frac{4\pi_0-2}{3\pi_0-1} \end{cases}$$

$$V(\pi_0) = \pi_0 R_0 + (1-\pi_0)R_1.$$

  Defining $\gamma := \frac{4\pi_0-2}{3\pi_0-1}$ for notational convenience and noting that γ is non-negative in region (ii), the conditional risk R_0 can be computed as

$$\begin{aligned} R_0 &= C_{10}\int_{\Gamma_1} p_0(y)\,dy \\ &= 2C_{10}\int_{\gamma}^{1} 1-y\,dy \\ &= C_{10}\left[1-2\gamma+\gamma^2\right] \end{aligned}$$

  Similarly, we can derive the risk R_1 as

$$\begin{aligned} R_1 &= C_{01}\int_{\mathcal{Y}\backslash\Gamma_1} p_1(y)\,dy \\ &= 2C_{01}\int_{0}^{\gamma} \frac{2-y}{4}\,dy \\ &= C_{10}\left[2\gamma - \frac{\gamma^2}{2}\right]. \end{aligned}$$

  where we have used the fact that $C_{01} = 2C_{10}$ in the last equality.

All the pieces are in place now to find the minimax decision rule. Since $V(\pi_0) = C_{10}\pi_0$ in region (i), we know that the we can equalize the risks only in region (ii). Applying the equalizer rule in region (ii), we just need to solve $1-2\gamma+\gamma^2 = 2\gamma - \frac{\gamma^2}{2}$. The positive solution to this equation is

$$\gamma = \frac{4-\sqrt{10}}{3}.$$

which yields the desired least favorable prior

$$\pi_{lf} = \frac{5+\sqrt{10}}{15} \approx 0.5442$$

which is in region (ii). The minimax decision rule is then

$$\rho^{mm}(y) = \delta^{B\pi_{lf}}(y) = \begin{cases} 1 & \text{if } \frac{4-\sqrt{10}}{3} \le |y| \le 2 \\ 0 & \text{otherwise} \end{cases}$$

and the risk can be calculated as

$$R_0 = R_1 = C_{10}(1 - 2\gamma + \gamma^2) \approx 0.5195 C_{10} = 0.2597 C_{01}.$$

5. 4 points. Poor textbook Chapter II, Problem 11.

   Solution: For the Bayes part of this problem, we are allowed to assume equally likely priors. We can write the likelihood ratio as

$$L(y) = \frac{p_1(y)}{p_0(y)} = e^{y-1/2} \ge \frac{\pi_0(C_{10} - C_{00})}{\pi_1(C_{01} - C_{11})} = \frac{1}{N}$$

   which implies a critical region Γ_1 where we decide $\mathcal{H}_1$ as

$$\Gamma_1 = \{y : y \ge \tau_N\}$$

   where $\tau_N := \frac{1}{2} + \ln \frac{1}{N}$. The Bayes decision rule is simply to decide $\mathcal{H}_0$ if $y \notin \Gamma_1$ or decide $\mathcal{H}_1$ if $y \in \Gamma_1$. The Bayes risk is of this decision rule is then

$$\begin{aligned} r(\pi_0 = 0.5) &= 0.5 \cdot R_0 + 0.5 \cdot R_1 \\ &= 0.5 \cdot \int_{\tau_N}^{\infty} p_0(y)\,dy + 0.5 \cdot N \int_{-\infty}^{\tau_N} p_1(y)\,dy \\ &= 0.5 \cdot (1 - \Phi(\tau_N)) + 0.5 \cdot N\Phi(\tau_N - 1) \end{aligned}$$

   where $\Phi(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = 1 - Q(x)$ is CDF of a zero-mean unit-variance Gaussian random variable. It should be clear that $r(\pi_0 = 0.5)$ converges to 0.5 as $N \to \infty$ because $\tau_N \to -\infty$ and $\Phi(\tau_N) \to 0$ (faster than $1/N$) as $N \to \infty$. Hence, when N gets large enough, the Bayes decision rule under equal priors will almost always decide $\mathcal{H}_1$ to avoid the huge cost of being wrong by deciding $\mathcal{H}_0$ when the true hypothesis is $\mathcal{H}_1$.

   To investigate the minimax decision rule, we need to consider *general priors*. In this case we can write the likelihood ratio as

$$L(y) = \frac{p_1(y)}{p_0(y)} = e^{y-1/2} \ge \frac{\pi_0(C_{10} - C_{00})}{\pi_1(C_{01} - C_{11})} = \frac{\pi_0}{(1-\pi_0)N}$$

   which implies a critical region

$$\Gamma_1 = \{y : y \ge \tau_N(\pi_0)\}$$

   where $\tau_N(\pi_0) := \frac{1}{2} + \ln \frac{\pi_0}{(1-\pi_0)N}$.

   For fixed N, applying the equalizer rule to the conditional risks allows us to implicitly specify a least favorable prior π_{lf} as the solution to

$$1 - \Phi(\tau_N(\pi_{lf})) = N\Phi(\tau_N(\pi_{lf}) - 1)$$

which can be rearranged as

$$\Phi(\tau_N(\pi_{lf})) + N\Phi(\tau_N(\pi_{lf}) - 1) = 1.$$

Recall that CDFs are monotonically increasing with range on $[0, 1]$. Hence, given any $N \geq 0$, we can always find $x \in \mathbb{R}$ such that $\Phi(x) + N\Phi(x - 1) = 1$. Also note that, since a CDF can't be negative,

$$\Phi(\tau_N(\pi_{lf}) - 1) \leq \frac{1}{N}.$$

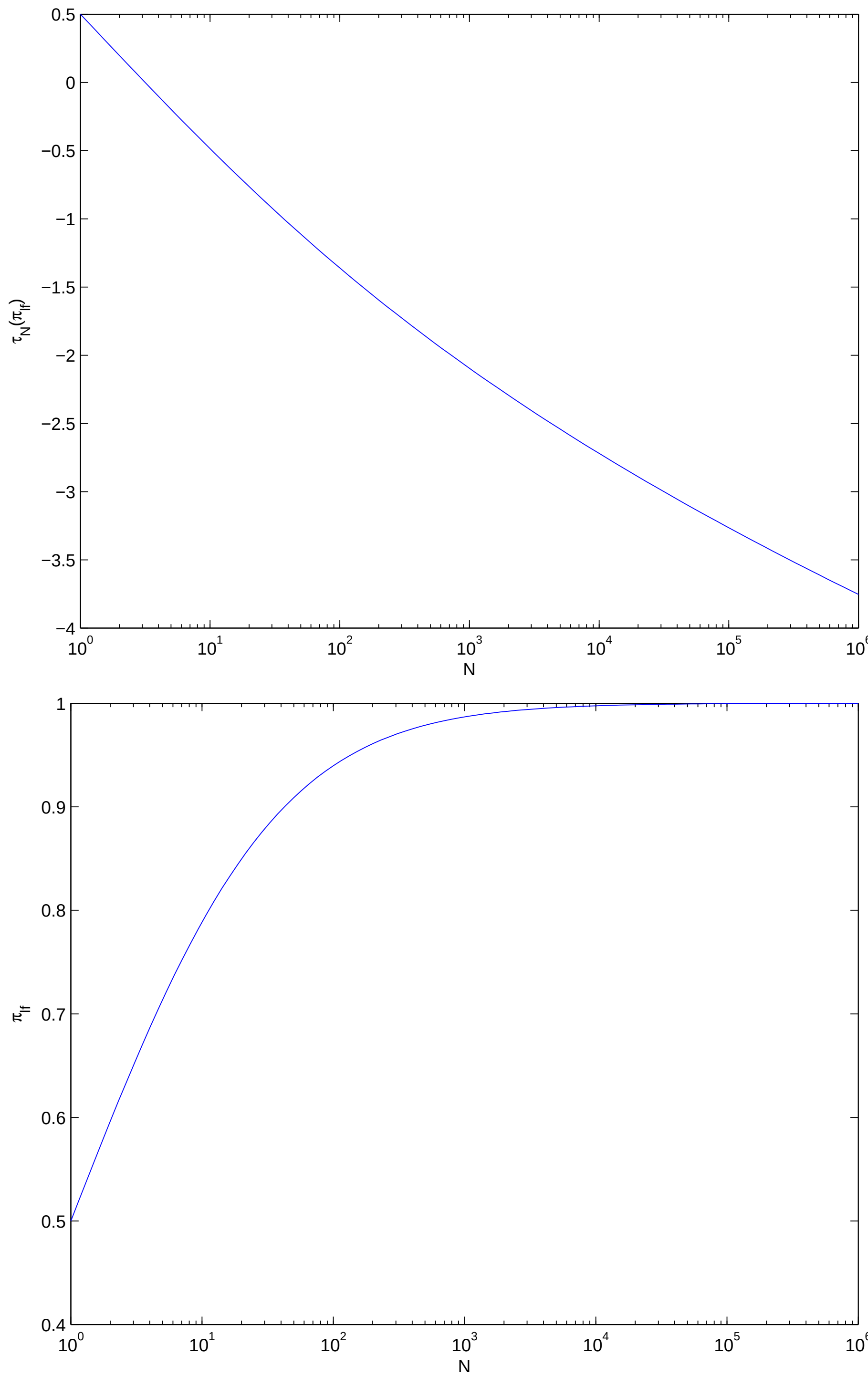
Hence, as $N \to \infty$, $\Phi(\tau_N(\pi_{lf}) - 1)$ must go to zero. This implies that $\tau_N(\pi_{lf}) \to -\infty$ as $N \to \infty$. Hence, the critical region Γ_1 of the minimax decision rule must converge to $\mathbb{R}$ as $N \to \infty$, i.e. we always decide $\mathcal{H}_1$.

What does the least favorable prior do as $N \to \infty$? This is difficult to determine analytically, but the following Matlab code numerically determines the optimum threshold and corresponding prior.

```
% ECE531 HW3 problem 5
% DRB 12-Feb-2009

% This part of the code determines the
% decision threshold as a function of N
Ntest = logspace(0,6,100);
tau = zeros(1,100);          % this is the threshold tau_N(pi_0)
fval = zeros(1,100);
exitflag = zeros(1,100);
i1=0;
for N=Ntest,
    i1=i1+1;
    [tau(i1) fval(i1) exitflag(i1)] = fsolve(@(x) Phi(x)+N*Phi(x-1)-1,0);
end
figure(1)
semilogx(Ntest,tau);
xlabel('N');
ylabel('\tau_N(\pi_{lf})')

% This part of the code determines the
% prior that corresponds to the threshold tau_N(pi_0)
X = Ntest.*exp(x-0.5);
pi0 = X./(1+X);
figure(2)
semilogx(Ntest,pi0);
xlabel('N');
ylabel('\pi_{lf}')
```

where we used the function

```
% Phi function
% Phi=1-Q;

function y=Phi(x)

x=erfc(x/sqrt(2))/2;
y = 1-x;
```



Note that the threshold $\tau_N(\pi_{lf})$ is going to $-\infty$ as we expected, but it isn't going there very fast (much slower than the rate at which N is going to infinity). This means that $\pi_{lf} \to 1$ as $N \to \infty$. Intuitively, the least favorable prior when N is large is for $\mathcal{H}_1$ to be the true hypothesis much smaller probability than $\mathcal{H}_0$. The Bayes decision is in the worst possible situation here because it wants to decide $\mathcal{H}_0$ based on the fact that this is what the priors say will occur most frequently, but the penalty for deciding $\mathcal{H}_0$ when the true state is $\mathcal{H}_1$ is very high. So the Bayes decision rule will have to guess $\mathcal{H}_1$ with enough frequency to avoid these massive penalties, and this will incur a unit cost each time $\mathcal{H}_1$ when the true state is $\mathcal{H}_0$.