S620 - Introduction To Statistical Theory - Homework 3

Enrique Areyan
February 6, 2014

[S420] Complete Exercises 2.2, 2.3, and 2.4.

(2.2) Setup:

$\Theta = \{1, 2\}$. $\theta = 1$ denote that the guard I question is a Wizard; let $\theta = 2$ the guard I question is a Muggle.

$\mathcal{X} = \{0, 1\}$, where $x = 0$ is the answer *no* to the question Are you a Wizard? and $x = 1$ is the answer *yes*.

$P_1(0) = 0, P_1(1) = 1; P_2(0) = 2/3, P_2(1) = 1/3$

$\mathcal{A} = \{1, 2\}$, where 1 = choose the guard I question and 2 = choose the guard I don't question.

$L(1, 1) = 0, L(1, 2) = 1; L(2, 1) = 1, L(2, 2) = 0$, here 1 = 1000 galleons.

1) Write down an exhaustive set of non-randomized decision rules and, by drawing the associated risk set, determine the minimax decision rule.

The following are an exhaustive set of non-randomized decision rules:

$$\begin{aligned} d_1(x) &= 1\text{, choose the guard I question always} \\ d_2(x) &= 2\text{, choose the guard I dont question always} \end{aligned}$$

$$d_3(x) = \begin{cases} 1 & \text{if } x = 0\text{, disregard the guard's response} \\ 2 & \text{if } x = 1 \end{cases}$$

$$d_4(x) = \begin{cases} 1 & \text{if } x = 1\text{, follow the guard's response} \\ 2 & \text{if } x = 0 \end{cases}$$

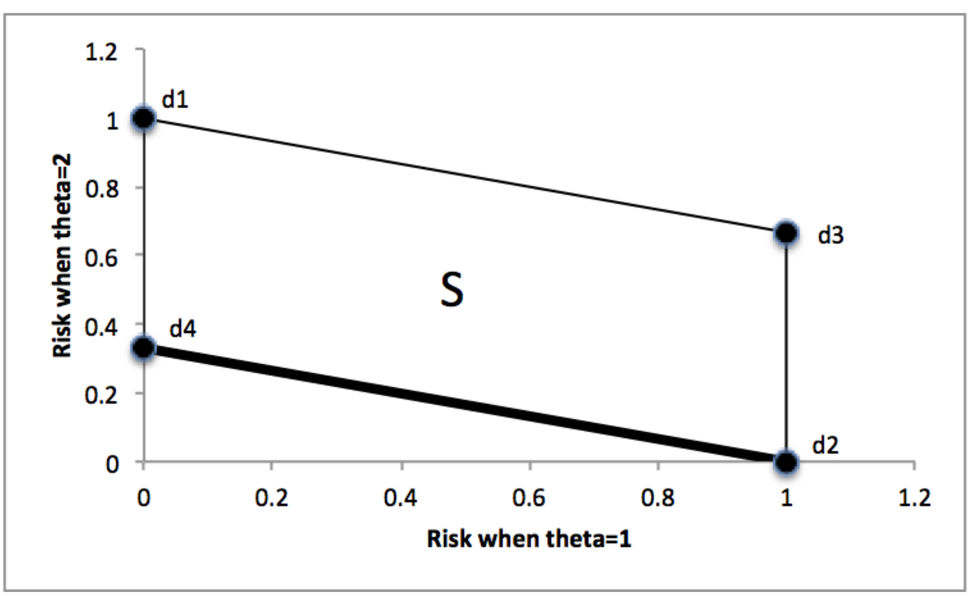
Now we can compute the risk associated with each rule (note that I will use 1 to denote a loss of 1000 galleons):

$$R(1, d_1) = E_1 L(1, d_1) = P_1(0)L(1, d_1(0)) + P_1(1)L(1, d_1(1)) = 0 \cdot 0 + 1 \cdot 0 = 0$$
$$R(2, d_1) = E_2 L(2, d_1) = P_2(0)L(2, d_1(0)) + P_2(1)L(2, d_1(1)) = \tfrac{2}{3} \cdot 1 + \tfrac{1}{3} \cdot 1 = 1$$
$$R(1, d_2) = E_1 L(1, d_2) = P_1(0)L(1, d_2(0)) + P_1(1)L(1, d_2(1)) = 0 \cdot 1 + 1 \cdot 1 = 1$$
$$R(2, d_2) = E_2 L(2, d_2) = P_2(0)L(2, d_2(0)) + P_2(1)L(2, d_2(1)) = \tfrac{2}{3} \cdot 0 + \tfrac{1}{3} \cdot 0 = 0$$
$$R(1, d_3) = E_1 L(1, d_3) = P_1(0)L(1, d_3(0)) + P_1(1)L(1, d_3(1)) = 0 \cdot 0 + 1 \cdot 1 = 1$$
$$R(2, d_3) = E_2 L(2, d_3) = P_2(0)L(2, d_3(0)) + P_2(1)L(2, d_3(1)) = \tfrac{2}{3} \cdot 1 + \tfrac{1}{3} \cdot 0 = \tfrac{2}{3}$$
$$R(1, d_4) = E_1 L(1, d_4) = P_1(0)L(1, d_4(0)) + P_1(1)L(1, d_4(1)) = 0 \cdot 1 + 1 \cdot 0 = 0$$
$$R(2, d_4) = E_2 L(2, d_4) = P_2(0)L(2, d_4(0)) + P_2(1)L(2, d_4(1)) = \tfrac{2}{3} \cdot 0 + \tfrac{1}{3} \cdot 1 = \tfrac{1}{3}$$

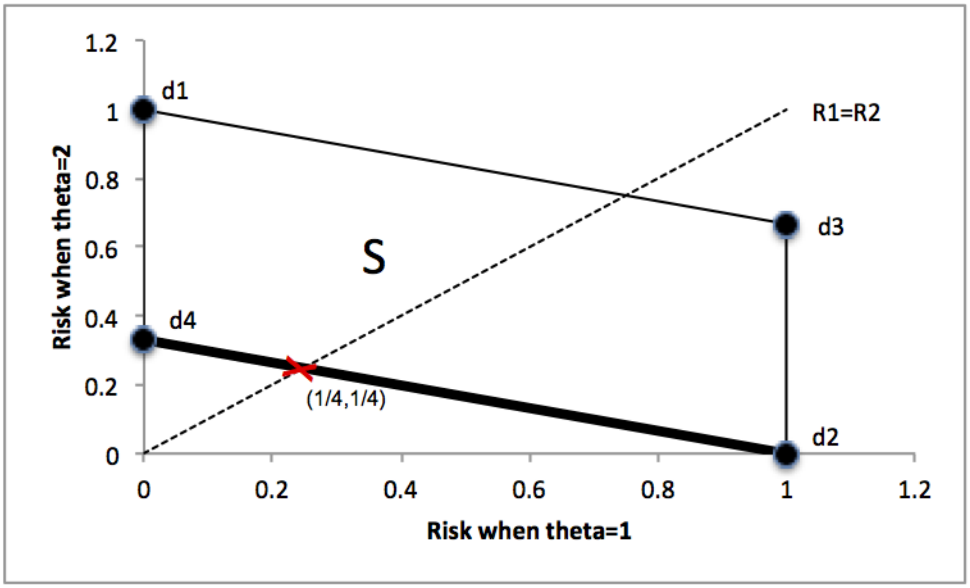
Hence, we have the points in the (R_1, R_2), i.e., (risk when $\theta = 1$, risk when $\theta = 2$)-plane:

$$d_1 = (0, 1) \quad d_2 = (1, 0) \quad d_3 = (1, \tfrac{2}{3}) \quad d_4 = (0, \tfrac{1}{3})$$

The associated risk set is:



Right away we can see that rules d_1 and d_3 are inadmissible. In fact, the collection of admissible (including randomized and nonrandomized decision rules), corresponds to the points on the lower left-hand boundary (represented by the thick line) in the previous graph. Now, we can compute the minimax rule within the collection of *nonrandomized decision rules*, i.e., $min\{max\{R(1, d_2), R(2, d_2)\}, max\{R(1, d_4), R(2, d_4\}\} = min\{max\{1, 0\}, max\{0, 1/3\}\} = min\{1, 1/3\} = 1/3$, corresponding to rule d_4. This makes intuitive sense: in absence of any other information we should follow the guard's response as a conservative strategy. However, if we were to include *randomized decision rules*, then our minimax rule will change as shown in the following graph:

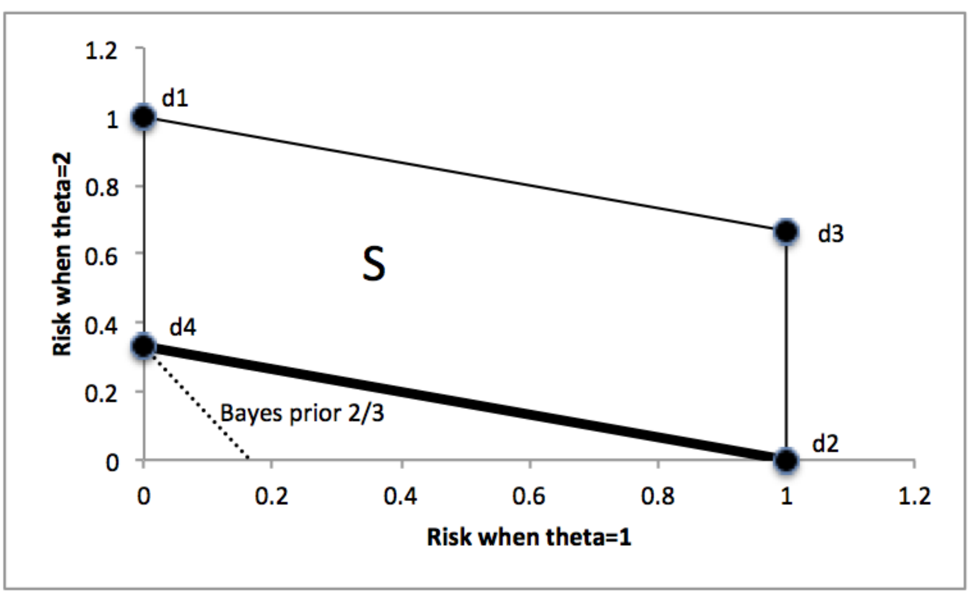


The intersection of the line $R_1 = R_2$ with the lower left-hand boundary of our risk set is at the point $(1/4, 1/4)$.

2) Let the prior probability be 2/3 that the guard being asked is indeed a Wizard. What is the Bayes decision rule?
Let us compute the Bayes risk for each rule:

$$\begin{aligned} r(\pi, d_1) &= \tfrac{2}{3}R(1, d_1) + \tfrac{1}{3}R(2, d_1) &&= 1/3 \\ r(\pi, d_2) &= \tfrac{2}{3}R(1, d_2) + \tfrac{1}{3}R(2, d_2) &&= 2/3 \\ r(\pi, d_3) &= \tfrac{2}{3}R(1, d_3) + \tfrac{1}{3}R(2, d_3) &&= 8/9 \\ r(\pi, d_4) &= \tfrac{2}{3}R(1, d_4) + \tfrac{1}{3}R(2, d_4) &&= \boxed{1/9} \Rightarrow d_4 \text{ is the Bayes rule with respect to prior } \psi = 2/3 \end{aligned}$$

Again, this makes intuitive sense. Since we have a suspicious that the guard being asked is the Wizard, it makes sense to follow his directions, i.e., apply rule d_4. We can also see this graphically by plotting the Bayes level curve that intersects the risk set S with prior 2/3, i.e., $\frac{2}{3}R_1 + \frac{1}{3}R_2 = \frac{1}{9}$ (dashed line in the following graph:)



(2.3) Setup

$\Theta = \{0, 1\}$. $\theta = 0$ there will not be snow tomorrow; $\theta = 1$ there will be snow tomorrow.

$\mathcal{X} = \{0, 1, 2\}$, where x denote the number of radio stations that forecast snow.

$P_0(0) = 1/4, P_0(1) = 2/4, P_0(2) = 1/4; P_1(0) = 1/16, P_1(1) = 6/16, P_1(2) = 9/16$

$\mathcal{A} = \{0, 1\}$, where $0 =$ don't close school and $1 =$ close school.

$L(0, 0) = 0, L(0, 1) = 1; L(1, 0) = 2, L(1, 1) = 1$

1) Write down an exhaustive set of non-randomized decision rules based on x.

The following are an exhaustive set of non-randomized decision rules:

$$d_1(x) = 0, \text{ never close school} \qquad d_2(x) = 1, \text{ always close school}$$

$$d_3(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ 1 & \text{if } x = 2 \end{cases} \qquad d_4(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ 1 & \text{if } x = 2 \end{cases}$$

$$d_5(x) = \begin{cases} 1 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ 0 & \text{if } x = 2 \end{cases} \qquad d_6(x) = \begin{cases} 0 & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ 1 & \text{if } x = 2 \end{cases}$$

$$d_7(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x = 1 \\ 0 & \text{if } x = 2 \end{cases} \qquad d_8(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ 0 & \text{if } x = 2 \end{cases}$$

2) Find the superintendent's admissible decision rules, and his minimax rule.
First, let us compute the risk associated with each rule:

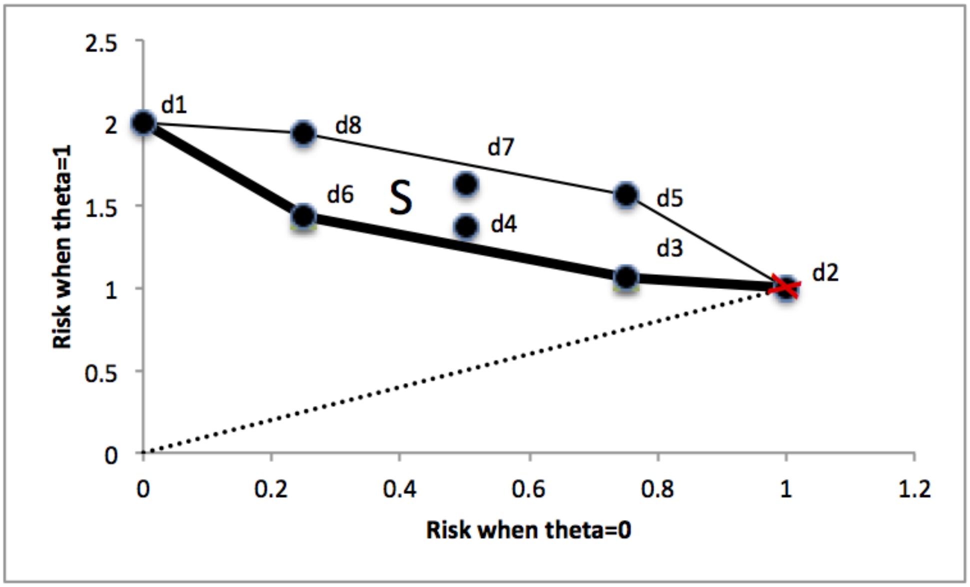
$R(0, d_1) = E_0 L(0, d_1) = P_0(0)L(0, d_1(0)) + P_0(1)L(0, d_1(1)) + P_0(2)L(0, d_1(2)) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 0 = 0$ $R(1, d_1) = E_1 L(1, d_1) = P_1(0)L(1, d_1(0)) + P_1(1)L(1, d_1(1)) + P_1(2)L(1, d_1(2)) = \frac{1}{16} \cdot 2 + \frac{3}{8} \cdot 2 + \frac{9}{16} \cdot 2 = 2$
$R(0, d_2) = E_0 L(0, d_2) = P_0(0)L(0, d_2(0)) + P_0(1)L(0, d_2(1)) + P_0(2)L(0, d_2(2)) = \frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 1 = 1$ $R(1, d_2) = E_1 L(1, d_2) = P_1(0)L(1, d_2(0)) + P_1(1)L(1, d_2(1)) + P_1(2)L(1, d_2(2)) = \frac{1}{16} \cdot 1 + \frac{3}{8} \cdot 1 + \frac{9}{16} \cdot 1 = 1$
$R(0, d_3) = E_0 L(0, d_3) = P_0(0)L(0, d_3(0)) + P_0(1)L(0, d_3(1)) + P_0(2)L(0, d_3(2)) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 1 = \frac{3}{4}$ $R(1, d_3) = E_1 L(1, d_3) = P_1(0)L(1, d_3(0)) + P_1(1)L(1, d_3(1)) + P_1(2)L(1, d_3(2)) = \frac{1}{16} \cdot 2 + \frac{3}{8} \cdot 1 + \frac{9}{16} \cdot 1 = \frac{17}{16}$
$R(0, d_4) = E_0 L(0, d_4) = P_0(0)L(0, d_4(0)) + P_0(1)L(0, d_4(1)) + P_0(2)L(0, d_4(2)) = \frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 1 = \frac{1}{2}$ $R(1, d_4) = E_1 L(1, d_4) = P_1(0)L(1, d_4(0)) + P_1(1)L(1, d_4(1)) + P_1(2)L(1, d_4(2)) = \frac{1}{16} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{9}{16} \cdot 1 = \frac{11}{8}$
$R(0, d_5) = E_0 L(0, d_5) = P_0(0)L(0, d_5(0)) + P_0(1)L(0, d_5(1)) + P_0(2)L(0, d_5(2)) = \frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 0 = \frac{3}{4}$ $R(1, d_5) = E_1 L(1, d_5) = P_1(0)L(1, d_5(0)) + P_1(1)L(1, d_5(1)) + P_1(2)L(1, d_5(2)) = \frac{1}{16} \cdot 1 + \frac{3}{8} \cdot 1 + \frac{9}{16} \cdot 2 = \frac{25}{16}$
$R(0, d_6) = E_0 L(0, d_6) = P_0(0)L(0, d_6(0)) + P_0(1)L(0, d_6(1)) + P_0(2)L(0, d_6(2)) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 1 = \frac{1}{4}$ $R(1, d_6) = E_1 L(1, d_6) = P_1(0)L(1, d_6(0)) + P_1(1)L(1, d_6(1)) + P_1(2)L(1, d_6(2)) = \frac{1}{16} \cdot 2 + \frac{3}{8} \cdot 2 + \frac{9}{16} \cdot 1 = \frac{23}{16}$
$R(0, d_7) = E_0 L(0, d_7) = P_0(0)L(0, d_7(0)) + P_0(1)L(0, d_7(1)) + P_0(2)L(0, d_7(2)) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 0 = \frac{1}{2}$ $R(1, d_7) = E_1 L(1, d_7) = P_1(0)L(1, d_7(0)) + P_1(1)L(1, d_7(1)) + P_1(2)L(1, d_7(2)) = \frac{1}{16} \cdot 2 + \frac{3}{8} \cdot 1 + \frac{9}{16} \cdot 2 = \frac{13}{8}$
$R(0, d_8) = E_0 L(0, d_8) = P_0(0)L(0, d_8(0)) + P_0(1)L(0, d_8(1)) + P_0(2)L(0, d_8(2)) = \frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot 0 = \frac{1}{4}$ $R(1, d_8) = E_1 L(1, d_8) = P_1(0)L(1, d_8(0)) + P_1(1)L(1, d_8(1)) + P_1(2)L(1, d_8(2)) = \frac{1}{16} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{9}{16} \cdot 2 = \frac{31}{16}$

Hence, we have the points in the (R_1, R_2), i.e., (risk when $\theta = 0$, risk when $\theta = 1$)-plane:

$$d_1 = (0, 2) \qquad d_2 = (1, 1) \qquad d_3 = (\tfrac{3}{4}, \tfrac{17}{16}) \qquad d_4 = (\tfrac{1}{2}, \tfrac{11}{8})$$

$$d_5 = (\tfrac{3}{4}, \tfrac{25}{16}) \qquad d_6 = (\tfrac{1}{4}, \tfrac{23}{16}) \qquad d_7 = (\tfrac{1}{2}, \tfrac{13}{8}) \qquad d_8 = (\tfrac{1}{4}, \tfrac{31}{16})$$

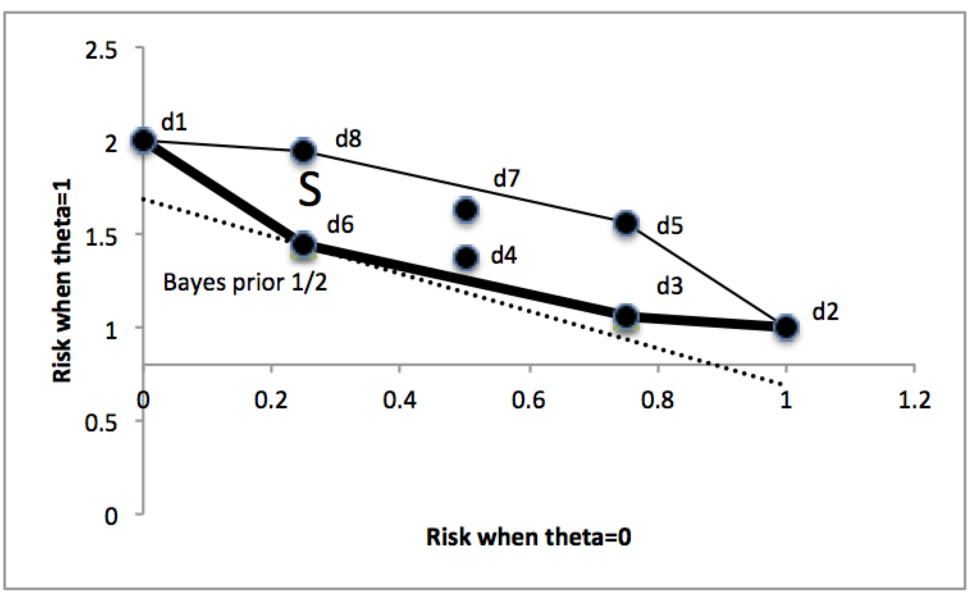
By looking at the next graph we can conclude that d_8, d_7, d_4 and d_5 are inadmissible. In fact, the collection of admissible (including randomized and nonrandomized decision rules), corresponds to the points on the lower left-hand boundary (represented by the thick line) in the next graph. The graph also shows the minimax rule (red x) if we consider the set of all rules. The minimax rule is d_2.



3) Before listening to the forecasts, he believes there will be snow with probability 1/2; find the Bayes rule with respect to this prior.
Let us compute the Bayes risk for each rule:

$$
\begin{array}{lcl}
r(\pi, d_1) &=& \frac{1}{2}R(0,d_1) + \frac{1}{2}R(1,d_1) &=& 1 \\
r(\pi, d_2) &=& \frac{1}{2}R(0,d_2) + \frac{1}{2}R(1,d_2) &=& 1 \\
r(\pi, d_3) &=& \frac{1}{2}R(0,d_3) + \frac{1}{2}R(1,d_3) &=& 29/32 \\
r(\pi, d_4) &=& \frac{1}{2}R(0,d_4) + \frac{1}{2}R(1,d_4) &=& 15/16 \\
r(\pi, d_5) &=& \frac{1}{2}R(0,d_5) + \frac{1}{2}R(1,d_5) &=& 37/32 \\
r(\pi, d_6) &=& \frac{1}{2}R(0,d_6) + \frac{1}{2}R(1,d_6) &=& \boxed{27/32} \Rightarrow d_6 \text{ is the Bayes rule with respect to prior } \psi = 1/2 \\
r(\pi, d_7) &=& \frac{1}{2}R(0,d_7) + \frac{1}{2}R(1,d_7) &=& 17/16 \\
r(\pi, d_8) &=& \frac{1}{2}R(0,d_8) + \frac{1}{2}R(1,d_8) &=& 35/32
\end{array}
$$

This result makes intuitive sense because if we believe there is equal chance of snow, then we will be better off closing the school having at least two radio station confirm that believe. Moreover, the next graph confirm that this rule is also the Bayes rule within the set of all decision rules:



(2.4) Setup:

$\Theta = \{0, 1\}$. $\theta = 0$ denote that the component is not functioning; let $\theta = 1$ otherwise.

$\mathcal{X} = \{0, 1\}$, where $x = 0$ denotes warning light off and $x = 1$ warning light on.

$P_0(0) = 1/3, P_0(1) = 2/3; P_1(0) = 3/4, P_1(1) = 1/4$

$\mathcal{A} = \{0, 1\}$, where $0 =$ don't launch $1 =$ go ahead with launch.

$L(0, 0) = 0, L(0, 1) = 10; L(1, 0) = 5, L(1, 1) = 0$. Units in billions of dollars.

1) First, let us do the same analysis as before in this case: The following are an exhaustive set of non-randomized decision rules:

$$\begin{aligned} d_1(x) &= 0, \text{ always stop launch} \\ d_2(x) &= 1, \text{ always go on with launch} \end{aligned}$$

$$\begin{aligned} d_3(x) &= \begin{cases} 0 & \text{if } x = 0, \text{ disregard the warning light} \\ 1 & \text{otherwise} \end{cases} \\ d_4(x) &= \begin{cases} 0 & \text{if } x = 1, \text{ follow the warning light} \\ 1 & \text{otherwise} \end{cases} \end{aligned}$$

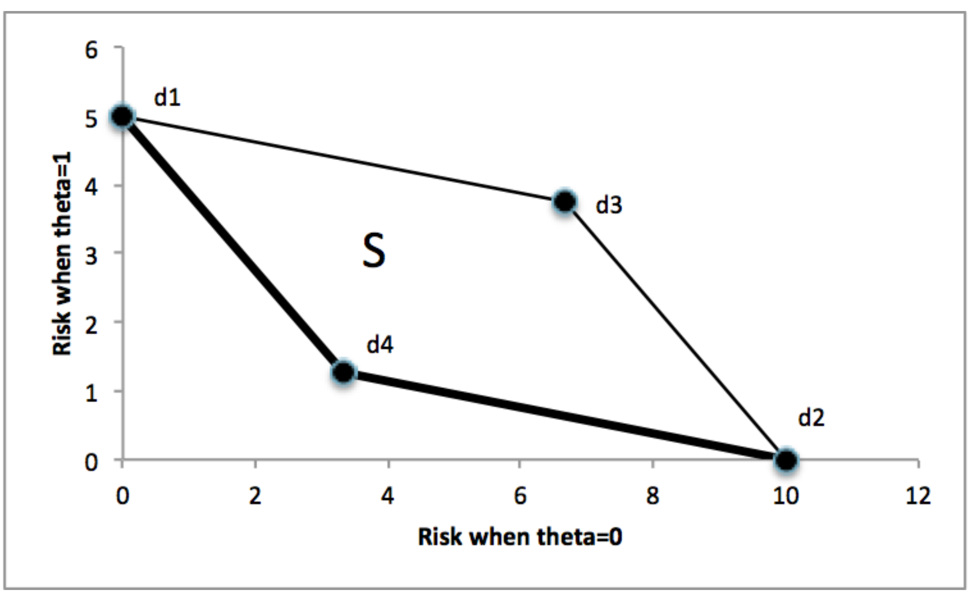
Now we can compute the risk associated with each rule:

$$R(0, d_1) = E_0 L(0, d_1) = P_0(0)L(0, d_1(0)) + P_0(1)L(0, d_1(1)) = \tfrac{1}{3} \cdot 0 + \tfrac{2}{3} \cdot 0 = 0$$
$$R(1, d_1) = E_1 L(1, d_1) = P_1(0)L(1, d_1(0)) + P_1(1)L(1, d_1(1)) = \tfrac{3}{4} \cdot 5 + \tfrac{1}{4} \cdot 5 = 5$$
$$R(0, d_2) = E_0 L(0, d_2) = P_0(0)L(0, d_2(0)) + P_0(1)L(0, d_2(1)) = \tfrac{1}{3} \cdot 10 + \tfrac{2}{3} \cdot 10 = 10$$
$$R(1, d_2) = E_1 L(1, d_2) = P_1(0)L(1, d_2(0)) + P_1(1)L(1, d_2(1)) = \tfrac{3}{4} \cdot 0 + \tfrac{1}{4} \cdot 0 = 0$$
$$R(0, d_3) = E_0 L(0, d_3) = P_0 0 L(0, d_3 0) + P_0(1)L(0, d_3(1)) = \tfrac{1}{3} \cdot 0 + \tfrac{2}{3} \cdot 10 = \tfrac{20}{3}$$
$$R(1, d_3) = E_1 L(1, d_3) = P_1(0)L(1, d_3(0)) + P_1(1)L(1, d_3(1)) = \tfrac{3}{4} \cdot 5 + \tfrac{1}{4} \cdot 0 = \tfrac{15}{4}$$
$$R(0, d_4) = E_0 L(0, d_4) = P_0(0)L(0, d_4(0)) + P_0(1)L(0, d_4(1)) = \tfrac{1}{3} \cdot 10 + \tfrac{2}{3} \cdot 0 = \tfrac{10}{3}$$
$$R(1, d_4) = E_1 L(1, d_4) = P_1(0)L(1, d_4(0)) + P_1(1)L(1, d_4(1)) = \tfrac{3}{4} \cdot 0 + \tfrac{1}{4} \cdot 5 = \tfrac{5}{4}$$

Hence, we have the points in the (R_0, R_1), i.e., (risk when $\theta = 0$, risk when $\theta = 1$)-plane: The associated risk set is:

$$d_1 = (0, 5) \quad d_2 = (10, 0) \quad d_3 = (\tfrac{20}{3}, \tfrac{15}{4}) \quad d_4 = (\tfrac{10}{3}, \tfrac{5}{4})$$

The associated risk set is:

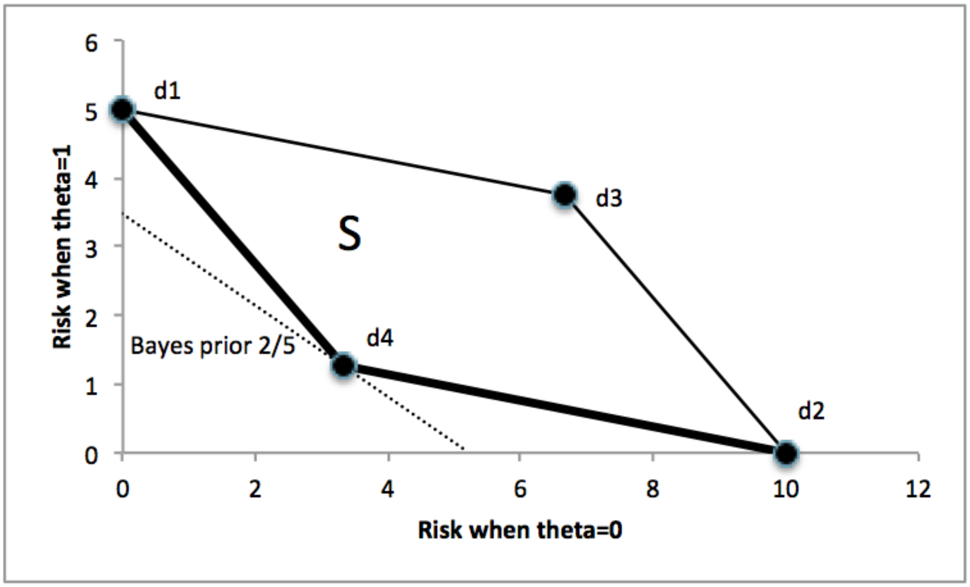


Thus, the only admissible rules are d_1, d_2, d_4, with d_4 being the minimax rule within the set of *non-randomized* decision rules. This makes sense because we would follow the warning light by using decision rule d_4. This would be the conservative approach.

2) Suppose the prior probability that the component is not functioning is $\psi = 2/5$. If the warning light does not go on, what is the decision according to the Bayes rule?
Let us compute what the Bayes rule would be:

$$\begin{aligned}
r(\pi, d_1) &= \tfrac{2}{5}R(0,d_1) + \tfrac{3}{5}R(1,d_1) &&= 3\\
r(\pi, d_2) &= \tfrac{2}{5}R(0,d_2) + \tfrac{3}{5}R(1,d_2) &&= 4\\
r(\pi, d_3) &= \tfrac{2}{5}R(0,d_3) + \tfrac{3}{5}R(1,d_3) &&= 59/12\\
r(\pi, d_4) &= \tfrac{2}{5}R(0,d_4) + \tfrac{3}{5}R(1,d_4) &&= \boxed{25/12} \text{ minimizes Bayes risk, i.e., bayes rule.}
\end{aligned}$$

Hence, the Bayes rule is d_4, and if the warning light does not go on we decide: $d_4(0) = 1$, to go on with the launch. We can plot this information (dashed line):



3) For what values of the prior probability ψ is the Bayes decision to launch the rocket, even if the warning light comes on?
Note that the only decisions rules that launch the rocket, even if the warning light comes on, are d_2, d_3. However, rule d_3 is inadmissible since it is strictly dominated by d_4. By Theorem 2.3, we known that the Bayes rule we seek must be admissible and hence, the only candidate is d_2. So our original problem reduces to finding the values of ψ for which d_2 is the Bayes rule.

Now, if $\psi = 0$, then the Bayes rule is d_2 since the Bayes level curve becomes $R1 = c$, and picking $c = 0$ the Bayes curve intersects the risk set at d_2 (through R1-axis). This will be the case up to the point where the Bayes curve coincides with the line joining d_4 and d_2. So, we want to find the value of ψ just before that happens.
We know the form of a Bayes curve: $\psi R_0 + (1-\psi)R_1 = c$. We need to solve for the curve that connects d_2 and d_4, i.e., that contains the points $d_2 = [R(0,d_2), R(1,d_2)]$ and $d_4 = [R(0,d_4), R(1,d_4)]$. We proceed:

$$\begin{aligned}
&\psi\tfrac{10}{3} + (1-\psi)\tfrac{5}{4} = c\\
&\psi 10 + (1-\psi)0 = c \quad \Rightarrow c = 10\psi
\end{aligned}$$

Replacing $c = 10\psi$ in the first equation: $\psi\frac{10}{3} + (1-\psi)\frac{5}{4} = 10\psi \Rightarrow (\frac{10}{3} - \frac{5}{4} - 10)\psi = -\frac{5}{4} \Rightarrow \psi = \frac{3}{19}$.
Hence, for $\boxed{\psi \in [0, 3/19)}$ the Bayes rule is d_2 and we choose to launch the rocket, even if the warning light comes on.

[S620]

1) Prove Theorem 2.4.: If a Bayes rule is unique, it is admissible.

Proof: (*by Contradition*). Let d_π be the unique Bayes rule with respect to the prior distribution π. Suppose that d_π is inadmissible. By definition of inadmissibility, there exists another rule $d \in \mathcal{D}$ such that $d \succ d_\pi$, i.e.,

$$R(\theta, d) \leq R(\theta, d_\pi) \text{ for every } \theta \in \Theta, \text{ and } R(\theta, d) < R(\theta, d_\pi), \text{ for at least one } \theta \in \Theta$$

Now, since π is a probability distribution on Θ we know that $\pi(\theta) \geq 0$ for every $\theta \in \Theta$. Hence,

$$R(\theta, d)\pi(\theta) \leq R(\theta, d_\pi)\pi(\theta), \text{ for every } \theta \in \Theta$$

Summing (integrating) over all states of nature $\theta \in \Theta$, we get that:

$$\int_\Theta R(\theta, d)\pi(d\theta) \leq \int_\Theta R(\theta, d_\pi)\pi(d\theta) \iff r(\pi, d) \leq r(\pi, d_\pi), \text{ by definition of Bayes risk}$$

We can analyze the last inequality by cases:

1) $r(\pi, d) < r(\pi, d_\pi) \Longrightarrow d$ has lower Bayes risk than d_π, a contradiction since d_π is Bayes.
2) $r(\pi, d) = r(\pi, d_\pi) \Longrightarrow d$ has the same risk as d_π, so d is Bayes, a contradiction since d_π is the unique Bayes rules for this π.

In any case we reach a contradiction. Therefore, d_π is admissible.

2) Exercise 2.8, part (ii):
In a Bayes decision problem, a prior distribution π is said to be *least favourable* if $r_\pi \geq r_{\pi'}$, for all prior distributions π', where r_π denotes the Bayes risk of the Bayes rule d_π with respect to π.
Suppose that π is a prior distribution, such that

$$\int R(\theta, d_\pi)\pi(\theta)d\theta = \sup_\theta R(\theta, d_\pi).$$

Show that π is least favourable.

Proof: Let π^* be an arbitrary prior distribution. Let π be the prior with the given property. Then:

$$\begin{aligned}
r(\pi^*, d_{\pi^*}) &= \int_\Theta R(\theta, d_{\pi^*})\pi^*(\theta)d\theta && \text{by definition of Bayes risk} \\
&\leq \int_\Theta R(\theta, d_\pi)\pi^*(\theta)d\theta && \text{since } d_{\pi^*} \text{ is the Bayes rule with respect to } \pi^* \\
&\leq \sup_\theta R(\theta, d_\pi) && \text{since } \pi^* \text{ is a probability distribution } (**) \\
&= \int_\Theta R(\theta, d_\pi)\pi(\theta)d\theta && \text{by hypothesis} \\
&= r(\pi, d_\pi) && \text{by definition of Bayes risk}
\end{aligned}$$

Hence, $\boxed{r(\pi^*, d_{\pi^*}) \leq r(\pi, d_\pi) \iff r_\pi \geq r_{\pi^*}}$, showing the result.

To see why $(**)$ holds, consider the following argument: we know, by definition of sup., that: $\sup_\theta R(\theta, d) \geq R(\theta, d)$ for every $\theta \in \Theta$. Now multiply by $\pi(\theta)$ each side. Since $\pi(\theta) \geq 0$ (a prob. distrib.) the inequality does not change: $\pi(\theta)\sup_\theta R(\theta, d) \geq \pi(\theta)R(\theta, d)$ now sum (integrate) over all values: $\int \pi(\theta)\sup_\theta R(\theta, d)d\theta \geq \int \pi(\theta)R(\theta, d)d\theta$. But $\sup_\theta R(\theta, d)$ is a constant on the left integral. The other term adds up to one, so we get the result: $\sup_\theta R(\theta, d) \geq \int \pi(\theta)R(\theta, d)d\theta$