Exercise 2.10

Part (a)

- Given that X can only take three different values $(0, 1, 2)$ there are $2^3 = 8$ non-randomized decision rules $(d_1, d_2, \ldots, d_8)$ (as in the notes) that can exists.
- The convex hull is the smallest convex set containing the risk points
$$\begin{pmatrix} R(\theta_1, d_i) \\ R(\theta_2, d_i) \end{pmatrix}$$
for $i = 1, 2, \ldots, 8$ and this is illustrated in the graph on page 5.
- To illustrate the calculations of the (x, y) coordinates:
$$\begin{aligned} R(\theta_1, d_1) &= L(\theta_1, a_1) \cdot 0.81 + L(\theta_1, a_1) \cdot 0.18 + L(\theta_1, a_1) \cdot 0.01 \\ &= 0 \cdot 0.81 + 0 \cdot 0.18 + 0 \cdot 0.01 \\ &= 0 \end{aligned}$$
$$\begin{aligned} R(\theta_2, d_1) &= L(\theta_2, a_1) \cdot 0.25 + L(\theta_2, a_1) \cdot 0.50 + L(\theta_2, a_1) \cdot 0.25 \\ &= 3 \cdot 0.25 + 3 \cdot 0.5 + 3 \cdot 0.25 \\ &= 3 \end{aligned}$$
The risk point that corresponds to d_1 is with coordinates $(0, 3)$. We can do the same for d_2:
$$\begin{aligned} R(\theta_1, d_2) &= L(\theta_1, a_1) \cdot 0.81 + L(\theta_1, a_1) \cdot 0.18 + L(\theta_1, a_2) \cdot 0.01 \\ &= 0 + 0 + 1 \cdot 0.01 \\ &= 0.01 \end{aligned}$$
$$\begin{aligned} R(\theta_2, d_1) &= L(\theta_2, a_1) \cdot 0.25 + L(\theta_2, a_1) \cdot 0.50 + L(\theta_2, a_2) \cdot 0.25 \\ &= 3 \cdot 0.25 + 3 \cdot 0.5 + 0 \cdot 0.25 \\ &= 2.25 \end{aligned}$$
Make sure that you can work out the remaining risk points and can reproduce the following table:

Rule	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
$R(\theta_1, d_i)$	0	0.01	0.18	0.19	0.81	0.82	0.99	1
$R(\theta_2, d_i)$	3	2.25	1.5	0.75	2.25	1.5	0.75	0

- For a prior $(p_1, p_2) = (p_1, 1 - p_1)$ on (θ_1, θ_2), the risk points that have the same value b of their Bayes risk are on the line with equation
$$p_1 x + p_2 y = b \qquad \text{or equivalently} \qquad p_1 x + (1 - p_1) y = b$$
Note that: if $x = y$ we get $x = y = b$ and hence the x (and equivalently y) coordinate of the intersection of the line $p_1 x + p_2 y = b$ with the line $x = y$ also represents the value of this risk.

Part (b)

- The minimax rule in the set $\mathcal{D}$ of randomize decision rules is obtained by examining the intersection of the line $y = x$ with the "most south-west" part of the convex risk set. Therefore, we need to solve the system

$$\begin{cases} y = x \\ y - 0 = \dfrac{0.75 - 0}{0.19 - 1}(x - 1) \end{cases}$$

which leads to

$$x = -\frac{75}{81}(x - 1)$$

and gives the solution $x = y = \frac{25}{52}$.

- Therefore the point $(25/52, 25/52)$ is the risk point that corresponds to the minimax decision rule δ^* in the set of all randomized decision rules $\mathcal{D}$ that are generated by the set $D = \{d_1, d_2, \ldots, d_8\}$ of the non-randomized decision rules.

- The value of the minimax risk is given by

$$\sup_{\theta \in \Theta} R(\theta, \delta^*) = \sup \left\{ \frac{25}{52}, \frac{25}{52} \right\} = \frac{25}{52}$$

- Important: if we were only looking for the minimax decision rule in D (not $\mathcal{D}$) then the answer will be different. It is given as follows: Now the minimum of these

Rule	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8
$R(\theta_1, d_i)$	0	0.01	0.18	0.19	0.81	0.82	0.99	1
$R(\theta_2, d_i)$	3	2.25	1.5	0.75	2.25	1.5	0.75	0
Max	3	2.25	1.5	**0.75**	2.25	1.5	0.99	1

maxima is 0.75. Hence, d_4 is the minimax decision rule with minimax risk of 0.75. This risk is naturally greater than the risk of the minimax decision rule in $\mathcal{D}$ of $25/52$.

- To represent the minimax rule δ^* in the set $\mathcal{D}$ as a randomization of the rules d_4 and d_8, we need to find $\alpha \in [0, 1]$ to say that δ^* choose d_4 with probability α and d_8 with probability $(1 - \alpha)$. Therefore,

$$\alpha \cdot 0.19 + (1 - \alpha) \cdot 1 = \frac{25}{52}$$

must hold. Solving this gives $\alpha = 0.641$ and we can claim that

$$\delta^* = \begin{cases} \text{choose } d_4 \text{ with probability } 0.641 \\ \text{choose } d_8 \text{ with probability } 0.359 \end{cases}$$

Part (c)

- For the least favourable prior, we need to maximize the Bayes risk when we start manipulating the priors $(p, 1-p)$.

- Since for any such prior the value of the Bayes risk will be geometrically represented as an x (or equivalently y) coordinate on the line that connects $(0,0)$ with $(25/52, 25/52)$, we maximize when we end up with $(25/52, 25/52)$. That is, the minimax solution.

- We are looking for a prior in the form $(p, 1-p)$ for which $(25/52, 25/52)$ would be the Bayes solution. This implies that $(p, 1-p)$ should be perpendicular ($\perp$) to the line $\overline{d_4 d_8}$. This requirement is the same as specifying the slope of the line $\overline{d_4 d_8}$:

$$\frac{0.75-0}{0.19-1} = -\frac{25}{27}$$

  to be equivalent to

$$-\frac{p}{1-p}$$

  since the line $px + (1-p)y = c_1$ (where c_1 is a constant) has slope $-p/(1-p)$, in fact, $y = -(p/(1-p))x + c_2$ (where c_2 is another constant).

- Hence we need to find p such that

$$-\frac{p}{1-p} = -\frac{25}{27}$$

  which leads to $p = 25/52$ and the least favourable prior is $(25/52, 27/52)$.

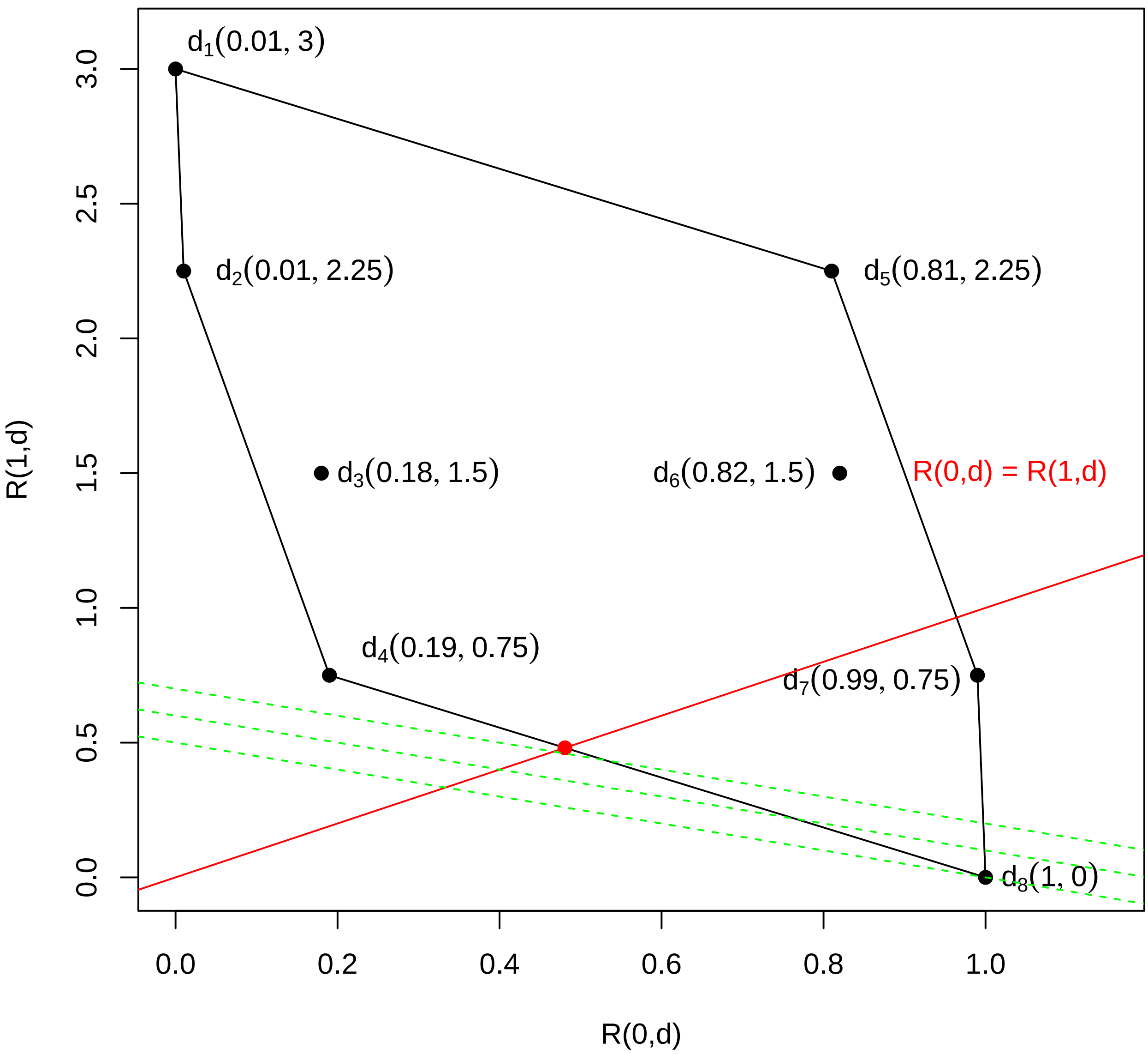
Part (d)

- The line
$$\frac{1}{3}x + \frac{2}{3}y = c_3$$
where c_3 is a constant, represents points (x, y) in the risk set with equivalent values of their Bayes risk. The slope of the lines is
$$\frac{-1/3}{2/3} = -\frac{1}{2}.$$

- Hence, to find the Bayes rule with respect to this prior, we need to move lines with a slope (-1/2) "most south-west" while still maintaining an intersection with the risk set.

- By doing so, you will see geometrically that you will end up with the rule d_8. Therefore, d_8 (1,0) represents the risk pont that corresponds to the Bayesian decision rule with respect to the prior (1/3,2/3) on (θ_1, θ_2). That is, d_8 is the Bayesian decision rule with respect to the prior (1/3,2/3).

- Its Bayes risk is equal to
$$\frac{1}{3}R(\theta_1, d_8) + \frac{2}{3}R(\theta_2, d_8) = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 0 = \frac{1}{3}.$$


d1(0.01, 3)
d2(0.01, 2.25)
d3(0.18, 1.5)
d4(0.19, 0.75)
d5(0.81, 2.25)
d6(0.82, 1.5)
d7(0.99, 0.75)
d8(1, 0)
R(0,d) = R(1,d)
R(1,d)
R(0,d)
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0