Solutions to the Exercises of Section 2.10.

2.10.1. (a) $r(\pi,(x,y)) = \pi R(1/3,(x,y)) + (1-\pi)R(2/3,(x,y))$,

$R(1/3,(x,y)) = \frac{1}{3}(\frac{1}{3}-x)^2 + \frac{2}{3}(\frac{1}{3}-y)^2$ and $R(2/3,(x,y)) = \frac{2}{3}(\frac{2}{3}-x)^2 + \frac{1}{3}(\frac{2}{3}-y)^2$.

To find the value of (x,y) that minimizes $r(\pi,(x,y))$, we solve $\partial r/\partial x = 0$ and $\partial r/\partial y = 0$, and find that r is minimized by

$$(x(\pi), y(\pi)) = (\frac{4-3\pi}{3(2-\pi)}, \frac{2}{3(1+\pi)})$$

which gives the Bayes rule with respect to π.

(b) The set of all Bayes rules is the parametric curve traced out by $(x(\pi), y(\pi))$ as π goes from 0 to 1 inclusive. This is the arc of the linear fractional curve, $y = 2(1-x)/(7-9x)$ from $(1/3,1/3)$ to $(2/3,2/3)$.

(c) To show the set of non-randomized Bayes rules is essentially complete, we may use Theorem 2.10.3. However, we cannot take C in that theorem to be the set of all non-randomized rules, since that set is the whole plane, R^2, which is not compact. Instead, we take $C = [1/3,2/3] \times [1/3,2/3]$ which is compact and argue that C is essentially complete: First, the set R^2 of all non-randomized rules is essentially complete since the loss is convex in (x,y) for all $\theta \in \Theta$. Next, any rule, (x,y), with $x < 1/3$ can be strictly improved by putting $x = 1/3$ (see the formula for R), and any rule with $x > 2/3$ can be improved by putting $x = 2/3$. Similarly for $y < 1/3$ and $y > 2/3$.

Now, since C is essentially complete and compact and since $R(\theta,(x,y))$ is continuous in (x,y) for all θ, the extended Bayes rules in C forms an essentially complete class, and since Θ is finite, the extended Bayes rules are the Bayes rules.

To show minimality, it is sufficient to note that each of these Bayes rules is admissible. This follows from Theorem 2.3.1 since each of these rules is a unique Bayes rule for the given prior.

2.10.2. (a) The set D of non-randomized decision rules may be represented by the unit square with $(x,y) \in D$ representing the rule that estimates θ to be x if $X = 1$ and estimates θ to be y if $X = 0$. The risk function is $R(\theta,(x,y)) = \theta(\theta-x)^2 + (1-\theta)(y-\theta)^2 = \theta^2[1-2x+2y] + \theta[x^2-2y-y^2] + y^2$. Hence, the Bayes risk with respect to a prior, τ, is

$$r(\tau,(x,y)) = (\mathrm{E}\theta^2)[1-2x+2y] + (\mathrm{E}\theta)[x^2-2y-y^2] + y^2.$$

Setting the derivatives of r with respect to x and y to zero yields the Bayes rules as follows. If $0 < \mathrm{E}\theta < 1$, then $x = \mathrm{E}\theta^2/\mathrm{E}\theta$ and $y = (\mathrm{E}\theta - \mathrm{E}\theta^2)/(1-\mathrm{E}\theta)$ is unique Bayes. If $\mathrm{E}\theta = 1$, then $x = 1$ and any y is Bayes. If $\mathrm{E}\theta = 0$, then $y = 0$ and any x is Bayes.

(b) For a given $\mathrm{E}\theta$, the value of $\mathrm{E}\theta^2$ ranges from $(\mathrm{E}\theta)^2$ to $\mathrm{E}\theta$, inclusive. Under the map, $(\mathrm{E}\theta, \mathrm{E}\theta^2) \to (x,y)$ provided by the Bayes rules found above, the set $T = \{(\mathrm{E}\theta, \mathrm{E}\theta^2) : 0 \le (\mathrm{E}\theta)^2 \le \mathrm{E}\theta^2 \le \mathrm{E}\theta \le 1\}$ maps onto the set $B = \{(x,y) : 0 \le y \le x \le 1\}$. This is the set of Bayes rules.

(c) D is essentially complete since L is convex. To show that B is essentially complete, we use first use Theorem 2.10.3 with $C = D$ (since D is compact and R is continuous on D) to conclude that the extended Bayes rules in D form a complete class. But since the Bayes risk is continuous on T and T is compact, the set of extended Bayes rules is B, the set of Bayes rules.

2.10.3. The parenthetical remark is false. B^* is not the class of Bayes rules. We may only conclude that B^* contains the class of Bayes rules. A stronger statement, proved below, is as follows. For a prior distribution π, let B_π denote the class of nonrandomized Bayes rules with respect to π, and let B^*_π be the class of randomized rules that give all their mass to B_π. Then for a decision problem with finite Θ and a risk set that is bounded from below and closed from below, the class $\bigcup_{\pi \in D^*} B^*_\pi$ of elements of D^* is a complete class of decision rules.

From Theorem 2, the class of Bayes rules is complete. We show that the class, $\mathcal{B}$, of Bayes rules is equal to $\bigcup_{\pi \in D^*} B^*_\pi$. Clearly, any probability mixture of elements of B_π is still Bayes with respect to π, so $\bigcup B^*_\pi \subset \mathcal{B}$. We must now show that $\mathcal{B} \subset \bigcup B^*_\pi$. For this we use the argument found in the second paragraph of Section 1.8. Suppose $\delta \in \mathcal{B}$ and find π such that δ is Bayes with respect to π. Let Z denote the random variable with values in D whose distribution is given by δ. Then $r(\pi,\delta) = \mathrm{E}\, r(\pi, Z)$, but because δ is Bayes with respect to π, we have $r(\pi,\delta) \le r(\pi,d)$ for all $d \in D$. This entails $r(\pi,Z) = r(\pi,\delta)$ with probability one, so that Z gives all of its weight to nonrandomized Bayes rules with respect to π. Hence, $\delta \in B^*_\pi$, completing the proof.

2.10.4. (a) The sample space consists of $n+1$ points, $\mathcal{X} = \{0, 1, \ldots, n\}$, and the action space of two, $\mathcal{A} = \{a_1, a_2\}$, so, using p_x to represent $\mathrm{P}(a_1|X = x)$, the space of behavioral rules may be written as

$$\mathcal{D} = \{\mathbf{p} = (p_0, p_1, \ldots, p_n) : 0 \le p_x \le 1 \quad \text{for all x}\}.$$

(b) The risk function of the behavioral rule, $\mathbf{p}$, is

$$\hat{R}(\theta_2, \mathbf{p}) = \sum_{x=0}^{n} \binom{n}{x} \theta_2^x (1-\theta_2)^{n-x} p_x \quad \text{and} \quad \hat{R}(\theta_1, \mathbf{p}) = \sum_{x=0}^{n} \binom{n}{x} \theta_1^x (1-\theta_1)^{n-x} (1-p_x).$$

(c) The Bayes risk of $\mathbf{p}$ with respect to $\underline{\pi} = (\pi, 1-\pi)$ is

$$r(\underline{\pi}, \mathbf{p}) = \sum_{x=0}^{n} \binom{n}{x} \left[\pi \theta_2^x (1-\theta_2)^{n-x} p_x + (1-\pi)\theta_1^x (1-\theta_1)^{n-x} (1-p_x)\right].$$

The behavioral rule, $\mathbf{p}$, is a Bayes rule with respect to $\underline{\pi}$ if $p_x = 0$ when the coefficient of p_x in this expression is positive, and $p_x = 1$ when it is negative. The Bayes rules with respect to $\underline{\pi}$ are those $\mathbf{P}$ such that

$$p_x = \begin{cases} 0 & \text{if } \pi\theta_2^x(1-\theta_2)^{n-x} > (1-\pi)\theta_1^x(1-\theta_1)^{n-x} \\ \text{any} & \text{if } \pi\theta_2^x(1-\theta_2)^{n-x} = (1-\pi)\theta_1^x(1-\theta_1)^{n-x} \\ 1 & \text{if } \pi\theta_2^x(1-\theta_2)^{n-x} < (1-\pi)\theta_1^x(1-\theta_1)^{n-x} \end{cases} = \begin{cases} 0 & \text{if } \rho^x < c\pi/(1-\pi) \\ \text{any} & \text{if } \rho^x = c\pi/(1-\pi) \\ 1 & \text{if } \rho^x > c\pi/(1-\pi) \end{cases}$$

where $\rho = \theta_1(1-\theta_2)/[\theta_2(1-\theta_1)] > 1$ and $c = [(1-\theta_2)/(1-\theta_1)]^n$. The class of all Bayes rules is therefore

$$\mathcal{B} = \{\mathbf{p} : \text{for some } j,\ p_i = 0 \text{ for } i < j,\ 0 \le p_j \le 1, \text{ and } p_i = 1 \text{ for } i > j\}.$$

This same class of rules is the class of Bayes rules for this problem for any θ_1 and θ_2, provided $0 < \theta_2 < \theta_1$.

(d) Yes. The parameter space, Θ, is finite, the set of nonrandomized decision rules, D, is finite, so the nonradomiized risk set is finite, and hence the risk set is compact. The result follows from Theorem 2.

2.10.5. Let $\mathcal{B}$ denote the class of Bayes rules (resp. extended Bayes rules), and assume that $\mathcal{B}$ is essentially complete. We are to show that $\mathcal{B}$ is complete.

Let δ be a rule not in $\mathcal{B}$. We must show there exists a $\delta_0 \in \mathcal{B}$ such that δ_0 is better than δ. Using essential completeness of $\mathcal{B}$, find $\delta_0 \in \mathcal{B}$ such that δ_0 is as good as δ. Suppose, contrary to what we are to show, that δ_0 is not better than δ. Then δ is equivalent to δ_0; that is, $R(\delta, \theta) = R(\delta_0, \theta)$ for all $\theta \in \Theta$. But this implies that δ is a Bayes rule (resp. extended Bayes rule), contradicting the assumption that δ is not in $\mathcal{B}$. Therefore δ_0 is better than δ.