Solutions to Exercises on Decision Theory

(1a) We have $\Psi(\Theta) = \{0,1\}$ and

$$\int_{\{0,1\}} \text{Loss}(\theta,\psi)\,\delta(x,d\psi) = \text{Loss}(\theta,0)\,\delta(x,\{0\}) + \text{Loss}(\theta,1)\,\delta(x,\{1\})$$

	$\theta = a$ ($\Psi(\theta)=1$)	$\theta = b$ ($\Psi(\theta)=0$)	$\theta = c$ ($\Psi(\theta)=0$)
$x=1$	0	1	1
$x=2$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$x=3$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

$$R(\theta,\delta) = \int_{\{1,2,3\}} \int_{\{0,1\}} \text{Loss}(\theta,\psi)\,\delta(x,d\psi)\,P_\theta(dx)$$

$\theta = a$: $= \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{3} = \frac{21}{54} = \frac{7}{18}$

$\theta = b$: $= \frac{1}{3}\cdot 1 + \frac{2}{3}\cdot\frac{1}{2} + 0\cdot\frac{1}{3} = \frac{2}{3}$

$\theta = c$: $= 0\cdot 1 + \frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{3} = \frac{21}{54} = \frac{7}{18}$

(b) The MLE of θ is $\theta(1) = a$, $\theta(2) = b$, $\theta(3) = c$

The profile MLE of Ψ is $d = \Psi(\theta(1)) = 1$, $\Psi(\theta(2)) = 0$, $\Psi(\theta(3)) = 0$ or $\Psi(\theta(1)) = 0$, $\Psi(\theta(2)) = 0$, $\Psi(\theta(3)) = 0$ so d can be thought of as a profile MLE as it is not unique. Using the first MLE,

$$R(\theta,d) = \int_{\{1,2,3\}} L(\theta,d(x))\,P_\theta(dx)$$

$\theta = a$: $= P_a(d(x)=0) = P_a(\{2,3\}) = 2/3$

$\theta = b$: $= P_b(d(x)=1) = P_b(\{1\}) = 1/3$

$\theta = c$: $= P_c(d(x)=1) = P_c(\{1\}) = 0$

(c) $R(a,\delta) < R(a,d)$ but $R(b,\delta) > R(b,d)$
so neither is preferred to the other.

(d) We have that when x is known the likelihood is known and conversely when the likelihood is known then x is known. Therefore x is a mss. Note that the conditional distribution of x given x is degenerate at x so $\delta_T = \delta$ and $d_T = d$ and there is no randomization.

(e) To be unbiased d has to satisfy, for every $\theta, \theta' \in \Theta$

$$R(\theta, d) \le \int_{\{1,2,3\}} \text{Loss}(\theta'; d(x)) P_\theta(dx)$$

$$\overset{\theta=a,\ \theta'=b,\ \Psi(\theta)=0}{=} \frac{1}{3}1 + \frac{1}{3}0 + \frac{1}{3}0 = \frac{1}{3}$$

and note $R(a,d) = 2/3 > \frac{1}{3}$ and so d is not unbiased.

3

(2) (a) $R(a,d) = \frac{1}{4}L(a,a) + \frac{1}{4}L(a,a) + 0 \cdot L(a,a) + \frac{1}{2}L(a,b)$

$= \frac{1}{2}$

$R(b,d) = \frac{1}{2}L(b,a) + 0L(b,a) + \frac{1}{4}L(b,a) + \frac{1}{4}L(b,b)$

$= \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$

(b) Consider the dec. fn given by $d_0(1) = b$, $d_0(2) = a$, $d_0(3) = b$, $d_0(4) = a$. Then

$R(a,d_0) = \frac{1}{4}L(a,b) + \frac{1}{4}L(a,a) + 0L(a,b) + \frac{1}{2}L(a,a)$

$= \frac{1}{4}$

$R(b,d_0) = \frac{1}{2}L(b,b) + 0L(b,a) + \frac{1}{4}L(b,b) + \frac{1}{4}L(b,a)$

$= \frac{1}{4}$

Therefore $R(\theta,d_0) < R(\theta,d) \quad \forall \theta$ and d is ~~not~~ admissible.

(c) $r(d) = \frac{1}{4} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{3}{4} = \frac{1}{8} + \frac{9}{16} = \frac{11}{16}$

(d) We first need to determine the posterior distribution of θ for each data value.

Posterior distribution

$x=1 \quad \theta=a \quad \frac{1}{4}\cdot\frac{1}{4} \Big/ \left(\frac{1}{4}\cdot\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{2}\right) = 1/7$

$\quad\quad\;\; \theta=b \quad \frac{3}{4}\cdot\frac{1}{2} \Big/ \left(\frac{1}{4}\cdot\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{2}\right) = 6/7$

$x=2 \quad \theta=a \quad \frac{1}{4}\cdot\frac{1}{4} \Big/ \frac{1}{4}\cdot\frac{1}{4} = 1$

$\quad\quad\;\; \theta=b \quad \frac{3}{4}\cdot 0 \Big/ \frac{1}{4}\cdot\frac{1}{4} = 0$

$x=3 \quad \theta=a \quad \frac{1}{4}\cdot 0 \Big/ \frac{3}{4}\cdot\frac{1}{4} = 0$

$\quad\quad\;\; \theta=b \quad \frac{3}{4}\cdot\frac{1}{4} \Big/ \frac{3}{4}\cdot\frac{1}{4} = 1$

$x=4 \quad \theta=a \quad \frac{1}{4}\cdot\frac{1}{2} \Big/ \left(\frac{1}{4}\cdot\frac{1}{2}+\frac{3}{4}\cdot\frac{1}{4}\right) = 2/5$

$\quad\quad\;\; \theta=b \quad \frac{3}{4}\cdot\frac{1}{4} \Big/ \left(\frac{1}{4}\cdot\frac{1}{2}+\frac{3}{4}\cdot\frac{1}{4}\right) = 3/5$

Now $\int_a L(\theta,a)\,\delta(x,da) = \delta(x,\{a\})L(\theta,a)+\delta(x,\{b\})L(\theta,b)$

$= \delta(x,\{a\})L(\theta,a) + (1-\delta(x,\{a\}))L(\theta,b)$

When $x=1$ posterior risk is

$(1-\delta(1,\{a\}))\frac{1}{7} + \delta(1,\{a\})\frac{6}{7} = \frac{1}{7} + \frac{5}{7}\delta(1,\{a\})$

and this is minimized by taking $\delta(1,\{a\}) = 0$

When $x=2$ posterior risk is

$(1-\delta(2,\{a\}))\cdot 1 + \delta(2,\{a\})\cdot 0 = 1-\delta(2,\{a\})$

and this is minimized by taking $\delta(2,\{a\}) = 1$

When $x=3$ posterior risk is

$$(1-\delta(3,\{a\}))\,0 + \delta(3,\{a\})\,1 = \delta(3,\{a\})$$

and this is minimized by taking $\delta(3,\{a\}) = 0$.

When $x=4$ posterior risk is

$$(1-\delta(4,\{a\}))\tfrac{2}{5} + \delta(4,\{a\})\tfrac{3}{5} = \tfrac{2}{5} + \delta(4,\{a\})\tfrac{1}{5}$$

and this is minimized by taking $\delta(4,\{a\}) = 0$.

Therefore the Bayes rule is nonrandomized and is given by

$d(1) = b$ $\qquad$ $d(3) = b$
$d(2) = a$ $\qquad$ $d(4) = b$

$$R(\theta,d) = \int_{\{1,2,3,4\}} L(\theta, d(x))\, P_\theta(dx)$$

$\theta = a$
$$= \tfrac{1}{4}\cdot 1 + \tfrac{1}{4}\cdot 0 + 0\cdot 1 + \tfrac{1}{2}\cdot 1 = 3/4$$

$\theta = b$
$$= \tfrac{1}{2}\cdot 0 + 0\cdot 1 + \tfrac{1}{4}\,0 + \tfrac{1}{4}\cdot 0 = 0$$

d is admissible because to beat it another δ would have to have $R(b,\delta)=0$ which means it could only differ from d when $x=2$ which would choosing a with some positive probability which would make $R(a,\delta) > R(a,d)$.

3 So $\underline{\Theta}(\theta)=\theta$ and $\underline{\mathcal{A}}(\theta)=\Theta$ with

loss $\quad \mathrm{Loss}(\theta, a) = I_{\{a\}}(\theta)$. A δ takes the form

$$\delta_{p_1 p_2}(x, \{a\}) = \begin{cases} p_1, & x=1, \ a=1 \\ 1-p_1, & x=1, \ a=1/2 \\ p_2, & x=0, \ a=1 \\ 1-p_2, & x=0, \ a=1/2 \end{cases}$$

$$R(\theta, \delta_{p_1 p_2}) = \int_{\{0,1\}} \int_{\{\frac{1}{2},1\}} \mathrm{Loss}(\theta, a)\, \delta_{p_1 p_2}(x, da)\, P_\theta(dx)$$

$$= \left(\mathrm{Loss}(\theta,1) p_1 + \mathrm{Loss}(\theta,\tfrac{1}{2})(1-p_1)\right) P_\theta(\{1\})$$

$$+ \left(\mathrm{Loss}(\theta,1) p_2 + \mathrm{Loss}(\theta,\tfrac{1}{2})(1-p_2)\right) P_\theta(\{0\})$$

$$= \begin{cases} p_1 \frac{1}{2} + p_2 \frac{1}{2}, & \theta = 1/2 \\ (1-p_2), & \theta = 1 \end{cases}$$

$$\max_\theta R(\theta, \delta_{p_1 p_2}) = \begin{cases} \frac{p_1+p_2}{2} & \text{when } \frac{p_1+p_2}{2} \geq 1-p_1 \text{ iff } p_1 \geq \frac{2}{3} - \frac{p_2}{3} \\ 1-p_1 & \text{otherwise.} \end{cases}$$

$$\therefore \min_{p_1, p_2} \max_\theta R(\theta, \delta_{p_1 p_2}) = \frac{1}{3}$$

min $\frac{p_1+p_2}{2}$ when $p_1 = \frac{2}{3} - \frac{p_2}{3}$ and $p_2 = 0$ equals $1/3$ and similarly $\min(1-p_1)$ when $p_1 = 2/3$,

by $\delta_{\frac{2}{3},0}$ and so randomization is necessary to reduce maximum risk

4 (a) Let $G = \{e, g_1, g_2\}$ where $g_1: 1 \to 2 \to 3 \to 1$, $g_2: 1 \to 3 \to 2 \to 1$, $g_1^{-1} = g_2$.
Then clearly $P_\theta \circ g^{-1} \in \{P_\theta : \theta \in \Theta\}$ $\forall \theta \in \Theta, g \in G$.
Since G is the largest group of 1-1 transformations on $\{1,2,3\}$ this is the largest group leaving the model invariant.

(b) We have that the group $\bar{G}$ of transformations on Θ is the set of all permutations on $\Theta = \{a,b,c\}$. Since $\bar{G}$ acts transitively on Θ this implies that this is a group model.

(c) Since $\Psi(\theta_1) = \Psi(\theta_2)$ iff $\theta_1 = \theta_2$ we have that $\Psi(\bar{g}(\theta_1)) = \Psi(\bar{g}(\theta_2))$ whenever $\theta_1 = \theta_2$ and so Ψ is equivariant.

Suppose $\Psi(\theta_1) = \Psi(\theta_2)$ then $\theta_1, \theta_2 \in \{a,b\}$ or $\theta_1 = \theta_2 = c$. But $\bar{g}(a) = b$, $\bar{g}(b) = c$, $\bar{g}(c) = a$ is s.t. $\bar{g} \in \bar{G}$ but while $\Psi(a) = \Psi(b)$ we have $\Psi(\bar{g}(a)) = \Psi(b) = 1 \neq \Psi(\bar{g}(b)) = \Psi(c) = 0$. Therefore Ψ is not equivariant.

(d) Since $\Psi(\theta) = \theta$ and Ψ is equivariant we have $\bar{G} = G_*$. Then noting $Loss(\theta, \psi) = 1$ whenever $\psi \neq \theta$ and $Loss(\theta, \psi) = 0$ when $\psi = \theta$ we have $Loss(\bar{g}(\theta), \bar{g}(\psi)) = 0$ when $\theta = \psi$ and since $\bar{g}$ is 1-1 $\bar{g}(\theta) \neq \bar{g}(\psi)$ when $\theta \neq \psi$ so $Loss(\bar{g}(\theta), \bar{g}(\psi)) = 1 = Loss(\theta, \psi)$ which proves the loss is invariant.

(e) The MLE is given by $\hat{\theta}(1) = b$, $\hat{\theta}(2) = c$, $\hat{\theta}(3) = a$. Now $\hat{\theta}(x) = \hat{\theta}(y)$ iff $x = y$ so if $\hat{\theta}(x) = \hat{\theta}(y)$ we have $\hat{\theta}(g(x)) = \hat{\theta}(g(y))$ and conversely so the MLE is equivariant.

To determine the risk of the MLE we need only do this for one value of θ because $\bar{G}$ is transitive on Θ. So $R(\theta, \hat{\theta}) = R(a, \hat{\theta})$
$= \int_X L(a, \hat{\theta}(x)) P_a(dx) = P_a(\{1\}) + P_a(\{2\}) = 1/6 + 1/3 = 1/2$.

Any other equivariant estimator d has constant risk $R(a,d) = P_a(d(x)=b) + P_a(d(x)=c)$. Now suppose $d(1)=d(2)$ and suppose $g(1)=2$, $g(2)=3$ so $d(g(1)) = d(2) = d(g(2)) = d(3)$ and so $d(1)=d(2)=d(3)$ and d is constant. But if d is constant then $d(1) = d(g(1)) = \bar{g}(d(1))$ and if $\bar{g}: a \to b, b \to c, c \to a$ this cannot hold no matter $d(1)$ equals. A similar argument shows that d must be 1-1 (can't have $d(2)=d(3)$ or $d(1)=d(3)$). Then we must have $P_a(d(x)=b) + P_a(d(x)=c) \geq 1/6 + 1/3 = 1/2$ (these are the smallest error probabilities when $\theta = a$. Therefore $\hat{\theta}$ is optimal equivariant.

(f) For general invariant δ we need only consider the risk when $\theta = a$. Now

$$\int_{\{a,b,c\}} L(a,t)\,\delta(x,dt) = \delta(x,\{b\}) + \delta(x,\{c\}) \text{ so}$$
$$= 1 - \delta(x,\{a\})$$

so $R(a,\delta) = E_a(1-\delta(x,\{a\})) = 1 - \frac{1}{6}\delta(1,\{a\}) - \frac{1}{3}\delta(2,\{a\}) - \frac{1}{2}\delta(3,\{a\})$

and note by invariance $\delta(2,\{a\}) = \delta(1,\{c\})$ under g_1 and $\delta(3,\{a\}) = \delta(1,\{b\})$ under g_2. Therefore

$$R(a,\delta) = 1 - \frac{1}{6}\delta(1,\{a\}) - \frac{1}{3}\delta(1,\{c\}) - \frac{1}{2}\delta(1,\{b\})$$
$$= 1 - \frac{1}{6}\delta(1,\{a\}) - \frac{1}{3}\delta(1,\{c\}) - \frac{1}{2}\left(1 - \delta(1,\{a\}) - \delta(1,\{c\})\right)$$
$$= \frac{1}{2} + \frac{1}{3}\delta(1,\{a\}) + \frac{1}{6}\delta(1,\{c\}) \geq \frac{1}{2}$$

and so the MLE is optimal invariant.