1 Bayesian Analysis I (Shawn Lyu)

1. Write out the posterior expected loss as
\[ \int_{\Theta} w(\theta)(\theta - a)^2 dF^{\pi(\theta|x)}(\theta) \]

The loss, a weighted sum of convex functions, is a Take the derivative wrt a and set it to 0 we have
\[ 0 = \frac{d}{da} \int_{\Theta} w(\theta)(\theta - a)^2 dF^{\pi(\theta|x)}(\theta) \]

which can be expanded to
\[ 0 = \frac{d}{da} \int_{\Theta} (w(\theta)\theta^2 - 2aw(\theta)\theta + a^2w(\theta))^2 dF^{\pi(\theta|x)}(\theta) = 2E_{\pi(\theta|x)}[\theta w(\theta)] - 2aE_{\pi(\theta|x)}[w(\theta)] \]

Solving this just shows that the Bayes rule is
\[ \delta^\pi(x) = \frac{E_{\pi(\theta|x)}[\theta w(\theta)]}{E_{\pi(\theta|x)}[w(\theta)]} \]

2. (a) From the Berger book, section 4.4.2 result 3, we know that the Bayes rule for this kind of loss is
\[ \delta^\pi(x) = E_{\pi}[\theta] \]

The goal is then for us to find the posterior distribution \( \pi(\theta|X = x) \). Since Beta distribution is the conjugate prior to binomial distribution we know that
\[ \theta|X \sim Beta(x + 3, 14 - x) \]

and that
\[ \delta^\pi(x) = \frac{x + 3}{17} \]

(b) From the Berger book, section 4.4.2 result 5 the Bayes estimator is the median of the posterior, which tells us directly that
\[ \delta^\pi(x) = \frac{3x + 8}{49} \]
(c) From the Berger book, section 4.4.2 result 4 (or the result above), we know that the Bayes rule for weighted loss is

\[ \delta^\pi(x) = \frac{\mathbb{E}_{\pi}[\theta w(\theta)]}{\mathbb{E}_{\pi}[w(\theta)]} \]

Note that since \( \theta|x \sim Beta(x + 3, 14 - x) \), expanding the pdf yields

\[
\mathbb{E}_{\pi}[\theta w(\theta)] = \mathbb{E}_{\pi}[\theta] \cdot w(\theta) = \int_0^1 \theta^{x+2} (1 - \theta)^{13-x} \frac{1}{B(x + 3, 14 - x) \theta(1 - \theta)} d\theta
\]

\[
= \frac{\Gamma(17) \Gamma(x + 3) \Gamma(14 - x)}{\Gamma(15) \Gamma(x + 2) \Gamma(13 - x) \Gamma(15)} \frac{\Gamma(15)}{\Gamma(x + 2) \Gamma(13 - x)} \theta^{x+1} (1 - \theta)^{12-x} d\theta
\]

\[
= \frac{\Gamma(17) \Gamma(x + 2) \Gamma(13 - x)}{\Gamma(16) \Gamma(x + 3) \Gamma(14 - x) \Gamma(16)} \int_0^1 \frac{\Gamma(15)}{\Gamma(x + 2) \Gamma(13 - x)} \theta^{x+1} (1 - \theta)^{12-x} d\theta
\]

\[
= \frac{16!}{14!(x + 2)(13 - x)}
\]

We then find the weighted expected value of \( \theta \), which is just

\[
\mathbb{E}_{\pi}[\theta w(\theta)] = \mathbb{E}_{\pi}[\theta] \cdot w(\theta) = \frac{\Gamma(17) \Gamma(x + 3) \Gamma(14 - x)}{\Gamma(16) \Gamma(x + 3) \Gamma(14 - x)} \theta^{x+2} (1 - \theta)^{12-x} d\theta
\]

\[
= \frac{\Gamma(17) \Gamma(x + 3) \Gamma(13 - x)}{\Gamma(16) \Gamma(x + 3) \Gamma(14 - x)} \int_0^1 \frac{\Gamma(16)}{\Gamma(x + 3) \Gamma(13 - x)} \theta^{x+2} (1 - \theta)^{12-x} d\theta
\]

\[
= \frac{16}{(13 - x)}
\]

Plug these two together we have

\[ \delta^\pi(x) = \frac{x + 2}{15} \]
2 Bayesian Analysis II (Shawn Lyu)

1. Based on the constraint write the decision rule as

$$\delta^*(x) = (a_1, 1 - a_1)^T$$

we have that the loss can be written as

$$\int \theta_1^2 a_1^2 + \theta_2^2 + \theta_2^2 a_1^2 + 1 + 2\theta_1 \theta_2 a_1 - 2\theta_1 \theta_2 a_1 - 2\theta_1 a_1 - 2\theta_2 a_1 - 2\theta_2 a_1 dF^\pi(\theta)$$

Set the derivative wrt to $a_1$ to 0 we have

$$0 = 2E[\theta_1^2]a_1 + 2E[\theta_2^2]a_1 + 2E[\theta_1 \theta_2] - 4E[\theta_1 \theta_2]a_1 - 2E[\theta_1] - 2E[\theta_2^2] + 2E[\theta_2]$$

Since the conjugate prior of Gaussian is Gaussian, we have

$$\theta | x \sim N\left(\frac{x + \mu}{2}, \frac{1}{2}I_2\right)$$

We then have

$$E[\theta_1^2] = \frac{(x_1 + \mu_1)^2 + 2}{4}$$

$$E[\theta_2^2] = \frac{(x_2 + \mu_2)^2 + 2}{4}$$

$$E[\theta_1] = \frac{x_1 + \mu_1}{2}$$

$$E[\theta_2] = \frac{x_2 + \mu_2}{2}$$

$$E[\theta_1 \theta_2] = \frac{(x_1 + \mu_1)(x_2 + \mu_2)}{4}$$

Plug these back on top to obtain

$$\frac{(x_1 + \mu_1)^2 + (x_2 + \mu_2)^2 + 4 - 2(x_1 + \mu_1)(x_2 + \mu_2)}{2} a_1 = -\frac{(x_1 + \mu_1)(x_2 + \mu_2)}{2} + x_1 + \mu_1 + \frac{(x_2 + \mu_2)^2}{2} + 1 - x_2 - \mu_2$$

So

$$a_1 = \frac{(x_2 + \mu_2)^2 + 2 + 2(x_1 + \mu_1) - 2(x_2 + \mu_2) - (x_1 + \mu_1)(x_2 + \mu_2)}{(x_1 + \mu_1)^2 + (x_2 + \mu_2)^2 + 4 - 2(x_1 + \mu_1)(x_2 + \mu_2)}$$

and the other parameter is

$$a_2 = \frac{(x_1 + \mu_1)^2 + 2 + 2(x_2 + \mu_2) - 2(x_1 + \mu_1) - (x_1 + \mu_1)(x_2 + \mu_2)}{(x_1 + \mu_1)^2 + (x_2 + \mu_2)^2 + 4 - 2(x_1 + \mu_1)(x_2 + \mu_2)}$$

Finally, our decision rule is just

$$\delta^*(x) = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

2. Expand the loss

$$\int \theta^T Q \theta - 2\theta^T Q a + a^T Q a dF^\pi(\theta)$$

Find the derivative wrt $a$ and set to 0 to obtain

$$0 = 2E[\theta^T]Q - 2a^T Q$$

We then easily know that

$$\delta^*(x) = E^\pi[\theta]$$
3 Minimax Analysis I (Karthika Nair and Yao-Hung Tsai)

1. Let \( \delta' \) be another decision rule. Since \( \delta \) is admissible, we can always find a point \( \theta' \) depending on \( \delta' \) s.t.
\[
R(\theta_0, \delta) \leq R(\theta', \delta') \leq \sup_\theta R(\theta, \delta')
\]
Since \( \delta \) has constant risk, we have the following inequalities
\[
\sup_\theta R(\theta, \delta) = R(\theta_0, \delta) \leq R(\theta', \delta') \leq \sup_\theta R(\theta, \delta').
\]
(4 pts) Note that \( \delta' \) can be arbitrary decision rule. Therefore, \( \delta \) is minimax. (1 pts) If unique Bayes, then unique minimax.

2. (a) \( A = \{a_1, a_2, a_3\} \)
where \( a_1 = \) Order of 15, \( a_2 = \) Order of 30, \( a_3 = \) Order of 45
\[
\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}
\]
where \( \theta_1 = \) Demand of 20, \( \theta_2 = \) Demand of 25 and \( \theta_3 = \) Demand of 30, \( \theta_4 = \) Demand of 45

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( a_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>-425</td>
<td>-350</td>
<td>-50</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>-400</td>
<td>-700</td>
<td>-400</td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>-375</td>
<td>-1050</td>
<td>-750</td>
</tr>
<tr>
<td>( \theta_4 )</td>
<td>-300</td>
<td>-975</td>
<td>-1800</td>
</tr>
</tbody>
</table>

(b) All actions are admissable.
(c) \( p(\pi, a_1) = 0.2(-425) + 0.4(-400) + 0.2(-375) + 0.2(-300) = -380 \)
\( p(\pi, a_2) = 0.2(-350) + 0.4(-700) + 0.2(-1050) + 0.2(-975) = -755 \)
\( p(\pi, a_3) = 0.2(-50) + 0.4(-400) + 0.2(-750) + 0.2(1800) = -680 \)
The Bayes action is \( a_2 \).
(d) \( \sup_\theta R(\theta, a_1) = -300 \)
\( \sup_\theta R(\theta, a_2) = -350 \)
\( \sup_\theta R(\theta, a_3) = -50 \)
The minimax action is \( a_2 \) and its corresponding value is -350.

3. (a) \( A = \{a_1; \text{without marketing campaign}, a_2; \text{with marketing campaign}\} \)
\( \Theta = [0,1] \)
\[
L(\theta, a_1) = \begin{cases} 
-(2 + 3\theta), & 0 \leq \theta < 0.7 \\
-5, & 0.7 \leq \theta \leq 1 
\end{cases}
\]
\[
L(\theta, a_2) = \begin{cases} 
-(4 + 1.5\theta), & 0 \leq \theta < 0.7 \\
-4, & 0.7 \leq \theta \leq 1 
\end{cases}
\]
(b) \( \rho(\pi, a_1) = E^\pi[L(\theta, a_1)] \)
\[
= \int_0^1 L(\theta, a_1) \pi(\theta) d\theta \\
= \int_0^0.7 -(2 + 3\theta) d\theta + \int_{0.7}^1 (-5) d\theta
\]
\[ \rho(\pi, a_2) = E^\pi[L(\theta, a_2)] \]
\[ = \int_0^1 L(\theta, a_2)\pi(\theta)d\theta \]
\[ = \int_0^{0.7} -(-4 + 1.5\theta)d\theta + \int_{0.7}^1 (-4)d\theta \]
\[ = -3.1675 - 1.2 \]
\[ = -4.3675 \]
Hence, \( a_2 \) is the Bayesian action.

(c) \( \sup_\theta (R(\theta, a_1)) = \sup_\theta (L(\theta, a_1)) = -2 \) (when \( \theta = 0 \))
\( \sup_\theta (R(\theta, a_2)) = \sup_\theta (L(\theta, a_2)) = -4 \) (when \( \theta = 0 \))
Hence, \( a_2 \) is the minimax action.
4 Minimax Analysis II (Karthika Nair and Yao-Hung Tsai)

1. For any other estimator $\delta'$, we have

$$\begin{align*}
\sup_{\theta} R(\theta, \delta') &\geq \int R(\theta, \delta') \pi(d\theta) \\
&\geq \int R(\theta, \delta^*_0) \pi(d\theta) = \sup_{\theta} R(\theta, \delta^*_0).
\end{align*}$$

(2 pts)

Therefore, $\delta^*_0$ is minimax. If $\delta^*_0$ is unique Bayes under $\pi$, then the second inequality is strict and thus makes $\delta^*_0$ to be a unique minimax. (1 pt)

Let $\pi'$ be another prior over $\Theta$. Then we have

$$\begin{align*}
R(\pi', \delta_{\pi'}) &= \int R(\theta, \delta_{\pi'}) \pi'(d\theta) \\
&\leq \int R(\theta, \delta^*_0) \pi'(d\theta) \\
&\leq \sup_{\theta} R(\theta, \delta^*_0) = R(\pi, \delta^*_0)
\end{align*}$$

(2 pts)

Therefore, $\pi$ is least favorable. (1 pt)

2. Let $\delta_0$ be an equalizer rule with $R(\theta, \delta_0) = C$. If $\delta_0$ was not minimax, $\exists \delta'$ s.t. $\sup_{\theta} R(\theta, \delta') = v$ would be strictly less than $C$. (2 pts)

For any $\epsilon > 0$ s.t. $0 < \epsilon < C - v$ and any prior $\pi$, we have

$$r(\pi, \delta_0) = c > v + \epsilon \geq r(\pi, \delta') + \epsilon.$$

(2 pts)

Thus, $\delta_0$ cannot be $\epsilon$-Bayes w.r.t. to any prior distribution, and so $\delta_0$ can not be an extended Bayes rule. We prove this by contradiction. (2 pts)
R(\theta, \delta) = \mathbb{E}_X \left( \frac{(\theta - \delta(X))^2}{1 - \theta} \right)
\begin{align*}
&= \sum_{x=0}^{\infty} (\theta^2 - 2\theta\delta(x) + \delta(x)^2)\theta^x \\
&= \delta(0)^2 + \theta(\delta(1)^2 - 2\delta(0)) + \sum_{x=2}^{\infty} (\delta(x)^2 - 2\delta(x-1) + 1)\theta^x
\end{align*}

(4 pts)

(b)
\begin{align*}
\delta(1)^2 &= 2\delta(0) \\
\delta(x)^2 &= 2\delta(x-1) - 1 \text{ for } x = 2, 3, \cdots
\end{align*}

(2 pts)

Since $2\delta(x-1) = \delta(x)^2 + 1 = (\delta(x)-1)^2 + 2\delta(x)$, so that $\delta(x)$ are non-decreasing from $x = 1$ on. (1 pt)

Therefore, $\delta(x)$ converges, and suppose its value is $c$ we have $c^2 = 2c - 1$. This implies $c = 1$. (1 pt)

To conclude, we have
\begin{align*}
\delta(0) &= \frac{1}{2} \text{ and } \\
\delta(x) &= 1 \text{ for } x = 1, 2, 3, \cdots
\end{align*}

(2 pts)

(c) From part (i),
\begin{align*}
r(\pi, \delta) &= \mathbb{E}R(\theta, \delta) \\
&= \delta(0)^2 + \mu_1(\delta(1)^2 - 2\delta(0)) + \sum_{x=2}^{\infty} \mu_x(\delta(x)^2 - 2\delta(x-1) + 1).
\end{align*}

(2 pts)

To find the Bayes rule, we take the derivative w.r.t. to each $\delta(x)$ for the above equation, setting to zero. (2 pts)

We then obtain
\begin{align*}
\delta_\pi(x) &= \frac{m_{x+1}}{m_x}
\end{align*}

for $x = 0, 1, 2, \cdots$. (1 pt)