Solutions to Problem 7 & 11
Homework Assignment 2

AMS 206B, WINTER 2016

Prepared by: Sharmistha Guha and Arthur Lui

January 27, 2016

Problem 7, HW 2

Consider the LINUX loss function defined by

$$L(\theta - \theta(x)) = e^{c(\theta-\theta(x))} - c(\theta - \theta(x)) - 1.$$

(a) Show that $L(\theta-\theta(x) \geq 0$ and plot this as a function of $(\theta-\theta(x))$ when $c = .1, .5, 1, 2$.
(b) Find $\hat{\theta}(x)$ the estimator that minimizes the Bayesian expected posterior loss.
(c) Find $\hat{\theta}(x)$ when $x_1, ...x_n \mid \theta \sim N(\theta, 1), \theta \sim N(\mu, \tau^2)$.

Solution:

(a)

Let $a = \theta - \theta(x)$, then $L(\theta - \theta(x)) = L(a) = e^a - a - 1$.

$$\begin{aligned}
\frac{d}{da}L(a) &:= 0 \\
e^a - 1 &= 0 \\
a &= \ln(1) \\
a &= 0
\end{aligned}$$

So, $L(a)$ reaches a local extrema at $a = 0 \Rightarrow L(a) = e^0 - 0 - 1 = 0$. That is, there is a local extrema at $(0, 0)$. And since $L''(a) = e^a > 0$, $(0, 0)$ is a local minimum.

Therefore, $L(\theta - \theta(x)) \geq 0$.

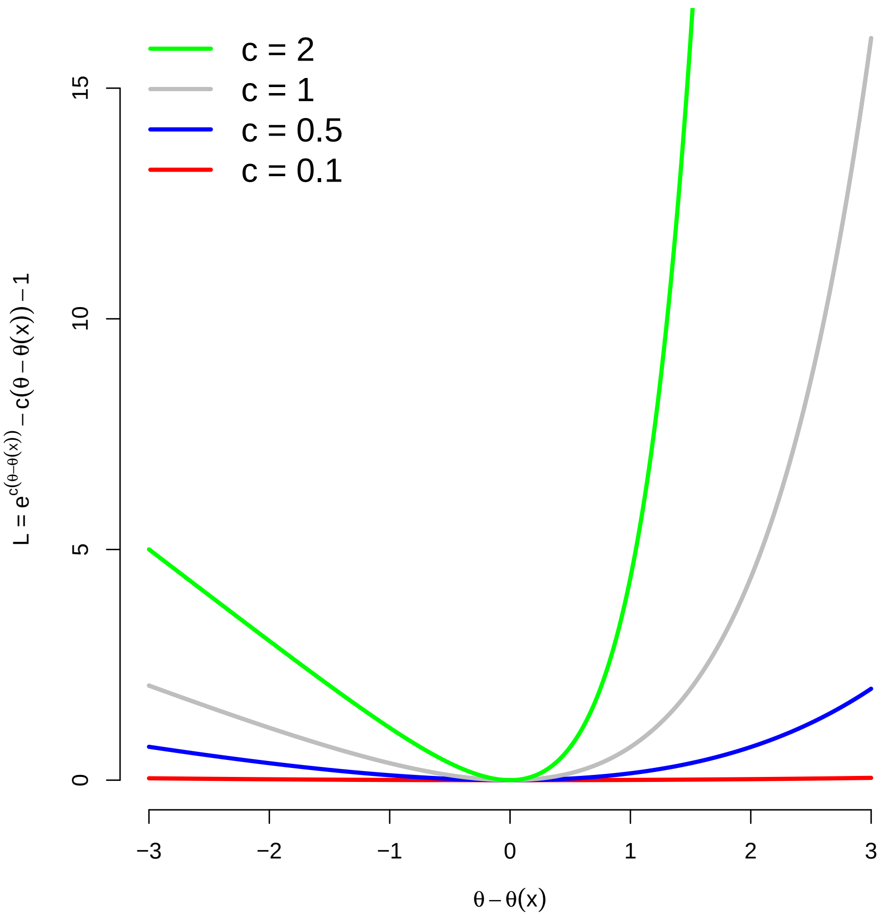

Figure 1: Plot for 7(a). Linex loss as a function of $\theta - \theta(x)$, evaluated at various c.

(b)

Let $f(\theta|\boldsymbol{x})$ be the posterior density for θ. Then, the expected posterior loss is

$$
\begin{aligned}
E[L(\theta - \theta(x))] &= \int_\Theta \left\{ e^{c(\theta - \theta(x))} - c(\theta - \theta(x)) - 1 \right\} f(\theta|x)\ d\theta \\
&= e^{-c\theta(x)} \int_\Theta e^{c\theta} f(\theta|x)\ d\theta - c \int_\Theta \theta f(\theta|x)\ d\theta + c\theta(x) \int_\Theta f(\theta|x)\ d\theta - \int_\Theta f(\theta|x)\ d\theta \\
&= e^{-c\theta(x)} E_{\theta|x}[e^{c\theta}|x] - cE_{\theta|x}[\theta|x] + c\theta(x) - 1
\end{aligned}
$$

To minimize the expected posterior loss,

$$\begin{aligned}
\frac{d}{d\theta(x)} E[L(\theta - \theta(x))] &:= 0 \\
-ce^{-c\hat{\theta}} E_{\theta|x}[e^{c\theta}|x] + c = 0 \\
e^{c\hat{\theta}} &= E_{\theta|x}[e^{c\theta}|x] \\
\\
\Rightarrow \hat{\theta} &= \frac{\ln(E_{\theta|x}[e^{c\theta}|x])}{c}
\end{aligned}$$

This is simply the log of the posterior moment generating function divided by c.

(c)

$\hat{\theta}(x)$ is simply the log of posterior mgf divided by c. So, all we need to do is compute the posterior of θ.

$$\begin{aligned}
f_{\theta|\boldsymbol{x}}(\theta|\boldsymbol{x}) &\propto \exp\left\{-\frac{\sum_{i=1}^{n}(x_i-\theta)^2}{2}\right\} \exp\left\{-\frac{(\theta-\mu)^2}{2\tau^2}\right\} \\
&\propto \exp\left\{\frac{-n\theta^2\tau^2 + 2n\bar{x}\theta\tau^2 - \theta^2 + 2\theta\mu}{2\tau^2}\right\} \\
&\propto \exp\left\{\frac{-\theta^2(n\tau^2+1) + 2\theta(n\bar{x}\tau^2+\mu)}{2\tau^2}\right\} \\
&\propto \exp\left\{\frac{-\theta^2 + 2\theta(\frac{n\bar{x}\tau^2+\mu}{n\tau^2+1})}{2\frac{\tau^2}{n\tau^2+1}}\right\}
\end{aligned}$$

$$\Rightarrow \theta|\boldsymbol{x} \sim N(\tilde{m}, \tilde{v}),$$

where $\tilde{m} = \dfrac{n\bar{x}\tau^2+\mu}{n\tau^2+1}$ and $\tilde{v} = \dfrac{\tau^2}{n\tau^2+1}$. Finally, using the MGF of the Normal distribution,

$$
\begin{aligned}
\hat{\theta}(x) &= \frac{\ln(\exp\{\tilde{m}c + \tilde{v}c^2/2\})}{c} \\
&= \frac{\tilde{m}c + \tilde{v}c^2/2}{c} \\
&= \tilde{m} + \frac{\tilde{v}c}{2}
\end{aligned}
$$

$$
\Rightarrow \hat{\theta}(x) = \frac{(n\bar{x} + c/2)\tau^2 + \mu}{n\tau^2 + 1}
$$

Problem 11, HW 2

Given that X_1, X_2 are two independent observations from

$$
P(X = \theta - 1|\theta) = P(X = \theta + 1|\theta) = \frac{1}{2} \tag{1}
$$

where θ is an integer.
We are provided with the 0 -1 Loss Function

$$
L(\theta, \theta(X_1, X_2)) = \begin{cases} 1 & \text{if } \theta(X_1, X_2) \neq \theta \\ 0 & \text{o.w.} \end{cases}
$$

(a)

(i) We proceed to find the risk of the estimator $\theta_0(X_1, X_2) = \frac{X_1+X_2}{2}$.
Let $\mathbf{X} = (X_1, X_2)'$

$$
\begin{aligned}
R(\theta, \theta_0(\mathbf{X})) &= \sum_{x_1 \in \{\theta-1, \theta+1\}} \sum_{x_2 \in \{\theta-1, \theta+1\}} L(\theta, \theta(x_1, x_2)) P(X_1 = x_1, X_2 = x_2) \\
&= 1.P(X_1 = \theta - 1, X_2 = \theta - 1) + 1.P(X_1 = \theta + 1, X_2 = \theta + 1) \\
&+ 0.P(X_1 = \theta - 1, X_2 = \theta + 1) + 0.P(X_1 = \theta + 1, X_2 = \theta - 1) \\
&= P(X_1 = \theta - 1)P(X_2 = \theta - 1) + P(X_1 = \theta + 1)P(X_2 = \theta + 1) \\
&= \frac{1}{2}.\frac{1}{2} + \frac{1}{2}.\frac{1}{2} = \frac{1}{2}.
\end{aligned} \tag{2}
$$

(ii) Now considering the second estimator $\theta_1(X_1, X_2) = X_1 + 1$.

$$\begin{aligned}
R(\theta, \theta_1(\mathbf{X})) &= \sum_{x_1 \in \{\theta-1, \theta+1\}} \sum_{x_2 \in \{\theta-1, \theta+1\}} L(\theta, \theta(x_1, x_2)) P(X_1 = x_1, X_2 = x_2) \\
&= 1.P(X_1 = \theta + 1, X_2 = \theta + 1) + 1.P(X_1 = \theta + 1, X_2 = \theta - 1) \\
&+ 0.P(X_1 = \theta - 1, X_2 = \theta + 1) + 0.P(X_1 = \theta - 1, X_2 = \theta - 1) \\
&= P(X_1 = \theta + 1)P(X_2 = \theta + 1) + P(X_1 = \theta + 1)P(X_2 = \theta - 1) \\
&= \frac{1}{2}.\frac{1}{2} + \frac{1}{2}.\frac{1}{2} = \frac{1}{2}. \qquad (3)
\end{aligned}$$

The two estimators are the same as far as "Frequentist Risk" is concerned.

(b) Need to find an estimator $\hat{\theta}(X_1, X_2)$ that minimizes Bayesian Expected Loss.

Bayesian Expected Loss $= E_{\theta|\mathbf{X}}(L(\theta, \theta(\mathbf{X})))$

$$\begin{aligned}
E_{\theta|\mathbf{X}}(L(\theta, \theta(\mathbf{X}))) &= 1.P_{\theta|\mathbf{X}}(\theta \neq \theta(\mathbf{X})) + 0.P_{\theta|\mathbf{X}}(\theta = \theta(\mathbf{X})) \qquad (4) \\
&= P_{\theta|\mathbf{X}}(\theta \neq \theta(\mathbf{X})) = 1 - P_{\theta|\mathbf{X}}(\theta = \theta(\mathbf{X})). \qquad (5)
\end{aligned}$$

In order to minimize the Bayesian Expected Loss, we need to maximize $P_{\theta|\mathbf{X}}(\theta = \theta(\mathbf{X}))$.
This implies $\hat{\theta}(\mathbf{X}) = arg \max_{\theta} P_{\theta|\mathbf{X}}$.

Let the prior distribution of θ be $\mathbf{p}(\cdot)$. Consider the following cases:

Case 1: $X_1 \neq X_2$

$P(\theta = \frac{X_1+X_2}{2} | X_1, X_2) = 1$.
So, the mode of the posterior distribution is $\frac{X_1+X_2}{2}$.
Therefore, when $X_1 \neq X_2$, $\hat{\theta}(\mathbf{X}) = \frac{X_1+X_2}{2}$.

Case 2: $X_1 = X_2$

The posterior distribution of θ has mass on $X_1 + 1$ and $X_1 - 1$ and the posterior p.m.f is given by

$$P(\theta = X_1 + 1 | \mathbf{X}) = \frac{\mathbf{p}(\theta = X_1 + 1)}{\mathbf{p}(\theta = X_1 + 1) + \mathbf{p}(\theta = X_1 - 1)} \qquad (6)$$

and

$$P(\theta = X_1 - 1 | \mathbf{X}) = \frac{\mathbf{p}(\theta = X_1 - 1)}{\mathbf{p}(\theta = X_1 + 1) + \mathbf{p}(\theta = X_1 - 1)}. \qquad (7)$$

Thus the mode of $\theta|\mathbf{X}$ is (X_1+1) if $\mathbf{p}(\theta = X_1+1) \geq \mathbf{p}(\theta = X_1-1)$. Also, the mode of $\theta|\mathbf{X}$ is (X_1-1) if $\mathbf{p}(\theta = X_1+1) < \mathbf{p}(\theta = X_1-1)$.
Hence

$$\hat{\theta}(X_1, X_2) = \begin{cases} \frac{X_1+X_2}{2}, & \text{If} \quad X_1 \neq X_2 \\ X_1+1, & \text{If } X_1 = X_2 \text{ and } \mathbf{p}(\theta = X_1+1) \geq \mathbf{p}(\theta = X_1-1) \\ X_1-1, & \text{If } X_1 = X_2 \text{ and } \mathbf{p}(\theta = X_1+1) < \mathbf{p}(\theta = X_1-1) \end{cases}$$