Q2 (a) $\boxed{\text{MSE}}$

Let k be the number of x_i in R_{jm}

$$\gamma_{jm} = \underset{r}{\operatorname{argmin}} \sum_{x_i \in R_{jm}} (y_i - f_{m-1}(x_i) - r)^2$$

$$0 = \frac{\partial}{\partial r} \sum_{x_i \in R_{jm}} (y_i - f_{m-1}(x_i) - r)^2$$

$$0 = \sum_{x_i \in R_{jm}} 2(y_i - f_{m-1}(x_i) - r)$$

$$0 = \sum_{x_i \in R_{jm}} 2y_i - 2f_{m-1}(x_i) - 2r$$

$$0 = \sum_{x_i \in R_{jm}} 2y_i - 2f_{m-1}(x_i) - \sum_{x_i \in R_{jm}} 2r$$

$$0 = \sum_{x_i \in R_{jm}} 2y_i - 2f_{m-1}(x_i) - 2kr$$

$$2kr = \sum_{x_i \in R_{jm}} 2y_i - 2f_{m-1}(x_i)$$

$$r_{jm}^* = \frac{1}{2k} \sum_{x_i \in R_{jm}} 2y_i - 2f_{m-1}(x_i)$$

$$\gamma_{jm}^* = \frac{1}{k} \sum_{x_i \in R_{jm}} y_i - f_{m-1}(x_i)$$

2(a) $\boxed{\text{Binomial deviance}}$

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

$$D = -2\left[y_i \cdot \log(p) + (1-y_i)\cdot \log(1-p)\right]$$

$$= -2\left[y_i \log(p) + \log(1-p) - y_i \log(1-p)\right]$$

$$= -2\left[y_i \left(\log(p) - \log(1-p)\right) + \log(1-p)\right]$$

$$= -2\left[y_i \log\left(\frac{p}{1-p}\right) + \log(1-p)\right]$$

$$= -2\left[y_i \log(odds) - \log(1-p)\right] \quad *$$

$$= -2\left[y_i \log(odds) - \log\left(1+e^{\log(odds)}\right)\right]$$

*

$$\log(1-p) = \log\left(1 - \frac{e^{\log(odds)}}{1+e^{\log(odds)}}\right) = \log\left(\frac{1+e^{\log(odds)}}{1+e^{\log(odds)}} - \frac{e^{\log(odds)}}{1+e^{\log(odds)}}\right)$$

$$= \log\left(\frac{1}{1+e^{\log(odds)}}\right)$$

$$= \log(1) - \log\left(1+e^{\log(odds)}\right)$$

$$= -\log\left(1+e^{\log(odds)}\right)$$

$$\frac{\partial}{\partial \log(odds)} -2\left[y_i \log(odds) - \log\left(1+e^{\log(odds)}\right)\right]$$

$$= -2\left[y_i - \frac{e^{\log(odds)}}{1+e^{\log(odds)}}\right]$$

$$= -2\left[y_i - p\right]$$

$$\frac{\partial^2}{\partial^2 \log(odds)} -2\left[y_i \log(odds) - \log(1+e^{\log(odds)})\right]$$

$$= 2 \frac{\partial}{\partial^2 \log(odds)}\left[-y_i + \frac{e^{\log(odds)}}{1+e^{\log(odds)}}\right]$$

$$= 2\left[-(1+e^{\log(odds)})^{-2} e^{\log(odds)} \cdot e^{\log(odds)} + (1+e^{\log(odds)})^{-1} e^{\log(odds)}\right.$$

$$= 2\left[\frac{-e^{2\log(odds)}}{(1+e^{\log(odds)})^2} + \frac{e^{\log(odds)}}{(1+e^{\log(odds)})}\right]$$

$$= 2\left[\frac{-e^{2\log(odds)}}{(1+e^{\log(odds)})^2} + \frac{e^{\log(odds)} + e^{2\log(odds)}}{(1+e^{\log(odds)})^2}\right]$$

$$= 2\left[\frac{-e^{2\log(odds)} + e^{\log(odds)} + e^{2\log(odds)}}{(1+e^{\log(odds)})^2}\right]$$

$$= 2 \frac{e^{\log(odds)} \times 1}{(1+e^{\log(odds)})(1+e^{\log(odds)})}$$

$$= 2 \cdot \frac{e^{\log(odds)}}{1+e^{\log(odds)}} \cdot \frac{1}{1+e^{\log(odds)}}$$

$$= 2p(1-p)$$

$$L(y_i, F_{m-1}(x_i)+r) \overset{\text{Taylor expansion}}{\approx} L(y_i, F_{m-1}(x_i)+r) + \frac{\partial}{\partial F} L(y_i, F_{m-1}(x_i))\, r + \frac{1}{2}\frac{\partial^2}{\partial^2 F} L(y_i, F_{m-1}(x_i))\, r^2$$

$$\because\ r_{jm} = \underset{r}{\arg\min} \sum_{x_i \in R_{jm}} L(y_i, F_{m-1}(x_i)+r)$$

$$0 = \frac{\partial}{\partial r} \sum_{x_i \in R_{jm}} L(y_i, F_{m-1}(x_i)+r)$$

$$0 = \frac{\partial}{\partial r} \sum_{x_i \in R_{jm}} \left\{ L(y_i, F_{m-1}(x_i)+r) + \frac{\partial}{\partial F} L(y_i, F_{m-1}(x_i))\, r + \frac{1}{2}\frac{\partial^2}{\partial^2 F} L(y_i, F_{m-1}(x_i))\, r^2 \right\}$$

$$\sum_{x_i \in R_{jm}} \left[\frac{\partial}{\partial F} L(y_i, F_{m-1}(x_i)) + \frac{\partial^2}{\partial F^2} L(y_i, F_{m-1}(x_i)) r \right] = 0$$

$$r = \frac{-\sum_{x_i \in R_{jm}} \frac{\partial}{\partial F} L(y_i, F_{m-1}(x_i))}{\sum_{x_i \in R_{jm}} \frac{\partial^2}{\partial^2 F} L(y_i, F_{m-1}(x_i))}$$

$$= \frac{\sum_{x_i \in R_{jm}} 2(y_i - p)}{\sum_{x_i \in R_{jm}} 2p(1-p)}$$

$$= \frac{\sum_{x_i \in R_{jm}} (y_i - p)}{\sum_{x_i \in R_{jm}} p(1-p)}$$

$$\because L(y_i, F_{m-1}(x_i)) = -2\left[y_i F_{m-1}(x_i) - \log(1 + e^{F_{m-1}(x_i)})\right]$$

$$\therefore \frac{\partial}{\partial F} L(y_i, F_{m-1}(x_i)) = -2(y_i - p)$$

$$\frac{\partial^2}{\partial^2 F} L(y_i, F_{m-1}(x_i)) = 2p(1-p)$$

$$\log\left(\frac{p}{1-p}\right) = \log(\text{odds}) = F_{m-1}(x_i)$$

$$e^{\log\left(\frac{p}{1-p}\right)} = e^{F_{m-1}(x_i)}$$

$$\frac{p}{1-p} = e^{F_{m-1}(x_i)}$$

$$p = e^{F_{m-1}(x_i)} - e^{F_{m-1}(x_i)} p$$

$$(1 + e^{F_{m-1}(x_i)}) p = e^{F_{m-1}(x_i)}$$

$$p = \frac{e^{F_{m-1}(x_i)}}{1 + e^{F_{m-1}(x_i)}}$$

$$r^*_{jm} = \frac{\sum_{x_i \in R_{jm}} (y_i - p)}{\sum_{x_i \in R_{jm}} p(1-p)} \quad \text{where} \quad p = \frac{e^{F_{m-1}(x_i)}}{1 + e^{F_{m-1}(x_i)}}$$

2(b) $$Obj(\theta)^t = \sum_{i=1}^{n} L(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + \text{Constant}$$

where $\Omega(f_t) = rT + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$

Let $g_i = \frac{\partial}{\partial \hat{y}^{(t-1)}} L(y_i, \hat{y}^{(t-1)})$ $\quad h_i = \frac{\partial^2}{\partial^2 \hat{y}^{(t-1)}} L(y_i, \hat{y}^{(t-1)})$

Using taylor approximation:

constant

$$Obj(\theta)^t \approx \sum_{i=1}^{n} \left[L(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) + C$$

$$\propto \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t)$$

$$= \sum_{i=1}^{n} \left[g_i w_{L(x_i)} + \frac{1}{2} h_i w_{L(x_i)}^2 \right] + rT + \lambda \frac{1}{2} \sum_{j=1}^{T} w_j^2$$

$$= \sum_{j=1}^{T} \left[\left(\textstyle\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\textstyle\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + rT$$

$$\boxed{\begin{aligned} &\sum_{j=1}^{T} \left(\textstyle\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \\ &= \sum_{j=1}^{T} \frac{1}{2} \left(\textstyle\sum_{i \in I_j} h_i \right) w_j^2 + \frac{1}{2} \lambda w_j^2 \\ &= \sum_{j=1}^{T} \frac{1}{2} \left(\textstyle\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \end{aligned}}$$

Let $G_j = \sum_{i \in I_j} g_i$

$\quad H_j = \sum_{i \in I_j} h_i$

$$\therefore \hat{Obj}(\theta)^t = \sum_{j=1}^{T} \left[G_j w_j + \frac{1}{2}(H_j + \lambda) w_j^2 \right] + rT$$

$$\frac{\partial \hat{Obj}(\theta)^t}{\partial w_j} = G_j + (H_j + \lambda) w_j = 0$$

$$w_j = -\frac{G_j}{H_j + \lambda}$$

- MSE :

$$g_i = \frac{\partial}{\partial \hat{y}^{(t-1)}} (y_i - \hat{y}^{(t-1)})^2$$

$$= -2(y_i - \hat{y}^{(t-1)})$$

$$h_i = \frac{\partial^2}{\partial^2 \hat{y}^{(t-1)}} (y_i - \hat{y}^{(t-1)})^2$$

$$= \frac{\partial}{\partial \hat{y}^{(t-1)}} -2y_i + 2\hat{y}^{(t-1)}$$

$$= 2$$

$$G_j = \Sigma_{i \in I_j} -2(y_i - \hat{y}^{(t-1)})$$

$$H_j = \Sigma_{i \in I_j} 2$$

$$w_j = - \frac{\Sigma_{i \in I_j} -2(y_i - \hat{y}^{(t-1)})}{\Sigma_{i \in I_j} 2 + \lambda}$$

- Binomial deviance :

$$\log\left(\frac{p}{1-p}\right) = \log(\text{odds}) = \hat{y}^{(t-1)}$$

$$p = \frac{e^{\hat{y}^{(t-1)}}}{1 + e^{\hat{y}^{(t-1)}}}$$

$$g_i = -2(y_i - p)$$

$$h_i = 2p(1-p)$$

$$G_j = \sum_{i \in I_j} -2(y_i - p)$$

$$H_j = \sum_{i \in I_j} 2p(1-p)$$

$$w_j = -\frac{\sum_{i \in I_j} -2(y_i - p)}{\sum_{i \in I_j} 2p(1-p) + \lambda} \quad \text{where } p = \frac{e^{\hat{y}^{(t-1)}}}{1 + e^{\hat{y}^{(t-1)}}}$$