Tsinghua-Berkeley Shenzhen Institute
INFERENCE AND INFORMATION
Fall 2017

Homework 3

赵丰 March 18, 2018

- **Acknowledgments:** This coursework refers to wikipedia: `https://en.wikipedia.org`.
- **Collaborators:** I finish this coursework by myself.

I use `enumerate` to generate answers for each question:

3.1. (a) $p_{\mathsf{H}}(H_0) = p, C_{ij} = \delta_{ij}$ 由似然比检验，对于给定的数据 y，当 $2y \geq \frac{p}{1-p}$ 拒绝 H_0; 当 $2y \leq \frac{p}{1-p}$ 时接受 H_0。由于 $y \in [0,1]$, 当 $p > \frac{2}{3}$ 时，$2y \leq \frac{p}{1-p}$ 恒成立，所以此时总是接受 H_0。

(b) 当 $p < \frac{2}{3}$ 时，

$$P_{\mathsf{D}} = \int_{\frac{p}{2(1-p)}}^{1} 2y dy = 1 - \left(\frac{p}{2(1-p)}\right)^2$$

$$P_{\mathsf{F}} = \int_{\frac{p}{2(1-p)}}^{1} dy = 1 - \frac{p}{2(1-p)}$$

$\Rightarrow$

$$P_{\mathsf{D}} = 2P_{\mathsf{F}} - P_{\mathsf{F}}^2, p \leq \frac{2}{3} \tag{1}$$

当 $p > \frac{2}{3}$ 时，$P_{\mathsf{D}} = P_{\mathsf{F}} = 0$。

(c) i. 令 $P_{\mathsf{D}} > P_{\mathsf{F}}$, 得 $P_F \leq 1 - \epsilon$, 将 $P_F = 1 - \epsilon$ 代入(1)中得 $P_{\mathsf{D}}^{\max}(\epsilon) = 1 - \epsilon^2$

ii. 令 $0 < P_F < 1$, 得 $0 < \epsilon < 1$。

iii. 令 $1 - \frac{p}{2} = P_F \leq 1 - \epsilon \Rightarrow p \geq \frac{2\epsilon}{1+2\epsilon}$。又 $p \leq 1 \Rightarrow p \in [\min\{\frac{2\epsilon}{1+2\epsilon}, \frac{2}{3}\}, 1]$

3.2. (a) i. 设 $\tilde{\varphi}(\mathsf{H}, H) = \mathbb{E}[C(\mathsf{H}, H)|\underline{\mathsf{y}} = \underline{y}]$ 当 $\hat{H}(\underline{y}) = H_1$ 时，

$$\begin{aligned}\tilde{\varphi}(H_1, H) =& C_{11}\mathbb{P}(\mathsf{H} = H_1|\underline{\mathsf{y}} = \underline{y}) + C_{12}\mathbb{P}(\mathsf{H} = H_2|\underline{\mathsf{y}} = \underline{y}) + C_{13}\mathbb{P}(\mathsf{H} = H_3|\underline{\mathsf{y}} = \underline{y}) \\ =& \mathbb{P}(\mathsf{H} = H_2|\underline{\mathsf{y}} = \underline{y}) + 2\mathbb{P}(\mathsf{H} = H_3|\underline{\mathsf{y}} = \underline{y})\end{aligned}$$

同理可求出:

$$\begin{aligned}\tilde{\varphi}(H_2, H) =& \mathbb{P}(\mathsf{H} = H_1|\underline{\mathsf{y}} = \underline{y}) + 2\mathbb{P}(\mathsf{H} = H_3|\underline{\mathsf{y}} = \underline{y}) \\ \tilde{\varphi}(H_3, H) =& 2\mathbb{P}(\mathsf{H} = H_1|\underline{\mathsf{y}} = \underline{y}) + 2\mathbb{P}(\mathsf{H} = H_2|\underline{\mathsf{y}} = \underline{y})\end{aligned}$$

设

$$j = \arg\min_{i=1,2,3} \tilde{\varphi}(H_i, H)$$

最优的决策准则为 $\hat{H}(\underline{\mathsf{y}}) = H_j$

ii. 利用概率的归一化条件有：

$$
\begin{aligned}
\tilde{\varphi}(H_1, H) =& \pi_2 + 2(1 - \pi_1 - \pi_2) \\
\tilde{\varphi}(H_2, H) =& \pi_1 + 2(1 - \pi_1 - \pi_2) \\
\tilde{\varphi}(H_3, H) =& 2\pi_1 + 2\pi_2
\end{aligned}
$$

根据 π_1, π_2 以及 $\pi_3 = 1 - \pi_1 - \pi_2$ 的非负性和 i. 中的结果，决策区域如图 1所示。

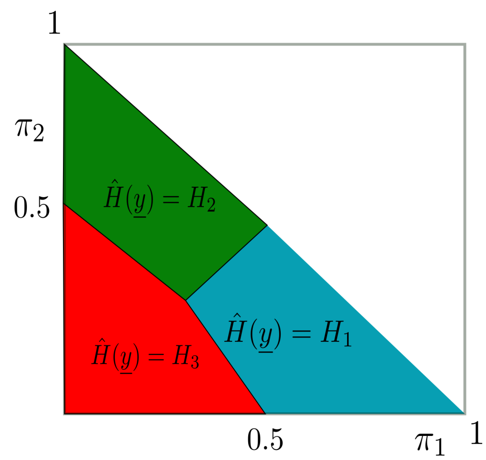

Figure 1: 决策区域

(b) 若 $L_i(\underline{\mathsf{y}}), i = 2, 3$ 已知，则

$$
\frac{p_{\underline{y}|\mathsf{H}}(\underline{y}|H_2)}{p_{\underline{y}|\mathsf{H}}(\underline{y}|H_3)} = \frac{L_2(\underline{y})}{L_3(\underline{y})}
$$

另一方面，因为 H_i 的先验均相等，基于后验概率的 $\underline{\pi}(\underline{y})$ 正比于

$$
\begin{bmatrix}
p_{\underline{y}|\mathsf{H}}(\underline{y}|H_1) \\
p_{\underline{y}|\mathsf{H}}(\underline{y}|H_2) \\
p_{\underline{y}|\mathsf{H}}(\underline{y}|H_3)
\end{bmatrix}
$$

所以最优决策准则可以用 $L_i(\underline{y}), i = 2, 3$ 给出。

$$
\begin{aligned}
p_{\underline{y}|\mathsf{H}}(\underline{y}|H_1) &= \frac{\exp(-\frac{(y_1-1)^2+y_2^2+y_3^2}{2\sigma^2})}{\sqrt{2\pi}^3\sigma^3} \\
p_{\underline{y}|\mathsf{H}}(\underline{y}|H_2) &= \frac{\exp(-\frac{y_1^2+(y_2-1)^2+y_3^2}{2\sigma^2})}{\sqrt{2\pi}^3\sigma^3} \\
p_{\underline{y}|\mathsf{H}}(\underline{y}|H_3) &= \frac{\exp(-\frac{y_1^2+y_2^2+(y_3-1)^2}{2\sigma^2})}{\sqrt{2\pi}^3\sigma^3}
\end{aligned}
$$

因此

$$L_2(\underline{y}) = \exp(-\frac{\ell_2(\underline{y})}{\sigma^2})$$
$$L_3(\underline{y}) = \exp(-\frac{\ell_3(\underline{y})}{\sigma^2})$$

3.3. (a) 根据 Bayes 风险最小的准则，$\hat{x} = -1$ 当后验概率满足

$$p_{x|\underline{\hat{x}}}(x = 1|\underline{\hat{x}}) \leq p_{x|\underline{\hat{x}}}(x = -1|\underline{\hat{x}})$$

否则取 $\hat{x} = 1$ 由 $P_{\mathsf{x}}(1) = P_{\mathsf{x}}(-1) = 1/2$, 得

$$p_{\underline{\hat{x}}|x}(\underline{\hat{x}}|x = 1) \leq p_{\underline{\hat{x}}|x}(\underline{\hat{x}}|x = -1)$$

记

$$\phi(k) = \begin{cases} \frac{3}{4} & k > 0 \\ \frac{1}{4} & k \leq 0 \end{cases}$$

则由独立性假设

$$\begin{aligned} p_{\underline{\hat{x}}|x}(\underline{\hat{x}}|x = 1) &= \prod_{i=1}^{n} p_{\hat{x_i}|x}(\hat{x_i}|x = 1) \\ &= \prod_{i=1}^{n} \phi(\hat{x_i}) \end{aligned}$$

同理

$$p_{\underline{\hat{x}}|x}(\underline{\hat{x}}|x = -1) = \prod_{i=1}^{n} \phi(-\hat{x_i})$$

因此最小错误概率的判决准则是：

$$\left(\frac{3}{4}\right)^{\sum_i 1_{\hat{x}_i=1}} \left(\frac{1}{4}\right)^{\sum_i 1_{\hat{x}_i=-1}} \underset{\hat{x}=-1}{\overset{\hat{x}=1}{\gtrless}} \left(\frac{3}{4}\right)^{\sum_i 1_{\hat{x}_i=-1}} \left(\frac{1}{4}\right)^{\sum_i 1_{\hat{x}_i=1}}$$

(b) 若 $\hat{x}_1(y_1) = 1(H_1)$, 期望损失为：

$$\begin{aligned} \tilde{\varphi}(H_1, y_1) =& C(1,1,1)p_{\mathsf{H},\hat{x}_2|\mathsf{y}}(\hat{x}_2 = 1, \mathsf{x} = 1|y_1) + C(1,1,-1)p_{\mathsf{H},\hat{x}_2|\mathsf{y}}(\hat{x}_2 = 1, \mathsf{x} = -1|y_1) \\ &+ C(1,-1,1)p_{\mathsf{H},\hat{x}_2|\mathsf{y}}(\hat{x}_2 = -1, \mathsf{x} = 1|y_1) + C(1,-1,-1)p_{\mathsf{H},\hat{x}_2|\mathsf{y}}(\hat{x}_2 = -1, \mathsf{x} = -1|y_1) \\ \Rightarrow \tilde{\varphi}(H_1, y_1)p_{\mathsf{y}_1}(y_1) =& C(1,1,1)p_{\mathsf{x}}(\mathsf{x} = 1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = 1|\mathsf{x} = 1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = 1) \\ &+ C(1,1,-1)p_{\mathsf{x}}(\mathsf{x} = -1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = 1|\mathsf{x} = -1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = -1) \\ &+ C(1,-1,1)p_{\mathsf{x}}(\mathsf{x} = 1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = -1|\mathsf{x} = 1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = 1) \\ &+ C(1,-1,-1)p_{\mathsf{x}}(\mathsf{x} = -1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = -1|\mathsf{x} = -1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = -1) \end{aligned}$$

同理给出 $\hat{x}_1(y_1) = -1(H_0)$, 期望损失 $\tilde{\varphi}(H_0, y_1)$ 满足：

$$\begin{aligned} \tilde{\varphi}(H_0, y_1)p_{\mathsf{y}_1}(y_1) =& C(-1,1,1)p_{\mathsf{x}}(\mathsf{x} = 1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = 1|\mathsf{x} = 1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = 1) \\ &+ C(-1,1,-1)p_{\mathsf{x}}(\mathsf{x} = -1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = 1|\mathsf{x} = -1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = -1) \\ &+ C(-1,-1,1)p_{\mathsf{x}}(\mathsf{x} = 1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = -1|\mathsf{x} = 1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = 1) \\ &+ C(-1,-1,-1)p_{\mathsf{x}}(\mathsf{x} = -1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2 = -1|\mathsf{x} = -1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x} = -1) \end{aligned}$$

当 $\tilde{\varphi}(H_0, y_1) \le \tilde{\varphi}(H_1, y_1)$ 时接受 H_0, 得
$\gamma_2 p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x}=1) \le \gamma_3 p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x}=-1)$ 其中

$$\begin{aligned}\gamma_2 =& C(-1,1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=1) + C(-1,-1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=1)\\ &-C(1,1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=1) - C(1,-1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=1)\end{aligned}$$

$$\begin{aligned}\gamma_3 =& C(-1,1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=-1)\\ &+C(-1,-1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=-1)\\ &-C(1,1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=-1) - C(1,-1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=-1)\end{aligned}$$

根据 $C(\hat{x}_1, \hat{x}_2, x)$ 的性质，
$C(-1,1,1) > C(1,1,1), C(-1,-1,1) > C(1,-1,1)$，所以 $\gamma_2 > 0$, 因此

$$\frac{p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x}=1)}{p_{\mathsf{y}_1|\mathsf{x}}(y_1|\mathsf{x}=-1)} \le \frac{\gamma_3}{\gamma_2} \triangleq \gamma_1$$

另解：

$$\begin{aligned}\bar{R} =& \sum_{x_1,x_2,\tilde{x}=-1,1} C(x_1,x_2,\tilde{x})P(\hat{\mathsf{x}}_1=x_1,\hat{\mathsf{x}}_2=x_2,\mathsf{x}=\tilde{x})\\ =& \sum_{x_1,x_2,\tilde{x}=-1,1} C(x_1,x_2,\tilde{x})P_{\mathsf{x}}(\tilde{x})P_{\hat{\mathsf{x}}_1|\mathsf{x}}(x_1|\tilde{x})P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|\tilde{x})\\ =& \sum_{x_2,\tilde{x}=-1,1} C(1,x_2,\tilde{x})P_{\mathsf{x}}(\tilde{x})P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|\tilde{x})\int_{y_1\in\mathcal{Y}_1} p_{\mathsf{y}_1|\mathsf{x}}(y_1|\tilde{x})dy_1\\ &+ \sum_{x_2,\tilde{x}=-1,1} C(-1,x_2,\tilde{x})P_{\mathsf{x}}(\tilde{x})P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|\tilde{x})\int_{y_1\in\mathcal{Y}_0} p_{\mathsf{y}_1|\mathsf{x}}(y_1|\tilde{x})dy_1\\ =& \sum_{x_2,\tilde{x}=-1,1} C(-1,x_2,\tilde{x})P_{\mathsf{x}}(\tilde{x})P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|\tilde{x}), \mathcal{Y}_1\cup\mathcal{Y}_0=\mathbb{R}\\ &+\int_{y_1\in\mathcal{Y}_1}\sum_{x_2,\tilde{x}=-1,1}\left(C(1,x_2,\tilde{x})-C(-1,x_2,\tilde{x})\right)P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|\tilde{x})P_{\mathsf{x}}(\tilde{x})p_{\mathsf{y}_1|\mathsf{x}}(y_1|\tilde{x})dy_1\end{aligned}$$

因此

$$\begin{aligned}y_1\in\mathcal{Y}_1 \iff& \sum_{x_2,\tilde{x}=-1,1}\left(C(1,x_2,\tilde{x})-C(-1,x_2,\tilde{x})\right)P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|\tilde{x})P_{\mathsf{x}}(\tilde{x})p_{\mathsf{y}_1|\mathsf{x}}(y_1|\tilde{x}) < 0\\ \iff& \sum_{x_2=-1,1}\left(C(1,x_2,1)-C(-1,x_2,1)\right)P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|1)P_{\mathsf{x}}(1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|1)\\ &< \sum_{x_2=-1,1}\left(C(-1,x_2,-1)-C(1,x_2,-1)\right)P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|-1)P_{\mathsf{x}}(-1)p_{\mathsf{y}_1|\mathsf{x}}(y_1|-1)\\ \iff& \frac{p_{\mathsf{y}_1|\mathsf{x}}(y_1|1)}{p_{\mathsf{y}_1|\mathsf{x}}(y_1|-1)} > \gamma_1\\ \gamma_1 =& \frac{P_{\mathsf{x}}(-1)\sum_{x_2=-1,1}\left(C(-1,x_2,-1)-C(1,x_2,-1)\right)P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|-1)}{P_{\mathsf{x}}(1)\sum_{x_2=-1,1}\left(C(1,x_2,1)-C(-1,x_2,1)\right)P_{\hat{\mathsf{x}}_2|\mathsf{x}}(x_2|1)}\end{aligned}$$

(c) 由对称性，根据 (b) 中的结果，可得最优的决策准则 $\hat{x}_2^*(\dot{})$ 是如下形式的似然比检验：

$$\frac{p_{\mathsf{y}_2|\mathsf{x}}(y_2|\mathsf{x}=1)}{p_{\mathsf{y}_2|\mathsf{x}}(y_2|\mathsf{x}=-1)} \underset{\hat{x}_2(y_1)=1}{\overset{\hat{x}_2(y_1)=-1}{\gtrless}} \frac{\gamma_2'}{\gamma_3'} \triangleq \gamma_1'$$

其中

$$\begin{aligned}\gamma_2' =& C(1,-1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=1|\mathsf{x}=1)+C(-1,-1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=-1|\mathsf{x}=1)\\ &-C(1,1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=1|\mathsf{x}=1)-C(-1,1,1)p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=-1|\mathsf{x}=1)\end{aligned}$$

$$\begin{aligned}\gamma_3' =& C(1,-1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=1|\mathsf{x}=-1)\\ &+C(-1,-1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=-1|\mathsf{x}=-1)\\ &-C(1,1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=1|\mathsf{x}=-1)-C(-1,1,-1)p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_1|\mathsf{x}}(\hat{x}_1=-1|\mathsf{x}=-1)\end{aligned}$$

(d) (b) 中 γ_2, γ_3 可化简为：

$$\begin{aligned}\gamma_2 =& p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=1)+Lp_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=1)\\ &-p_{\mathsf{x}}(\mathsf{x}=1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=1)\\ =& p_{\mathsf{x}}(\mathsf{x}=1)(1+(L-2)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=1))\end{aligned}$$

$$\begin{aligned}\gamma_3 =& p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=-1)\\ &-Lp_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=-1)-p_{\mathsf{x}}(\mathsf{x}=-1)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=-1|\mathsf{x}=-1)\\ =& -p_{\mathsf{x}}(\mathsf{x}=-1)(1+(L-2)p_{\hat{x}_2|\mathsf{x}}(\hat{x}_2=1|\mathsf{x}=-1))\end{aligned}$$

因此，当 $L=2$ 时，$\gamma=-\frac{p_{\mathsf{x}}(\mathsf{x}=1)}{p_{\mathsf{x}}(\mathsf{x}=-1)}$, 与 $\hat{x}_2$ 无关; 同理求出 $\gamma'=\gamma$.

3.4. 上机作业

3.5. Thanks to 陆石, who gives me this template.