CS 559 HW1

Chenxu Wang
10457625

problem1

Since $P(A|B) = \frac{P(AB)}{P(B)}$, $P(B|A) = \frac{P(AB)}{P(A)}$

In general, $P(A) \neq P(B)$

$\therefore$ $P(A|B)$ & $P(B|A)$ are not the same.

Example: $A = \{$ The team wins in the game $\}$

$B = \{$ The player Frank goals in the game $\}$. Suppose $P(A) = 0.6$ $P(B) = 0.5$, and when The Team won the game, Probability that Frank have a goal is 0.7.

$\because P(AB) = P(B|A) \cdot P(A) = 0.7 \times 0.6 = 0.42$

$\therefore P(A|B) = \frac{P(AB)}{P(B)} = \frac{0.42}{0.5} = 0.84$.

$\therefore$ $P(A|B) \neq P(B|A)$

problem 2.

2.1

$$\text{covariance}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$$

$$= E\{XY - XE(Y) - E(X)Y + E(X)E(Y)\}$$

$$= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$

$$= E(XY) - E(X)E(Y)$$

Since X & Y are independent.

$$\therefore E(XY) = E(X)E(Y)$$

$\therefore$ $\text{covariance}(X, Y) = 0$

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sigma X \cdot \sigma Y} = 0$$

So, X, Y are uncorrelated.

2.2 Suppose

X	-1	0	1
P	$1/4$	$1/2$	$1/4$

$y = x^2$, y is dependent on x.

$E(X) = -1 \times 1/4 + 0 \times \frac{1}{2} + 1 \times 1/4 = 0$

$E(X^3) = (-1)^3 \cdot 1/4 + 0^3 \times \frac{1}{2} + 1^3 \times \frac{1}{4} = 0$

$\mathrm{cov}(X, Y) = E(X - \bar{X}) \cdot (Y - \bar{Y})$

$= E(XY) - E(X) \cdot E(Y)$

$= E(X^3) - E(X) \cdot E(Y) = 0$

$\therefore \rho_{XY} = 0$

$\therefore$ X & Y are uncorrelated.

however, X^2 is dependent on X.

Problem 3.

Consider to chuose discriminant funcition:

$$g_j(x) = \ln P(x|w_j) + \ln P(w_j)$$

Since x_i is statistically independent for all x in w_j

$$\therefore \quad P(x|w_j) = \prod_{i=1}^{d} P(x_i|w_j)$$

$$= \prod_{i=1}^{d} P_{ij}^{x_i} \cdot (1-P_{ij})^{1-x_i}$$

Then, $g_j(x) = \ln \prod_{i=1}^{d} P_{ij}^{x_i} \cdot (1-P_{ij})^{1-x_i} + \ln P(w_j)$

$$= \sum_{i=1}^{d} \left[x_i \cdot \ln P_{ij} + (1-x_i) \ln (1-P_{ij}) \right] + \ln P(w_j)$$

$$= \sum_{i=1}^{d} \left[x_i \cdot \ln P_{ij} + \ln(1-P_{ij}) - x_i \ln(1-P_{ij}) \right] + \ln P(w_j)$$

$$= \sum_{i=1}^{d} \left[x_i \cdot \ln \frac{P_{ij}}{1-P_{ij}} + \ln(1-P_{ij}) \right] + \ln P(w_j)$$

$$= \sum_{i=1}^{d} x_i \cdot \ln \frac{P_{ij}}{1-P_{ij}} + \sum_{i=1}^{d} \ln(1-P_{ij}) + \ln P(w_j)$$

Problem 4.

Given $P(x|w_1) = N(4,1)$, $P(x|w_2) = N(8,1)$

$P(w_2) = \frac{1}{4}$, $P(w_1) = 1 - P(w_2) = \frac{3}{4}$

$$\lambda = \begin{bmatrix} 0_{11} & 1_{12} \\ 3_{21} & 0_{22} \end{bmatrix}$$

$R(\alpha_1|x) = \lambda_{11}P(w_1|x) + \lambda_{12}P(w_2|x)$

$R(\alpha_2|x) = \lambda_{21}P(w_1|x) + \lambda_{22}P(w_2|x)$

Decision Rule: decide w_1 if $R(\alpha_1|x) < R(\alpha_2|x)$

$\therefore (\lambda_{21} - \lambda_{11})P(w_1|x) > (\lambda_{12} - \lambda_{22})P(w_2|x)$

$(3-0)P(w_1|x) > (1-0)P(w_2|x)$

$$\frac{P(w_1|x)}{P(w_2|x)} > 1$$

$$P(x|w_1) = \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-(x-4)^2}{2}}$$

$$P(x|w_2) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-8)^2}{2}}$$

$$\Rightarrow \quad \frac{e^{-\frac{(x-4)^2}{2}}}{e^{-\frac{(x-8)^2}{2}}} > 1$$

$$\Rightarrow \quad e^{-\frac{(x-4)^2}{2}} > e^{-\frac{(x-8)^2}{2}}$$

Take logarithm on both sides:

$$-\frac{(x-4)^2}{2} > -\frac{(x-8)^2}{2}$$

$$x < 6$$

$\therefore$ when $x < 6$, we will decide to choose w_1.

Problem 5

5.1

when $i = 1, \dots c$,

$$R(\alpha_i | x) = \sum_{j=1}^{c} \lambda(\alpha_i | \omega_j) P(\omega_j | x)$$

$$= \lambda_s \sum_{j=1, j \neq i}^{c} P(\omega_j | x)$$

$$= \lambda_s [1 - P(\omega_i | x)]$$

when $i = c+1$,

$$R(\alpha_{c+1} | x) = \lambda_r$$

So, when we decide to chose ω_i

$$R(\alpha_i | x) \leq R(\alpha_{c+1} | x)$$

$$\therefore \lambda_s [1 - P(\omega_i | x)] \leq \lambda_r$$

$$\Rightarrow P(\omega_i | x) \geq 1 - \frac{\lambda_r}{\lambda_s}$$, and others will be rejected.

5.2. when $\lambda_r = 0$, then $P(\omega_i | x) \geq 1$

$\therefore$ it will always reject.

5.3. when $\lambda_r > \lambda_s$, $P(\omega_i | x) \geq b$ ($b < 0$).

then it will not reject anyone.

Problem 6.

6.1 Given $p(x|\eta) = h(x) \cdot e^{[\eta^T T(x) - A(\eta)]}$

Integrate both sides:

$$\int p(x|\eta) = \int h(x)\, e^{\eta^T \cdot T(x)} \cdot e^{-A(\eta)}\, dx$$

Since the integral of probability density equals to 1.

$$1 = \int h(x) \cdot e^{\eta^T T(x)} \cdot e^{-A(\eta)}\, dx$$

Multiply $e^{A(\eta)}$ on the both sides:

$$e^{A(\eta)} = \int h(x)\, e^{\eta^T T(x)}\, dx$$

Take logarithm of both sides:

$$A(\eta) = \ln\left[\int h(x)\, e^{\eta^T T(x)}\, dx\right]$$

6.2 $$\frac{\partial}{\partial \eta} A(\eta) = \frac{\partial}{\partial \eta} \ln \int (h(x)\, e^{\eta^T \cdot T(x)}\, dx)$$

$$= \frac{\int (h(x) e^{\eta^T T(x)} \cdot T(x)\, dx)}{\int h(x) \cdot e^{\eta^T \cdot T(x)}\, dx}$$

$$= \frac{\int h(x) \cdot e^{\eta^T T(x)} \cdot T(x)\, dx}{e^{A(\eta)}}$$

$$= \int h(x)\, e^{\eta^T \cdot T(x) - A(\eta)} \cdot T(x)\, dx$$

$$= \int p(x \mid \eta) \cdot T(x)\, dx = E_\eta[T(x)]$$

6.3

The probability density:

$$p(x|\eta) = h(x)\, e^{\eta^T T(x) - A(\eta)}$$

The likelihood function is:

$$L(\eta | x_1 \cdots x_n) = \prod_{i=1}^{n} h(x_i) \cdot e^{\eta^T T(x_i) - A(\eta)}$$

$$= \prod_{i=1}^{n} h(x_i) \cdot e^{\eta^T \sum_{i=1}^{n} T(x_i) - nA(\eta)}$$

Take log of both sides:

$$\ln L(\eta | x_1 \cdots x_n) = \ln \sum_{i=1}^{n} h(x_i) + \eta^T \sum_{i=1}^{n} T(x_i) - n \cdot A(\eta)$$

Lets differentiate $\ln L(\eta | x_1 \cdots x_n)$ w.r.t η.

$$\frac{\partial \ln L(\eta | x_1 \cdots x_n)}{\partial \eta} = 0 + 1 \cdot \sum_{i=1}^{n} T(x_i) - n \frac{\partial A(\eta)}{\partial \eta} = 0$$

$$\therefore \frac{\partial A(\eta)}{\partial \eta} = \frac{\sum_{i=1}^{n} T(x_i)}{n}$$

$$\therefore E_\eta(T(x)) = \frac{\sum_{i=1}^{n} T(x_i)}{n}$$

Problem 7.

7.1.

$$LL(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y \cdot \log H_\theta(x) + (1-y) \cdot \log(1 - H_\theta(x))\right], \quad y \in \{0, 1\}$$

$$H_\theta(x) = \frac{1}{1 + e^{-\theta^T \cdot x}}$$

Let, $LL(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} k(\theta)\right]$

$$k(\theta) = y \cdot \log H_\theta(x) + (1-y) \cdot \log(1 - H_\theta(x))$$

$$\therefore \frac{\partial LL(\theta)}{\partial \theta} = -\frac{1}{m} \cdot \frac{\partial \sum_{i=1}^{m} k(\theta)}{\partial \theta}$$

$$k'(\theta) = \frac{y \cdot H_\theta'(x)}{H_\theta(x)} + (1-y)\frac{[1 - H_\theta(x)]'}{1 - H_\theta(x)}$$

$$H_\theta'(x) = H_\theta(x) \cdot (1 - H_\theta(x)) \cdot (\theta^T \cdot x)'$$

$$[1 - H_\theta'(x)]' = -H_\theta(x)(\theta^T \cdot x)'$$

$$\therefore \frac{LL(\theta)}{\partial \theta} = \frac{1}{m}\left[\sum_{i=1}^{m} (H_\theta(x) - y) \cdot (\theta^T \cdot x)'\right]$$