資料科學 HW 2 Student ID: P10942A08 Name: 賴之浩.

Problem 1: Type I and II errors

(1) Type I error:

$\alpha = P(\text{Type I error})$

$= P(\text{Reject } H_0 / H_0 \text{ is true})$

$= P(\sum_{i=1}^{5} X_i = 5 / H_0 \text{ is true})$

$= P(\sum_{i=1}^{5} X_i = 5 / H_0 : p = \frac{1}{2})$

$= C^5_5 (\frac{1}{2})^5 (1-\frac{1}{2})^{5-5}$

$= \frac{1}{32} = 0.03125$

$\therefore$ The probability of Type I error is 0.03125

(2) Type II error

$\beta = P(\text{Type II error})$

$= P(\text{Accept } H_0 / H_a \text{ is true})$

$= P(\sum_{i=1}^{5} X_i < 5 / H_a \text{ is true})$

$= P(\sum_{i=1}^{5} X_i \le 4 / H_a : p = \frac{3}{4})$

$= \sum_{k=0}^{4} \binom{5}{k} (\frac{3}{4})^k (1-\frac{3}{4})^{5-k}$

$= (\frac{1}{4})^5 + 5 \times \frac{3}{4} \times (\frac{1}{4})^4 + 10 \times (\frac{3}{4})^2 (\frac{1}{4})^3 + 10 \times (\frac{3}{4})^3 \times (\frac{1}{4})^2 + 5 \times (\frac{3}{4})^4 \times (\frac{1}{4})$

$= \frac{1}{1024} + \frac{15}{1024} + \frac{90}{1024} + \frac{270}{1024} + \frac{405}{1024} = \frac{781}{1024} = 0.7627$

$\therefore$ The probability of Type II error is 0.7627.

Problem 2: Hypothesis testing

$H_0: \sigma^2 = 0.81$, $H_a: \sigma^2 > 0.81$

when $S = 1.2 \Rightarrow S^2 = 1.44$ and $n = 10$

the test statistic:

$$X^2 = \frac{(n-1)S^2}{\sigma} = \frac{9 \times 1.44}{0.81} = 16$$

from the Chi-Square Distribution table,

the degree of freedom $= (10-1) = 9$, and $\alpha = 0.05$.

the null hypothesis is rejected when $X^2 > 16.919$

$\because$ The X^2 statistic is not significant at the 0.05 level,

$\therefore$ that there is insufficient evidence to believe that $\sigma > 0.9$ year.

Problem 3: Markov Properties

(1) What are the pairwise, local and global Markov properties for Markov random fields?

Pairwise Markov property: Any two non-adjacent variables are conditionally independent given all other variables: $X_u \perp\!\!\!\perp X_v \mid X_{V \setminus \{u,v\}}$

Local Markov property: A variable is conditionally independent of all other variables given its neighbors: $X_v \perp\!\!\!\perp X_{V \setminus N'[v]} \mid X_{N(v)}$ where $N(v)$ is the set of neighbors of v, $N'[v] = v \cup N(v)$ is the closed neighbourhood of v.

Global Markov property: Any two subsets of variables are conditionally independent given a separating subset: $X_A \perp\!\!\!\perp X_B \mid X_S$ where every path from a node in A to a node in B passes through S.

(2) What is the local Markov properties for Bayesian network?

Each variable is conditionally independent of its non-descendants given its parent variables: $X_v \perp\!\!\!\perp X_{V \setminus de(v)} \mid X_{pa(v)}$ for all $v \in V$ where $de(v)$ denotes the set of descendants of v (thus $V \setminus de(v)$ is the set of non-descendants of v)

(3) What is a Markov blanket?

The Markov blanket of a node is the set of nodes consisting of its parents, its children, and any other parents of its children.

Problem 4: LDA.

Samples for class 1: $X_1 = (x_1, x_2) = \{(5,3), (3,5), (3,4), (4,5), (4,7), (5,6)\}$

Samples for class 2: $X_2 = (x_1, x_2) = \{(9,10), (7,7), (8,5), (8,8), (7,2), (10,8)\}$

The classes mean are:

$\mu_1 = \frac{1}{N_1} \sum_{x \in X_1} x = \frac{1}{6} [\binom{5}{3} + \binom{3}{5} + \binom{3}{4} + \binom{4}{5} + \binom{4}{7} + \binom{5}{6}] = \frac{1}{6} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$

$\mu_2 = \frac{1}{N_2} \sum_{x \in X_2} x = \frac{1}{6} [\binom{9}{10} + \binom{7}{7} + \binom{8}{5} + \binom{8}{8} + \binom{7}{2} + \binom{10}{8}] = \frac{1}{6} \begin{bmatrix} 49 \\ 40 \end{bmatrix} = \begin{bmatrix} \frac{49}{6} \\ \frac{20}{3} \end{bmatrix}$

Covariance matrix of the first class:

$S_1 = \frac{1}{n-1} \sum_{x \in X_1} (x-\mu_1)(x-\mu_1)^T = \frac{1}{5} (\begin{bmatrix} 5-4 \\ 3-5 \end{bmatrix} \begin{bmatrix} 5-4 & 3-5 \end{bmatrix} + \begin{bmatrix} 3-4 \\ 5-5 \end{bmatrix} \begin{bmatrix} 3-4 & 5-5 \end{bmatrix} +$

$\begin{bmatrix} 3-4 \\ 4-5 \end{bmatrix} \begin{bmatrix} 3-4 & 4-5 \end{bmatrix} + \begin{bmatrix} 4-4 \\ 5-5 \end{bmatrix} \begin{bmatrix} 4-4 & 5-5 \end{bmatrix} +$

$\begin{bmatrix} 4-4 \\ 7-5 \end{bmatrix} \begin{bmatrix} 4-4 & 7-5 \end{bmatrix} + \begin{bmatrix} 5-4 \\ 6-5 \end{bmatrix} \begin{bmatrix} 5-4 & 6-5 \end{bmatrix})$

$= \frac{1}{5} (\begin{bmatrix} 1 \\ -2 \end{bmatrix} \begin{bmatrix} 1 & -2 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 & -1 \end{bmatrix} +$

$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix})$

$= \frac{1}{5} (\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix})$

$= \frac{1}{5} \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix} = \begin{bmatrix} 0.8 & 0 \\ 0 & 2 \end{bmatrix}$

$$S_2 = \frac{1}{n-1}\sum_{x\in X_2}(x-\mu_2)(x-\mu_2)^T = \frac{1}{5}\left(\begin{bmatrix} 9-\frac{49}{6} \\ 10-\frac{20}{3}\end{bmatrix}\begin{bmatrix} 9-\frac{49}{6} & 10-\frac{20}{3}\end{bmatrix} + \begin{bmatrix} 7-\frac{49}{6} \\ 7-\frac{20}{3}\end{bmatrix}\begin{bmatrix} 7-\frac{49}{6} & 7-\frac{20}{3}\end{bmatrix} +\right.$$

$$\begin{bmatrix} 8-\frac{49}{6} \\ 5-\frac{20}{3}\end{bmatrix}\begin{bmatrix} 8-\frac{49}{6} & 5-\frac{20}{3}\end{bmatrix} + \begin{bmatrix} 8-\frac{49}{6} \\ 8-\frac{20}{3}\end{bmatrix}\begin{bmatrix} 8-\frac{49}{6} & 8-\frac{20}{3}\end{bmatrix} +$$

$$\left.\begin{bmatrix} 7-\frac{49}{6} \\ 2-\frac{20}{3}\end{bmatrix}\begin{bmatrix} 7-\frac{49}{6} & 2-\frac{20}{3}\end{bmatrix} + \begin{bmatrix} 10-\frac{49}{6} \\ 8-\frac{20}{3}\end{bmatrix}\begin{bmatrix} 10-\frac{49}{6} & 8-\frac{20}{3}\end{bmatrix}\right)$$

$$= \frac{1}{5}\left(\begin{bmatrix} \frac{5}{6} \\ \frac{10}{3}\end{bmatrix}\begin{bmatrix} \frac{5}{6} & \frac{10}{3}\end{bmatrix} + \begin{bmatrix} -\frac{7}{6} \\ \frac{1}{3}\end{bmatrix}\begin{bmatrix} -\frac{7}{6} & \frac{1}{3}\end{bmatrix} + \begin{bmatrix} -\frac{1}{6} \\ -\frac{5}{3}\end{bmatrix}\begin{bmatrix} -\frac{1}{6} & -\frac{5}{3}\end{bmatrix} +\right.$$

$$\left.\begin{bmatrix} -\frac{1}{6} \\ \frac{4}{3}\end{bmatrix}\begin{bmatrix} -\frac{1}{6} & \frac{4}{3}\end{bmatrix} + \begin{bmatrix} -\frac{7}{6} \\ -\frac{14}{3}\end{bmatrix}\begin{bmatrix} -\frac{7}{6} & -\frac{14}{3}\end{bmatrix} + \begin{bmatrix} \frac{11}{6} \\ \frac{4}{3}\end{bmatrix}\begin{bmatrix} \frac{11}{6} & \frac{4}{3}\end{bmatrix}\right)$$

$$= \frac{1}{5}\left(\begin{bmatrix} \frac{25}{36} & \frac{25}{9} \\ \frac{25}{9} & \frac{100}{9}\end{bmatrix} + \begin{bmatrix} \frac{49}{36} & -\frac{7}{18} \\ -\frac{7}{18} & \frac{1}{9}\end{bmatrix} + \begin{bmatrix} \frac{1}{36} & \frac{5}{18} \\ \frac{5}{18} & \frac{25}{9}\end{bmatrix} + \begin{bmatrix} \frac{1}{36} & -\frac{2}{9} \\ -\frac{2}{9} & \frac{16}{9}\end{bmatrix} +\right.$$

$$\left.\begin{bmatrix} \frac{49}{36} & \frac{49}{9} \\ \frac{49}{9} & \frac{196}{9}\end{bmatrix} + \begin{bmatrix} \frac{121}{36} & \frac{22}{9} \\ \frac{22}{9} & \frac{16}{9}\end{bmatrix}\right)$$

$$= \frac{1}{5}\begin{bmatrix} \frac{246}{36} & \frac{186}{18} \\ \frac{186}{18} & \frac{354}{9}\end{bmatrix} = \begin{bmatrix} 1.3667 & 2.0667 \\ 2.0667 & 7.8667\end{bmatrix}$$

Within-class scatter matrix:

$$S_W = S_1 + S_2 = \begin{bmatrix} 0.8 & 0 \\ 0 & 2\end{bmatrix} + \begin{bmatrix} 1.3667 & 2.0667 \\ 2.0667 & 7.8667\end{bmatrix} = \begin{bmatrix} 2.1667 & 2.0667 \\ 2.0667 & 9.8667\end{bmatrix}$$

Between-class scatter matrix:

$$S_B = (\mu_1-\mu_2)(\mu_1-\mu_2)^T = \left(\begin{bmatrix} 4 \\ 5\end{bmatrix} - \begin{bmatrix} \frac{49}{6} \\ \frac{10}{3}\end{bmatrix}\right)\left(\begin{bmatrix} 4 & 5\end{bmatrix} - \begin{bmatrix} \frac{49}{6} & \frac{20}{3}\end{bmatrix}\right) = \begin{bmatrix} -\frac{25}{6} \\ -\frac{5}{3}\end{bmatrix}\begin{bmatrix} -\frac{25}{6} & -\frac{5}{3}\end{bmatrix}$$

$$= \begin{bmatrix} \frac{625}{36} & \frac{125}{18} \\ \frac{125}{18} & \frac{25}{9}\end{bmatrix} = \begin{bmatrix} 17.3611 & 6.9444 \\ 6.9444 & 2.7778\end{bmatrix}$$

The LDA projection is then obtained as the solution of the generalized eigenvalue problem:

$$S_W^{-1} S_B w = \lambda w$$

$$\Rightarrow |S_W^{-1} S_B - \lambda I| = 0$$

$$\Rightarrow \left|\begin{bmatrix} 2.1667 & 2.0667 \\ 2.0667 & 9.8667\end{bmatrix}^{-1}\begin{bmatrix} 17.3611 & 6.9444 \\ 6.9444 & 2.7778\end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1\end{bmatrix}\right| = 0$$

$$\Rightarrow \left|\begin{bmatrix} 0.5768 & -0.1208 \\ -0.1208 & 0.1267\end{bmatrix}\begin{bmatrix} 17.3611 & 6.9444 \\ 6.9444 & 2.7778\end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1\end{bmatrix}\right| = 0$$

$$\Rightarrow \left|\begin{bmatrix} 9.175-\lambda & 3.67 \\ -1.2174 & -0.4869-\lambda\end{bmatrix}\right| = (9.175-\lambda)(-0.4869-\lambda) - (-1.2174)\times 3.67 = 0$$

$$\Rightarrow \lambda^2 - 8.6881\lambda = 0 \Rightarrow \lambda_1 = 0,\ \lambda_2 = 8.6881$$

Hence: $\begin{bmatrix} 9.175 & 3.67 \\ -1.2174 & -0.4869\end{bmatrix} W_1 = 0\begin{bmatrix} W_1 \\ W_2\end{bmatrix}$ and $\begin{bmatrix} 9.175 & 3.67 \\ -1.2174 & -0.4869\end{bmatrix} W_2 = 8.6881\begin{bmatrix} W_1 \\ W_2\end{bmatrix}$

thus: $W_2 = \begin{bmatrix} -2.202 \\ 0.2921\end{bmatrix}$

normalize $W_2 = \begin{bmatrix} \frac{-2.202}{\sqrt{(-2.202)^2+(0.2921)^2}} \\ \frac{0.2921}{\sqrt{(-2.202)^2+(0.2921)^2}}\end{bmatrix} = \begin{bmatrix} -0.9913 \\ 0.1315\end{bmatrix}$

$\therefore$ The optimal projection vector (normalized to unit length) is $\begin{bmatrix} -0.9913 \\ 0.1315\end{bmatrix}$

and its corresponding eigenvalue is 8.6881.