Motivation/Data

Based on 8 elderly controls (EC) and 7 mild cognition impairment (MCI) subjects, we perform the logistic discriminant analysis (LDA) on cortical thickness and cortical surface area to see if these two imaging biomarkers can be used to discriminate MCI from EC.

Data source: Sterling Johnson
Previous work on AD and cortical thickness

• Cortical thickness has been shown to characterize cortical atrophy in AD patients quite well (Lerch et al., NeuroImage, 2005).

• Question: It is unclear if cortical thickness will be an important biomarkers of discriminating EC vs. MCI.
<table>
<thead>
<tr>
<th>Group</th>
<th>gender</th>
<th>age</th>
<th>education</th>
<th>memory</th>
<th>thickness</th>
<th>surface area</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>63</td>
<td>16</td>
<td>47</td>
<td>2.7718</td>
<td>15.4768</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>66</td>
<td>16</td>
<td>48</td>
<td>2.7132</td>
<td>13.953</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>70</td>
<td>18</td>
<td>43</td>
<td>2.8094</td>
<td>13.5577</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>75</td>
<td>20</td>
<td>48</td>
<td>2.7371</td>
<td>13.821</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>75</td>
<td>16</td>
<td>59</td>
<td>2.6476</td>
<td>14.0353</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>64</td>
<td>13</td>
<td>42</td>
<td>2.8913</td>
<td>14.354</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>69</td>
<td>20</td>
<td>56</td>
<td>2.8284</td>
<td>13.3125</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>81</td>
<td>16</td>
<td>41</td>
<td>2.6412</td>
<td>14.463</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>75</td>
<td>12</td>
<td>29</td>
<td>2.7204</td>
<td>13.628</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>62</td>
<td>20</td>
<td>39</td>
<td>2.6992</td>
<td>13.6875</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>68</td>
<td>17</td>
<td>34</td>
<td>2.7632</td>
<td>13.823</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>77</td>
<td>14</td>
<td>26</td>
<td>2.6106</td>
<td>13.1918</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>80</td>
<td>18</td>
<td>28</td>
<td>2.4556</td>
<td>12.8165</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>78</td>
<td>20</td>
<td>37</td>
<td>2.6012</td>
<td>13.164</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>64</td>
<td>18</td>
<td>31</td>
<td>2.8094</td>
<td>13.1511</td>
</tr>
</tbody>
</table>
Data: Group, Thickness, Area
Response variable: discrete

• Group variable

\[ Y_i \sim \text{Bernoulli}(\pi_i) \]

• \( Y = 0 \) (EC), 1 (MCI)

\[ \pi_i = \bar{P}(Y_i = 1) \]
Inconsistency of liner model

\[ Y_i = x'_i \beta + \epsilon_i, \]

\[ x'_i = (1, x_{i1}, \cdots, x_{ip}) \quad \beta' = (\beta_0, \cdots, \beta_p) \]

Inconsistency:

\[ \mathbb{E}Y_j = \pi_i = x'_i \beta \]

The expectation has to be linked to predictors somehow
General setting for GLM

\[ g(\mathbf{E}y) = z \cdot \lambda + x \cdot \beta \]

\( g \) is a link function

\[ y = g^{-1}(z \cdot \lambda + x \cdot \beta) + \varepsilon \]
Logistic regression function

\[ \pi_i = g(x_i) = \frac{\exp(x_i' \beta_i)}{1 + \exp(x_i' \beta_i)} \]

Logit function - this is the link function we will be using

\[ \text{logit}(\pi_i) = \log \frac{\pi_i}{1 - \pi_i} = x_i' \beta_i \]
Likelihood function

\[ L(\beta|y_1, \cdots, y_n) = \prod_{i=1}^{n} \pi_i^{y_i} (1 - \pi_i)^{1-y_i} = \prod_{i=1}^{n} \left[ \frac{\exp(x_i'\beta_i)}{1 + \exp(x_i'\beta_i)} \right]^{y_i} \prod_{i=1}^{n} \left[ \frac{1}{1 + \exp(x_i'\beta)} \right]^{1-y_i} \]

The parameters are estimated by maximizing the likelihood function. See /theories for detail.
MATLAB command for doing GLM with logit link

>> beta=glmfit([area thick], [group ones(n,1)], 'binomial')

beta =
  188.7976
   -7.8718
  -29.6755

G(group) = 188.7976 –7.8718*area –29.6755*thick
$$>> \ beta=g \ lmfit([\text{education memory}], [\text{group ones(n,1)}], \ 'binomial')$$

\[\beta = \begin{array}{l}
55.8206 \\
3.6825 \\
-3.0639 \\
\end{array}\]

\[G(\text{group}) = 55.8206 + 3.6825 \times \text{education} - 3.0639 \times \text{memory}\]
Inference on parameters

\[ \logit(\pi_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p. \]

\[ \beta^{(1)} = (\beta_0, \cdots, \beta_q)', \quad \beta^{(2)} = (\beta_{q+1}, \cdots, \beta_p)'. \]

\[ H_0 : \beta^{(2)} = 0. \]

Loglikelihood ratio test

\[ 2[\log L(\hat{\pi}^{(p)}) - \log L(\hat{\pi}^{(q)})] = D_q - D_p \sim \chi^2_{p-q}. \]
Logistic discriminant

Classification rule (decision rule)

\[ P(Y_i = 1) > P(Y_i = 0) \]

\[ \pi_i > 1/2 \quad \text{---} \rightarrow \text{group 1} \]

\[ x_{i,1}^T \beta_i > 0 \]
Performance of classifier

• Error rate = overall probability of making wrong decision

• Simplest way to estimate it is to use the concept of cross-validation.

• Leave one out cross-validation
Correct decision

Incorrect decision
Classification boundary

Error rate: thickness --> 47%
Thickness + area --> 20%
area --> 20%
Correct diagnostic rate (1-error rate) computed at each point on surface
Lecture 9 Topics

Gaussian Kernel Smoothing