1. Assume a disease so rare that it is seen in only one person out of every million. Assume also that we have a test that is effective in that if a person has the disease, there is a 99 percent chance that the test result will be positive; however, the test is not perfect, and there is a one in a thousand chance that the test result will be positive on a healthy person. Assume that a new patient arrives and the test result is positive. What is the probability that the patient has the disease?

Answer:

Let us represent disease by $d$ and test result by $t$. We are given the following: $P(d = 1) = 10^{-6}, P(t = 1|d = 1) = 0.99, P(t = 1|d = 0) = 10^{-3}$. We are asked $P(d = 1|t = 1)$. We use Bayes’ rule.

$$P(d = 1|t = 1) = \frac{P(t = 1|d = 1)P(d = 1)}{P(t = 1)} = \frac{P(t = 1|d = 1)P(d = 1)}{P(t = 1|d = 1)P(d = 1) + P(t = 1|d = 0)P(d = 0)} = \frac{0.99 \cdot 10^{-6}}{0.99 \cdot 10^{-6} + 10^{-3} \cdot (1 - 10^{-6})} = 0.00098902$$

That is, knowing that the test result is positive increased the probability of disease from one in a million to one in a thousand. But since the disease is so rare, testing positive still has a low probability to indicate that the patient has the disease.

2. In a two-class problem, the log odds is defined as

$$\log \frac{P(C_1|x)}{P(C_2|x)}$$

Write the discriminant function in terms of the log odds.

Answer:

We define a discriminant function as

$$g(x) = \log \frac{P(C_1|x)}{P(C_2|x)}$$

and choose $\left\{ \begin{array}{ll} C_1 & \text{if } g(x) > 0 \\ C_2 & \text{otherwise} \end{array} \right.$

Note that log odds is the sum of log likelihood ratio and log of prior ratio:
\[ g(x) = \log \frac{p(x|C_1)}{p(x|C_2)} + \log \frac{P(C_1)}{P(C_2)} \]

If the priors are equal, the discriminant is just the log likelihood ratio.

3. In a two-class, two-action problem, if the loss function is \( \lambda_{11} = \lambda_{22} = 0, \lambda_{12} = 10, \) and \( \lambda_{21} = 5, \) write the optimal decision rule?

Answer: Let us calculate the expected risks of the two actions:

\[
R(\alpha_1|x) = 0 \cdot P(C_1|x) + 10 \cdot P(C_2|x) = 10 \cdot (1 - P(C_1|x))
\]

\[
R(\alpha_2|x) = 5 \cdot P(C_1|x) + 0 \cdot P(C_2|x) = 5 \cdot P(C_1|x)
\]

We choose \( \alpha_1 \) if

\[
R(\alpha_1|x) < R(\alpha_2|x)
\]

\[
10 \cdot (1 - P(C_1|x)) < 5 \cdot P(C_1|x)
\]

\[
P(C_1|x) > \frac{2}{3}
\]

If \( P(C_1|x) < \frac{2}{3}, \) we use action \( \alpha_2 \)

4. Given the following data of transactions at a shop, calculate the support and confidence values of milk \( \rightarrow \) bananas, bananas \( \rightarrow \) milk, milk \( \rightarrow \) chocolate, and chocolate \( \rightarrow \) milk.

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Items in basket</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>milk, bananas, chocolate</td>
</tr>
<tr>
<td>2</td>
<td>milk, chocolate</td>
</tr>
<tr>
<td>3</td>
<td>milk, bananas</td>
</tr>
<tr>
<td>4</td>
<td>chocolate</td>
</tr>
<tr>
<td>5</td>
<td>chocolate</td>
</tr>
<tr>
<td>6</td>
<td>milk, chocolate</td>
</tr>
</tbody>
</table>
The association rules and their support and confidence values are as follows:

- milk $\rightarrow$ bananas : Support = 2/6, Confidence = 2/4
- bananas $\rightarrow$ milk : Support = 2/6, Confidence = 2/2
- milk $\rightarrow$ chocolate : Support = 3/6, Confidence = 3/4
- chocolate $\rightarrow$ milk : Support = 3/6, Confidence = 3/5

Though only half of the people who buy milk buy bananas too, anyone who buys bananas also buys milk.

5. For the multinomial we discussed in class (e.g., with probability $p_i$, outcome $i$ will occur and there are $K$ different possible outcomes), prove that the MLE (or the log likelihood) is

$$\hat{p}_i = \frac{\sum_t x_{it}}{N}$$

**Answer:**

$$J(p_i) = \sum_i \sum_t x_{it} \log p_i + \lambda \left(1 - \sum_i p_i\right)$$

$$\frac{\partial J}{\partial p_i} = \frac{\sum_t x_{it}}{p_i} - \lambda = 0$$

$$\lambda = \frac{\sum_t x_{it}}{p_i} \Rightarrow p_i \lambda = \sum_t x_{it}$$

$$\sum_i p_i \lambda = \sum_i \sum_t x_{it} \Rightarrow \lambda = \sum_t \sum_i x_{it}$$

$$p_i = \frac{\sum_t x_{it}}{\sum_t \sum_i x_{it}} = \frac{\sum_t x_{it}}{N} \text{ because } \sum_i x_{it} = 1$$
6. Given two normal distributions \( p(x|C_1) \sim N(\mu_1, \sigma_1^2) \) and \( p(x|C_2) \sim N(\mu_2, \sigma_2^2) \) and \( P(C_1) \) and \( P(C_2) \), calculate the Bayes' discriminant points analytically.

**Answer:**

Given that

\[
p(x|C_1) = \mathcal{N}(\mu_1, \sigma_1^2) = \frac{1}{\sqrt{2\pi} \sigma_1} \exp \left[ -\frac{(x - \mu_1)^2}{2\sigma_1^2} \right]
\]

\[
p(x|C_2) = \mathcal{N}(\mu_2, \sigma_2^2)
\]

we would like to find \( x \) that satisfy \( P(C_1|x) = P(C_2|x) \), or

\[
p(x|C_1)P(C_1) = p(x|C_2)P(C_2)
\]

\[
\log p(x|C_1) + \log P(C_1) = \log p(x|C_2) + \log P(C_2)
\]

\[
-\frac{1}{2} \log 2\pi - \log \sigma_1 - \frac{(x - \mu_1)^2}{2\sigma_1^2} + \log P(C_1) = \cdots
\]

\[
- \log \sigma_1 - \frac{1}{2\sigma_1^2} \left( x^2 - 2x\mu_1 + \mu_1^2 \right) + \log P(C_1) = \cdots
\]

\[
\left( \frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_2^2} \right) x^2 + \left( \frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} \right) x + \left( \frac{\mu_2^2}{2\sigma_1^2} - \frac{\mu_1^2}{2\sigma_2^2} \right) + \log \frac{\sigma_2}{\sigma_1} + \log \frac{P(C_1)}{P(C_2)} = 0
\]

This is of the form \( ax^2 + bx + c = 0 \) and the two roots are

\[
x_1, x_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]

Note that if the variances are equal, the quadratic terms vanishes and there is one root, that is, the two posteriors intersect at a single \( x \) value.