7

AI Homework #5

Dominique
Owens
ddo3

1) a) $$P(\theta \mid y,n) = \frac{p(y \mid \theta,n) \cdot P(\theta \mid n)}{p(y \mid n)}$$

$$= \frac{p(y \mid \theta,n) \cdot P(\theta \mid n)}{\int_0^1 p(y \mid \theta,n) \cdot p(\theta \mid n)\, d\theta}$$

$$= \frac{P(y \mid \theta,n) \cdot P(\theta \mid n)}{\frac{1}{1+n}}$$

$$= P(y \mid \theta,n) \cdot P(\theta \mid n) \cdot (1+n)$$

b + c) on a seperate page

2) a+c) on a seperate page

b) For this problem, I will solve it using hypothesis one, but the equation will be the same for any hypothesis

$P(h1|d) = .9$

$$.9 = P(d|h1) \cdot P(h1)$$
$$= \left[\prod_j P(d_j|h1)\right] \cdot P(h1)$$
$$= P(l|h1)^k \cdot P(c|h1)^{n-k} \cdot P(h1)$$

n = # of candies, k = # of lime candies,
l = lime candy   c = cherry candy

$$P(c|h1)^{n-k} = \frac{.9}{P(h1) \cdot P(l|h1)^k}$$

$$(n-k) \cdot \log(P(c|h1)) = \log(.9) - \log(P(h1) \cdot P(l|h1)^k)$$

$$n-k = \frac{\log(.9) - \log(P(h1) \cdot P(l|h1)^k)}{\log(P(c|h1))}$$

$$n = \left[\frac{\log(.9) - \log(P(h1) \cdot P(l|h1)^k)}{\log(P(c|h1))}\right] + k$$

3) (a) $P(x) = P(X \mid c = C_1) \cdot P(c = C_1) + P(x \mid c = C_2) \cdot P(c = C_2)$

(b) If we plot the probability density function (pdf) of x given each of the classes, we get two curves that overlap at our decision boundary. The total error would be the total area of this overlapping section, so to find the total error, we need to integrate over this area using the pdf for each of our curves

(c) - The following equation is the integration over this area.

$$\text{total error} = \int \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{\frac{-(\theta-\mu_1)^2}{2\sigma_1^2}} + \int \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{\frac{-(\theta-\mu_2)^2}{2\sigma_2^2}}$$

3 ⓒ To minimize the probability of miscalculations, we set the two integrals from the previous question equal to each other and solve for theta

$$\int \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{\frac{-(\theta-\mu_1)^2}{2\sigma_1^2}} = \int \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{\frac{-(\theta-\mu_2)^2}{2\sigma_2^2}}$$

$$\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\theta-\mu_1}{\sigma_1\sqrt{2}}\right)\right] = \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\theta-\mu_2}{\sigma_2\sqrt{2}}\right)\right]$$

$$\frac{\theta-\mu_1}{\sigma_1\sqrt{2}} = \frac{\theta-\mu_2}{\sigma_2\sqrt{2}}$$

$$\frac{\theta-\mu_1}{\sigma_1} = \frac{\theta-\mu_2}{\sigma_2}$$

$$(\theta-\mu_1)\sigma_2 = (\theta-\mu_2)\sigma_1$$

$$\theta\sigma_2 - \mu_1\sigma_2 = \theta\sigma_1 - \mu_2\sigma_1$$

$$\theta\sigma_2 + \theta\sigma_1 = \mu_1\sigma_2 - \mu_2\sigma_1$$

$$\theta(\sigma_2+\sigma_1) = \mu_1\sigma_2 - \mu_2\sigma_1$$

$$\theta = \frac{\mu_1\sigma_2 - \mu_2\sigma_1}{\sigma_1+\sigma_2}$$

4) The vector form of μ_K is directly from the slides in lecture 19

Given our objective function:

$$D = \sum_{n=1}^{N} \sum_{K=1}^{K} r_{n,K} \| x_n - \mu_K \|^2$$

to get the update rule, we must first take the derivative of our objective function, set it equal to zero, and solve for μ_K.

$$\frac{\partial D}{\partial \mu_K} = 2 \sum_{n=1}^{N} r_{n,K} (x_n - \mu_K)$$

by setting this equation equal to zero and solving for μ_K, we get the following equation:

$$\mu_K = \frac{\sum_n r_{n,K} \cdot x_n}{\sum_n r_{n,K}}$$

To get the scalar form we need the relation:

$$\| x_n - \mu_K \|^2 = \sum_i (x_{n,i} - \mu_{K,i})^2$$

we get this relation from the fact that $\|x\|^2 = \sum_i x_i^2$

Given this relation, our new objective function is:

$$D = \sum^{N} \sum^{K} r_{n,k} \sum_i (x_{n,i} - \mu_{k,i})^2$$

Using this objective function we can can get the scalar form of the μ_k.

$$\mu_{k,i} = \frac{\sum_{n=1}^{N} r_{n,k} \cdot x_{n,i}}{\sum_{n=1}^{N} r_{n,k}}$$