Classifier based on Maximum Likelihood (ML) Event (MLE) and Maximum-A-Posterior Probability (MAP)

Q: (Math 1.1) Animal scientists have found that the probability of finding small ears in Cats and Dogs are 0.8 and 0.1, respectively. Suppose An animal is observed with Large Ears.

(i) What is the probability that the observed animal is a dog?

(ii) To which class the test animal will be classified by the MLE classifier - CAT or DOG?

Cats and Dogs

- Suppose we have these conditional probability mass functions for cats and dogs
  - P(small ears | dog) = 0.1, P(large ears | dog) = 0.9
  - P(small ears | cat) = 0.8, P(large ears | cat) = 0.2
- Observe an animal with large ears
  - Dog or a cat?

likelihood = P (feature | class) ... called likelihood of that particular feature in the given class

  - Makes sense to say dog because probability of observing large ears in a dog is much larger than probability of observing large ears in a cat

kelihood)

  - ***Pr***[large ears | dog] = 0.9 > 0.2= ***Pr***[large ears | cat] = 0.2

SO ===> Classify as "DOG" (Solved)

  - Core Idea of MLE classifier: Choose the event of largest probability, i.e. maximum likelihood event (here, the events are "being Dog" and "being Cat" and the Maximum Likelihood event is "being Dog"

(ii) MAP

$$P(\omega_i|x) = \frac{P(x|\omega_i) \cdot P(\omega_i)}{P(x)}$$

MAP ⇒ Maximum A Posterior.

(i) ML.

$$\boxed{P(x|\omega_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

x = feature,

Q: (Math 1.2) fish experts have found that the length of Salmon & Bass fishes follow Gaussian distribution (i.e., Normal distribution) with mean of 5 and 10 inches, respectively and varince of 1 and 4 inch^2, respectively. A fish is observed with length of 7 inch. Explain how an (i) MLE and (ii) MAP(Bayes) classifiers would classify it - Salmon or Bass? (iii) What will be the classification decision by them if (a) Salmon & Bass are equally likely (b) Salmon is twice as likely as Bass?

- Respected fish expert says that
  - Salmon' length has distribution $\boldsymbol{N}(5,1)$
  - Sea bass's length has distribution $\boldsymbol{N}(10,4)$

The variances are 1 and 4 cm^2, which means the STD values (standard deviations) are 1 and 2 cm, respectively.

- Recall if r.v. is $\boldsymbol{N}(\mu,\sigma^2)$ then it's density is ***Probability Density Function/PDF***

$$p(l)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(l-\mu)^2}{2\sigma^2}}$$

(Dont Miss the 'minus' in the exponent)

- Thus *class conditional* densities are

class conditional PDFs

$$p(l\,|\,salmon)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(l-5)^2}{2*1}} \qquad p(l\,|\,bass)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(l-10)^2}{2*4}}$$

Likelihood function

- Thus *class conditional densities* are

$$p(l \mid \underset{fixed}{salmon}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-5)^2}{2*1}} \qquad p(l \mid \underset{fixed}{bass}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-10)^2}{2*4}}$$

- Fix length, let fish class vary. Then we get *likelihood function* (it is ***not density*** and ***not probability mass function)***

This is called class conditional probability desnity value:

$$p(\underset{fixed}{l} \mid class) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-5)^2}{2}} & \text{if } class = salmon \\ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-10)^2}{8}} & \text{if } class = bass \end{cases}$$

(i.e., likelihood of finding the fixed length of 'l' in the class 'salmon' or 'bass')

ML (maximum likelihood) Classifier

- We would like to choose salmon if

$$Pr[length = 7 \mid salmon] > Pr[length = 7 \mid bass]$$

- However, since ***length*** is a continuous r.v.,

$$Pr[length = 7 \mid salmon] = Pr[length = 7 \mid bass] = 0$$

- Instead, we choose class which maximizes likelihood

$$p(l \mid salmon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-5)^2}{2}} \qquad p(l \mid bass) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-10)^2}{2*4}}$$

- ML classifier: for an observed ***l***:

$$p(l \mid salmon) \underset{salmon}{\overset{bass}{\lessgtr}} p(l \mid bass)$$

in words: if p(l | salmon) > p(l | bass), classify as salmon, else classify as bass

MATH 1.2

(i) ML or MLE Classifier

likelihood of $l=7$ $= P(l=7 \mid \text{Salmon}) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{(l-\mu)^2}{2\sigma^2}} = \frac{1}{1\cdot\sqrt{2\pi}} e^{-\frac{(7-5)^2}{2\cdot 1}}$

$$= \frac{1}{\sqrt{2\pi}} e^{\frac{-4}{2}} \quad \boxed{= \frac{1}{\sqrt{2\pi}}(0.135)}$$

likelihood of $l=7$

$= P(l=7 \mid \text{sea bass}) = \frac{1}{\underset{=2}{\sigma}\sqrt{2\pi}} e^{-\frac{(l-10)^2}{2\cdot 4}} = \frac{1}{2}\frac{1}{\sqrt{2\pi}} e^{-\frac{(7-10)^2}{8}}$

$$= \frac{1}{\sqrt{2\pi}} \quad \boxed{= \frac{1}{\sqrt{2\pi}}(0.16)}$$

Thus: $P(l=7 \mid \text{sea bass}) > P(l=7 \mid \text{Salmon})$

So: $\boxed{\text{classify as Sea bass}}$

(ii) Bayes Classifier

$$P(\text{salmon} / l=7) = \frac{P(l=7 \mid \text{salmon}) * P(\text{salmon})}{P(l=7)}$$

$= \frac{1}{2}$ as no prior probability given

$$P(\text{sea bass} \mid l=7) = \frac{P(l=7 \mid \text{sea bass}) * P(\text{sea bass})}{P(l=7)}$$

$= \frac{1}{2}$ as no prior probability given

$$\boxed{= \frac{\frac{1}{\sqrt{2\pi}}(0.135) * \frac{1}{2}}{P(l=7)}}$$

$$\boxed{= \frac{\frac{1}{\sqrt{2\pi}}(0.16) * \frac{1}{2}}{P(l=7)}}$$

→ This one Larger value

So: $\boxed{\text{classify as Seabass}}$

(iii) MLE classifier is $\boxed{\text{not}}$ affected by prior probability.
So: its output will be same as before (Sea bass)

But: Bayes Classifier (MAP classifier) considers the Prior probability. So its decision may change.

(a) when $P(\text{Salmon}) = P(\text{Sea bass}) = \frac{1}{2}$

$$P(\text{salmon} \mid l=7) = \frac{\frac{1}{\sqrt{2\pi}}(0.135) * \frac{1}{2}}{P(l=7)}$$ [see previous page]

$$P(\text{sea bass} \mid l=7) = \frac{\frac{1}{\sqrt{2\pi}}(0.16) * \frac{1}{2}}{P(l=7)}$$ [see previous page]

↓

Longer. So Classify as $\boxed{\text{Sea bass}}$

(b) ~~So: P~~

$P(\text{Salmon}) = \frac{2}{3}$

$P(\text{sea bass}) = \frac{1}{3}$

$$P(\text{salmon} \mid l=7) = \frac{P(l=7 \mid \text{Salmon}) * P(\text{Salmon})}{P(l=7)} = \frac{\frac{1}{\sqrt{2\pi}}(0.135) * \frac{2}{3}}{P(l=7)} = \boxed{\frac{\frac{1}{\sqrt{2\pi}} * (0.09)}{P(l=7)}}$$

⇓ Larger value

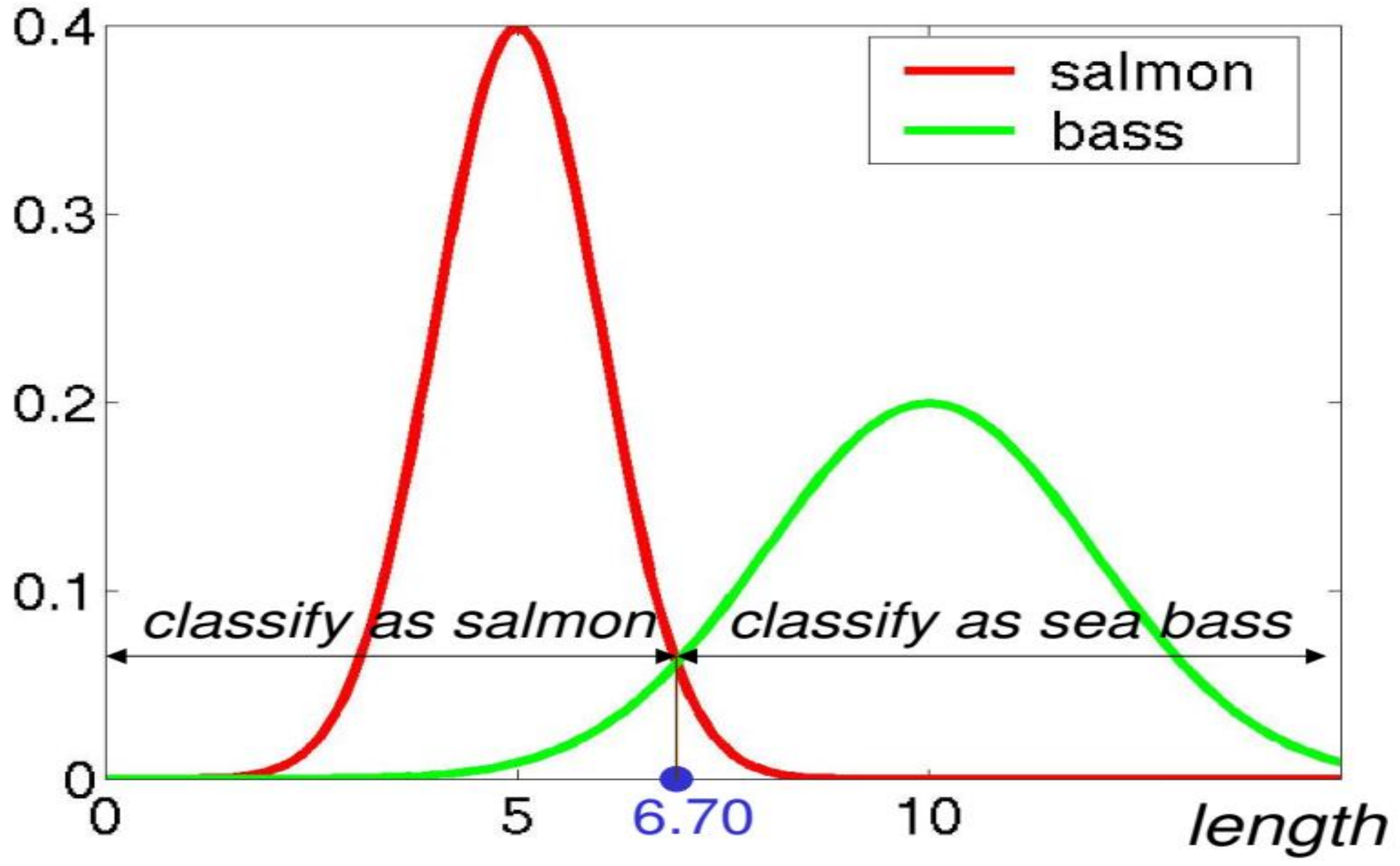
$$P(\text{sea bass} \mid l=7) = \frac{\frac{1}{\sqrt{2\pi}}(0.16) \cdot \frac{1}{3}}{P(l=7)} = \boxed{\frac{\frac{1}{\sqrt{2\pi}} * 0.053}{P(l=7)}}$$

So classify as $\boxed{\text{Salmon}}$

Q. (Math 1.3) Find the Decision Boundary between the Salmon and Bass classes based on their length, when no prior knowledge is available.

Decision Boundary

Q: What is decision boundary? The Decision Boundary between class i and class j is the set of All Points where both the classes have equal likelihood value (MLE classifier), or equal posterior probability value (Bayes classifier) or equal discrminant function value (for the Minimum error rate classifier)



MATH 1·3

at Decision Boundary l, both P(Salmon) and P(Sea Bass) will be equal.

~~P(l/salmon) = P(l/~~

For Bayes Classifier / MAP classifier, Start from Here:

$$P(\text{Salmon} \mid l) = P(\text{Sea Bass} \mid l)$$

$$\Rightarrow \frac{P(l \mid \text{Salmon}) * \overbrace{P(\text{Salmon})}^{\frac{1}{2} \text{ as no Prior given}}}{P(l)} = \frac{P(l \mid \text{SeaBass}) * \overbrace{P(\text{Sea bass})}^{\frac{1}{2} \text{ as No Prior given}}}{P(l)}$$

for MLE classifier, Just Start From Here:

$$\Rightarrow P(l \mid \text{Salmon}) = P(l \mid \text{Sea Bass})$$

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

$$\Rightarrow \frac{1}{1\sqrt{2\pi}} e^{\frac{-(l-5)^2}{2\cdot 1}} = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(l-10)^2}{2\cdot 4}}$$

$$\Rightarrow e^{-\frac{(l-5)^2}{2}} = \frac{1}{2} e^{\frac{-(l-10)^2}{8}}$$

take $\log_e$ of both sides

$$\Rightarrow -\frac{(l-5)^2}{2} = -\log_e 2 - \frac{(l-10)^2}{8}$$

Multiply by -8

$$\Rightarrow 4(l-5)^2 = 8 * (\log_e 2) + (l-10)^2$$

$$\Rightarrow 4l^2 - 40l + 100 = 5.54 + l^2 - 20l + 100$$

$$\Rightarrow 3l^2 - 20l = 5.54 \Rightarrow 3l^2 - 20l - 5.54 = 0 \rightarrow$$

দ্বিঘাত সমীকরণ $ax^2+bx+c=0$ এর roots: $\frac{-b+\sqrt{b^2-4ac}}{2a}, \frac{-b-\sqrt{b^2-4ac}}{2a}$

→ [illegible]:

$l = \underset{\text{ok}}{\underline{6.93}}, \underset{\text{invalid}}{-0.27}$

Decision Boundary: 6.93 inch

(Math 1.4) Fish experts have found that there are twice as many Salmon as sea bass. With this prior knowledge, How should the ML classifier and Bayes classifier classify a fish of length 7.0?

Ans: for MLE classifier, NO MATTER / no effect of "twice as likely" . For Bayes classifier, see scanned solution!

Q 1.5 Find the decision boundary for the above problem for the Bayes / MAP classifier

$\boxed{\text{Math 1.5}}$

At decision Boundary l, we have $P(\text{Salmon} \mid l) = P(\text{Bass} \mid l)$

$$P(\text{Salmon} \mid \text{length}=l) = P(\text{Sea bass} \mid \text{length}=l)$$

$$\Rightarrow \frac{P(l \mid \text{Salmon}) * P(\text{Salmon})}{P(\text{length}=l)} = \frac{P(l \mid \text{Sea Bass}) \times P(\text{SeaBass})}{P(\text{length}=l)}$$

$$\Rightarrow P(l \mid \text{Salmon}) * \frac{2}{3} = P(l \mid \text{Sea Bass}) * \frac{1}{3}$$

$$\Rightarrow \frac{1}{1\sqrt{2\pi}} e^{-\frac{(l-5)^2}{2 \cdot 1}} * \frac{2}{3} = \frac{1}{2\sqrt{2\pi}} e^{-\frac{(l-10)^2}{2 \cdot 4}} \cdot \frac{1}{3}$$

$$\Rightarrow 2e^{-\frac{(l-5)^2}{2}} = \frac{1}{2} e^{-\frac{(l-10)^2}{8}}$$

$$\Rightarrow 4e^{-\frac{(l-5)^2}{2}} = e^{-\frac{(l-10)^2}{8}}$$

Take $\log_e$ of Both sides $\quad \log_e 4 - \frac{(l-5)^2}{2} = -\frac{(l-10)^2}{8}$

Multiply by 8 $\quad 8(\log_e 4) - 4(l-5)^2 = -(l-10)^2$

$$\Rightarrow 4(l-5)^2 - 8(\log_e 4) = (l-10)^2$$

$$\Rightarrow 4l^2 - 40l + 100 - 11.09 = l^2 - 20l + 100$$

$$\Rightarrow \boxed{3l^2 - 20l - 11.09 = 0}$$

$ax^2+bx+c=0$, so roots $\frac{-b \pm \sqrt{b^2-4ac}}{2a}$

Roots $l = \underset{\text{OK}}{7.18}, \underset{\text{invalid}}{-0.51}$

So. Decision Boundary 7.18 inch