Exercise 21

Diana Solis, Shakiba Fadaei, Todor Gitchev

5/19/2019

Introduction

We calculate the Bayes classifier for one explanatory variable (feature) under the assumption that the class-conditional probability densities are normal. We want to distinguish between two fish species, sea brass (class 0) and salmon (class 1). Assume the class-conditional probability densities for the length x of the fishes are given by the following normal distributions:

$X|Y=0 \sim \mathcal{N}\left(\mu=7, \sigma^2=1\right)$ for sea brass and $X|Y=1 \sim \mathcal{N}\left(\mu=3, \sigma^2=1\right)$ for salmon

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

```
length <- seq(0, 12, by=0.1) # lengt measurments

fx_y0 <- function(length) dnorm(length, mean = 7, sd = 1)
fx_y1 <- function(length) dnorm(length, mean = 3, sd = 1)
```

a)

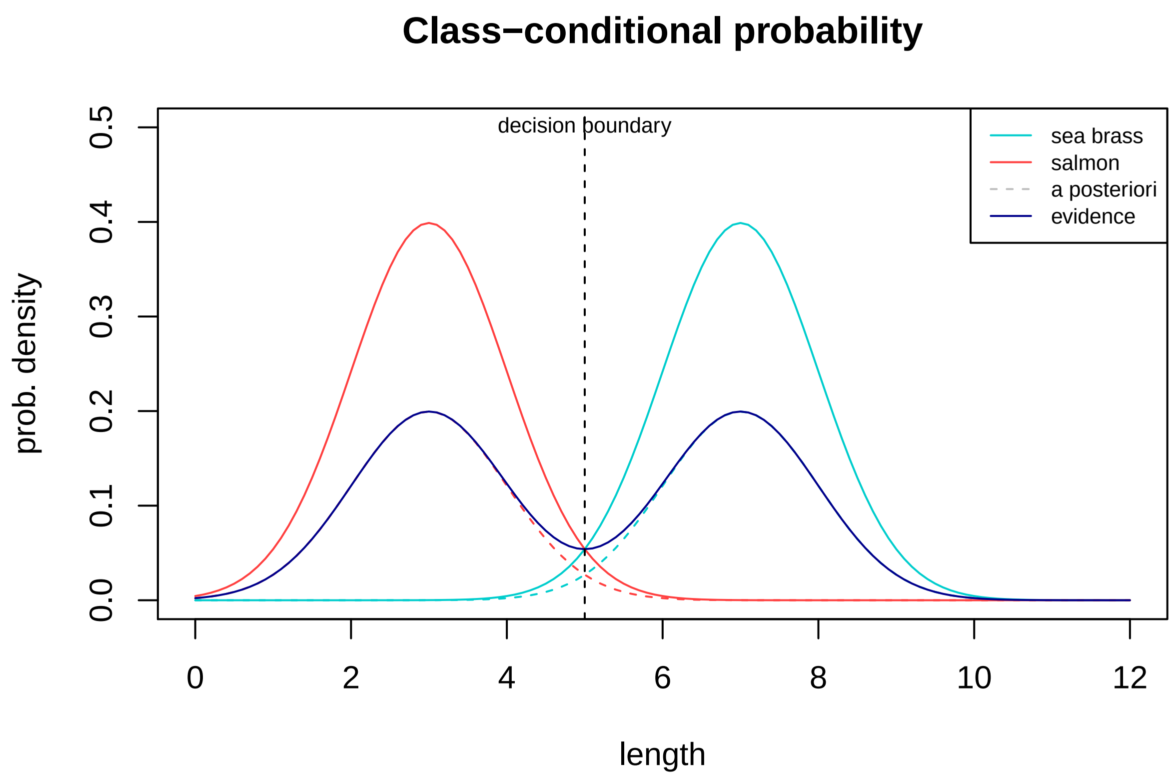
We have prior class probabilities: $\pi_0 = P[Y=0] = \pi_1 = P[Y=1] = 1/2$

posterior class probabilities: $P\left(y_j|x\right) = \frac{p(x|y_j)\pi_j}{p(x)} \quad j = 0, 1$

evidence: $p(x) = \sum_{i=0}^{1} p\left(x|y_i\right)\pi_i$

a posteriory: $p\left(x|y_i\right)\pi_i$

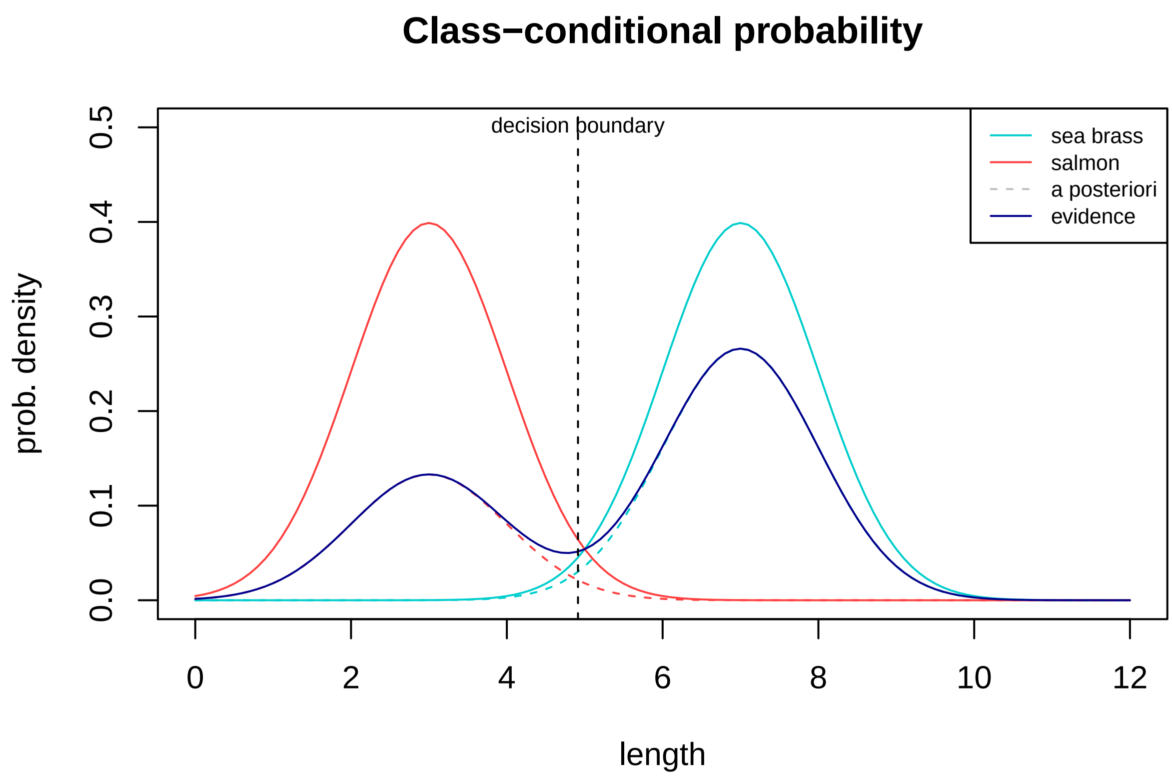
```
fx_evidence <- function(length, pi_0, pi_1) pi_0*fx_y0(length) + pi_1*fx_y1(length)

plot_class_prob <- function(pi_0, pi_1) {
  plot(length, fx_y0(length), main = "Class-conditional probability",  ylab="prob. density", ylim=c(0,0
  lines(length, pi_0*fx_y0(length), type="l", lty = 2, col="cyan3")
  lines(length, fx_y1(length),type="l", lty = 1, col="brown1")
  lines(length, pi_1*fx_y1(length), type="l", lty = 2, col="brown1")
  lines(length, fx_evidence(length, pi_0, pi_1), type="l", lty = 1, col="darkblue")
  decision_boundary <- (log(pi_1/pi_0) + 40) / 8
  abline(v=decision_boundary, lty=2, col="black" )
  text(decision_boundary, 0.5, "decision boundary", cex= 0.7)

  legend("topright", legend = c("sea brass", "salmon", "a posteriori", "evidence"), col=c("cyan3", "bro
}

plot_class_prob(0.5, 0.5)
```



b)

We have prior class probabilities: $\pi_0 = P[Y = 0] = 2/3$, $\pi_1 = P[Y = 1] = 1/3$

```
plot_class_prob(2/3, 1/3)
```



c)

Decision boundary:

$$\pi_0 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-7)^2}{2}} = \pi_1 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-3)^2}{2}}$$

$$\pi_0 \cdot e^{-(x-7)^2} = \pi_1 \cdot e^{-(x-3)^2}$$

$$log(\pi_0) - (x-7)^2 = log(\pi_1) - (x-3)^2$$

$$log(\pi_0) - x^2 + 14x - 49 = log(\pi_1) - x^2 + 6x - 9 => x = \frac{log(\pi_1/\pi_0) + 40}{8}$$