Solution to Bayes Exercise 2

Answer to Problem 1

First we should notice the prior information given,

$$P_{prior}(Ripe) = 0.15 \text{ and } P_{prior}(Unripe) = 0.85$$

Then we should note the likelihood of ripe and unripe fruits reflecting light of specific wavelengths. Let's denote the wavelength, *w*, then:

$$P_{like}(540 < w < 550 \mid \text{ripe}) = \Phi\left(\frac{550-600}{50}\right) - \Phi\left(\frac{540-600}{50}\right) = 0.044$$

and

$$P_{like}(540 < w < 550 \mid \text{unripe}) = \Phi\left(\frac{550-500}{50}\right) - \Phi\left(\frac{540-500}{50}\right) = 0.053$$

where

$$\Phi\left(\frac{w-\mu}{\sigma}\right)$$

is the normal cumulative density function (normcdf.m in Matlab) for the variable, *w*, distributed according to a normal distribution with mean, μ, and standard deviation, σ.

Now we're almost ready to calculate the posterior probability of the fruit with a wavelength between 540 and 550 nm is ripe by using Bayes' rule:

$$P_{post}(ripe \mid 540 < w < 550) = \frac{P_{like}(540 < w < 550 \mid ripe)P_{prior}(ripe)}{P(540 < w < 550)}$$

Note that the denominator is the overall, or unconditional probability of jujus (ripe or unripe) reflecting a wavelength between 540 and 55 nm. It can be calculated as a weighted sum:

$$\sum_{ripe} P_{like}(540 < w < 550 \mid ripe)P_{prior}(ripe) =$$
$$P_{like}(540 < w < 550 \mid ripe)P_{prior}(ripe) + P_{like}(540 < w < 550 \mid unripe)P_{prior}(unripe)$$

so that finally we arrive at

$$P_{post}(ripe \mid 540 < w < 550) =$$
$$\frac{P_{like}(540 < w < 550 \mid ripe)P_{prior}(ripe)}{P_{like}(540 < w < 550 \mid ripe)P_{prior}(ripe) + P_{like}(540 < w < 550 \mid unripe)P_{prior}(unripe)} \approx$$
$$\frac{0.044 \times 0.15}{0.044 \times 0.15 + 0.053 \times 0.85} \approx 0.1263$$

Answer to Problem 2

Again we need to calculate the posterior probability

$$P_{post}(ripe \mid 540 < w < 550)$$

but now we have three kinds of fruit and we only have information about prior probabilities and likelihoods specific to each fruit. The key to the solution is to realize that the type of fruit is irrelevant and needs to be *discounted*:

$$P_{post}(ripe \mid 540 < w < 550) = \sum_{fruit} P_{post}(ripe, fruit \mid 540 < w < 550)$$

We can then insert Bayes' rule

$$P_{post}(ripe, fruit \mid 540 < w < 550) =$$

$$\frac{P_{like}(540 < w < 550 \mid ripe, fruit)P_{prior}(ripe, fruit)}{P(540 < w < 550)} =$$

$$\frac{P_{like}(540 < w < 550 \mid ripe, fruit)P_{prior}(ripe, fruit)}{\sum_{fruit,ripe} P_{like}(540 < w < 550 \mid ripe, fruit)P_{prior}(ripe, fruit)}$$

The joint prior factorize so that $P_{prior}(ripe, fruit) = P_{prior}(ripe \mid \text{fruit})P_{prior}(fruit)$. Now you can insert everything in one big equation:

$$P_{post}(ripe \mid 540 < w < 550) =$$
$$\sum_{fruit} P_{post}(ripe, fruit \mid 540 < w < 550)$$

$$\sum_{fruit} \frac{P_{like}(540 < w < 550 \mid ripe, fruit)P_{prior}(ripe \mid \text{fruit})P_{prior}(fruit)}{\sum_{fruit,ripe} P_{like}(540 < w < 550 \mid ripe, fruit)P_{prior}(ripe \mid \text{fruit})P_{prior}(fruit)} \approx 0.402$$

So, chances are that the fruit is not ripe (<50% chance) and I wouldn't bother to pick it. If, however, I saw a monkey enjoying the fruit, I would think that the fruit was probably ripe. The monkey enjoying the fruit thus *explains away* the possibility of the fruit being unripe.

Answer to Problem 3

Simulating fruiting picking by sampling from a multinomial distribution (see the solution code for one way to do this). The sample gives you the number of jujus, mongos and chakavas as well as the number of them that were ripe. Now that you have the fruits you should generate a wavelength for each and every fruit by sampling from the likelihood function specific to that fruit and its ripeness:

$$P(\mathrm{w} \mid ripe, fruit) = \varphi((w-\mu)/\sigma)$$

Here, $\varphi((w-\mu)/\sigma)$, is the normal probability density. Now that you have 1000 wavelengths, you can determine the monkey's posterior probability for each fruit being a juju, mongo or chakava; ripe or unripe. This is exactly what you did in Problem 2. Now you just have to do it for many different wavelength intervals and not just for 540-550 nm. Finally let the monkey make a decision using the *max aposteriori* rule and count the number of times the monkey was right. A typical answer could look like this:

number of ripe jujus: 14
number of ripe jujus picked: 7
number of unripe jujus: 89
number of unripe jujus picked: 15
number of ripe mongos: 373
number of ripe mongos picked: 340
number of unripe mongos: 110
number of unripe mongos picked: 9
number of ripe chakavas: 42
number of ripe chakavas picked: 8
number of unripe chakavas: 372
number of unripe chakavas picked: 119

Note that the monkey makes a lot of mistakes picking unripe chakavas. It's because they look like ripe mongos.