Homework 1

Yuan Li, N19728558

Part I: Written Exercises

1. (a) Answer:

Set "p_i" as the probability of finding the correct $Label_i$ in a cave, $i = 1, 2, ..., 5$; "p_a" as the probability of incorrectly labeling a crab if the scientist predicts the species uniformly at random.

$$p_a = \sum_{i=1}^{5} 0.2 * (1 - p_i)$$
$$= 0.2 * (1 - 0.25) + 0.2 * (1 - 0.05) + 0.2 * (1 - 0.2) + 0.2 * (1 - 0.35) + 0.2 * (1 - 0.15)$$
$$= 0.15 + 0.19 + 0.16 + 0.13 + 0.17$$
$$= \mathbf{0.8}$$

1. (b) Answer:

Yes, the scientist can improve his error by picking the same label every time. The label which the scientist choose is **Label 4**. We can calculate the Bayes error rate. Set the probability as "p_b"

$$p_b = 1 - E(max p(p_i | p_b))$$
$$= 1 - p_4$$
$$= 1 - 0.35$$
$$= \mathbf{0.65}$$

2. Answer:
We can bring the value into the following formula:

$$p(x|\mu, \Sigma) = \frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}} exp(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu))$$

if $p(x|\mu_{blue}, \Sigma_{blue}) > p(x|\mu_{orange}, \Sigma_{orange})$, then the crab's color is blue, otherwise the color is orange.

$$p(x|\mu_{blue}, \Sigma_{blue}) = \frac{1}{\sqrt{2\pi}|\Sigma_{blue}|^{1/2}} exp(-\frac{1}{2}(x - \mu_{blue})^T \Sigma_{blue}^{-1}(x - \mu_{blue}))$$
$$= 2.166 * 10^{-277}$$

$$p(x|\mu_{orange}, \Sigma_{orange}) = \frac{1}{\sqrt{2\pi}|\Sigma_{orange}|^{1/2}} exp(-\frac{1}{2}(x - \mu_{orange})^T \Sigma_{orange}^{-1}(x - \mu_{orange}))$$
$$= 2.152 * 10^{-141}$$

Because $p(x|\mu_{blue}, \Sigma_{blue}) < p(x|\mu_{orange}, \Sigma_{orange})$, so it's more likely to be orange.

3. (a) Answer:

Because θ is the probability of Heads, So:

$$p(Heads) = \theta$$
$$p(Tails) = 1 - \theta$$

Then we can get $p(HHTHH)$ as:

$$p(HHTHH) = \theta * \theta * (1 - \theta) * \theta * \theta$$
$$= \theta^4 * (1 - \theta)$$

3. (b) Answer:

$$logp(HHTHH) = log(\theta^4 * (1 - \theta))$$
$$= log\theta^4 + log(1 - \theta)$$
$$= 4log\theta + log(1 - \theta)$$

3. (c) Answer:

Because the log function is a monotonically increasing function. so we can calculate the the maximum likelihood estimate of θ as follow:

$$0 = \frac{\partial logp(HHTHH)}{\partial \theta}$$
$$= \frac{\partial}{\partial \theta}[4log\theta + log(1 - \theta)]$$
$$= \frac{4}{\theta} - \frac{1}{1 - \theta}$$

So we can get the θ should be:

$$\theta = \mathbf{0.8}$$