Exercise 1

Florian Klein 412984
Krzysztof Kostrzewa 380029
Rico Stanosek 433500

Question 2

(a)

Probability of picking an apple:

$$\begin{aligned} P(Apple) &= P(Red) * P(Apple|Red) + P(Green) * P(Apple|Green) + P(Blue) * P(Apple|Blue) \\ &= \frac{3}{50} + \frac{9}{50} + \frac{1}{10} \\ &= \frac{17}{50} = 34\% \end{aligned}$$

(b)

Probability that if we picked an orange, it came from the green box:

$$P(Green|Orange) = \frac{P(Orange|Green)}{P(Orange)} = \frac{0.3 \cdot 0.6}{0.36} = 0.514 = 51.4\%$$

Question 3

L_{kj} is defined as the loss inferred for a decision C_j if the true class is C_k:

$$L_{kj} = \begin{cases} 0, & k = j \\ l_r, & j = N+1 \\ l_s, & otherwise \end{cases}$$

$$L_{kj} = \begin{pmatrix} 0 & l_s & \dots & \dots & l_s & l_r \\ l_s & 0 & \ddots & & \vdots & \vdots \\ \vdots & \ddots & 0 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 0 & l_s & \vdots \\ l_s & \dots & \dots & l_s & 0 & l_r \end{pmatrix}$$

(a)

$$\begin{aligned} E[L] &= \sum_k \sum_j \int_{R_j} L_{kj} \cdot p(x, C_k) dx \\ \Rightarrow &= \sum_k L_{kj} \cdot p(C_k|x) \\ &= \sum_{k=1}^{N} l_s \cdot p(C_k|x) \\ &= \sum_{k=1}^{N} l_s * \frac{p(x|C_k) * p(C_k)}{p(x)} \end{aligned}$$

(b)

If $l_r = 0$, then there is no loss if you reject the current sample. This also means that the sample rather gets rejected than assigned to a (even very) probable class. This may lead to wrong decisions or no decisions at all.

(c)

If $l_r > l_s$, then the loss of the rejection is higher than the loss incurred for making a substitution error. This means that a given sample is rather assigned to a class than rejected.

Additionally, this also can lead to faulty decisions while guaranteeing that the given sample is not rejected.

Question 4

Given that all N measurements $x_1, \ldots, x_N > 0$, we can derive that $g(x) = 1$.

$\Rightarrow p(x|\theta) = \theta^2 \cdot x \cdot \exp(-\theta \cdot x)$

$$\begin{aligned} L(p) &= ln(\theta^2 \cdot x \cdot \exp(-\theta \cdot x)) \\ &= ln(\theta^2) + ln(x) + ln(\exp(-\theta \cdot x)) \\ &= ln(\theta^2) + ln(x) - \theta \cdot x \end{aligned}$$

Derive $L(p)$ and set to 0:

$$\begin{aligned} \frac{\delta}{\delta\theta} L(p) &= L'(p) \\ &= \frac{2}{\theta} - x \\ \Rightarrow 0 &\stackrel{!}{=} \frac{2}{\theta} - x \\ 0 &\stackrel{!}{=} \frac{2}{\theta} - \sum_{i=1}^{N} X_i \\ \sum_{i=1}^{N} X_i &\stackrel{!}{=} \frac{2}{\theta} \\ \Rightarrow \theta &= \frac{2}{\sum_{i=1}^{N} X_i} \end{aligned}$$

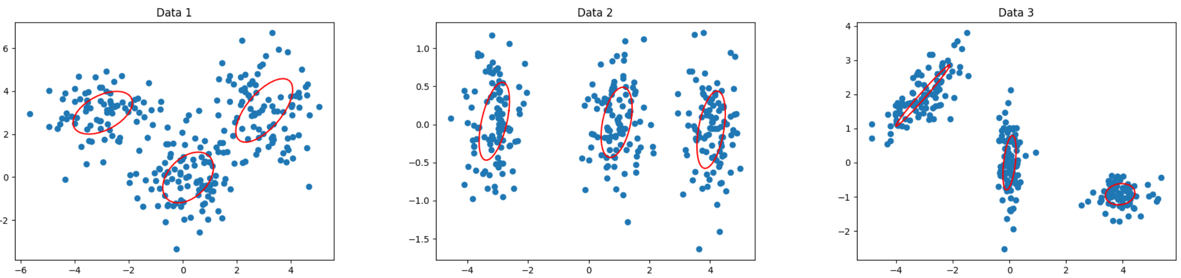
Question 5

Please see the included source code.

Question 6

Please see the included source code.

(f)



(a) data1 (b) data2 (c) data3

Figure 1: Visualisation of the provided synthetic data

etc). Examine how the likelihood of the computed model changes with K. Are these likelihoods comparable? Which is the best K and why?

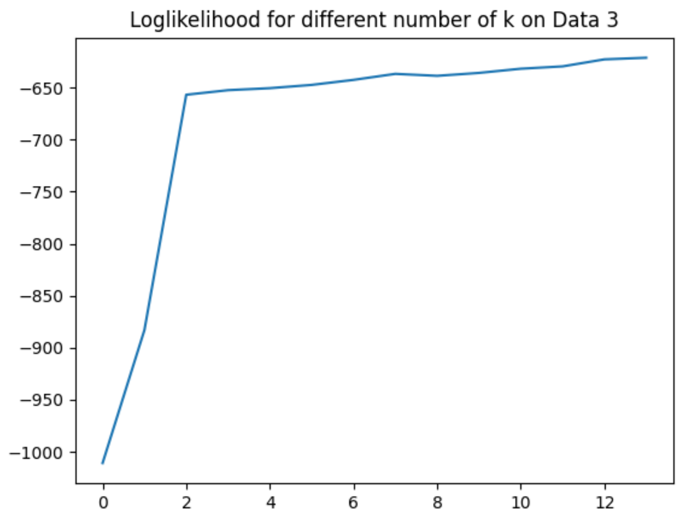


Figure 2: Change of the loglikelihood over different values of K

As you can see in figure 2, the likelihood given K has a rapid increase in value until $K = 2$. From there on the increase is noticeable but marginal.

A value for K of at least 2 is necessary to have decent results.