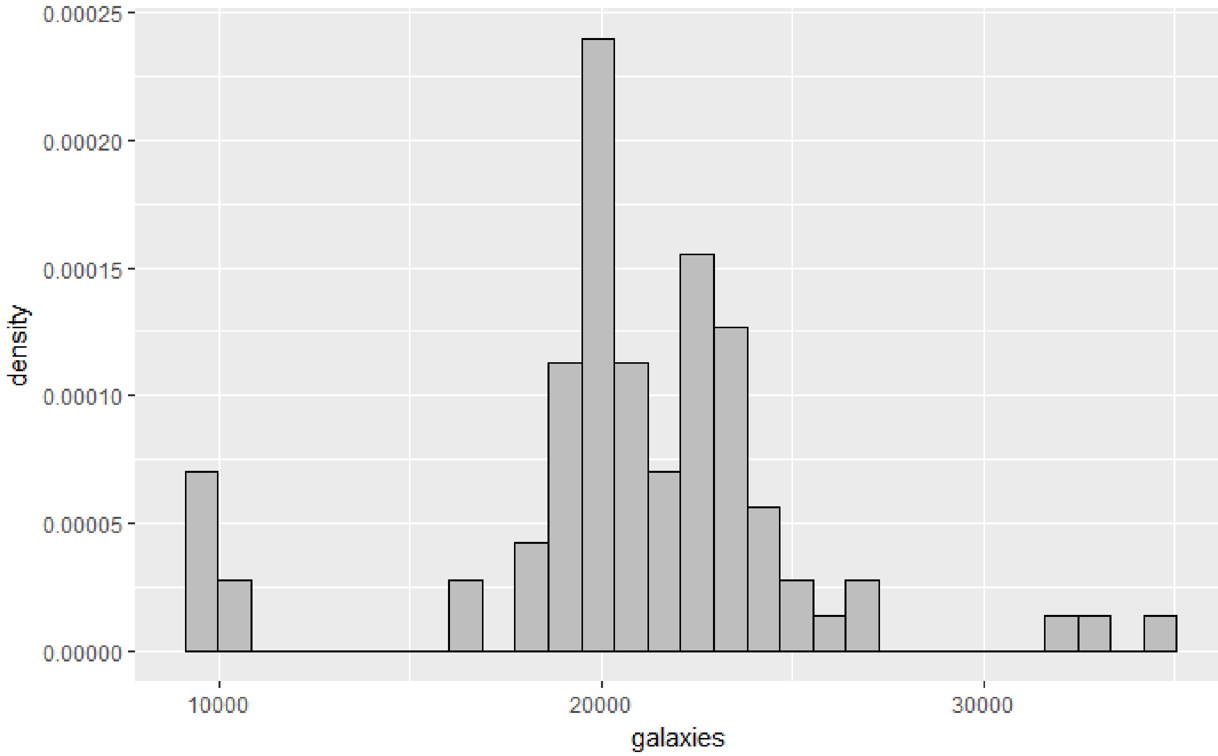
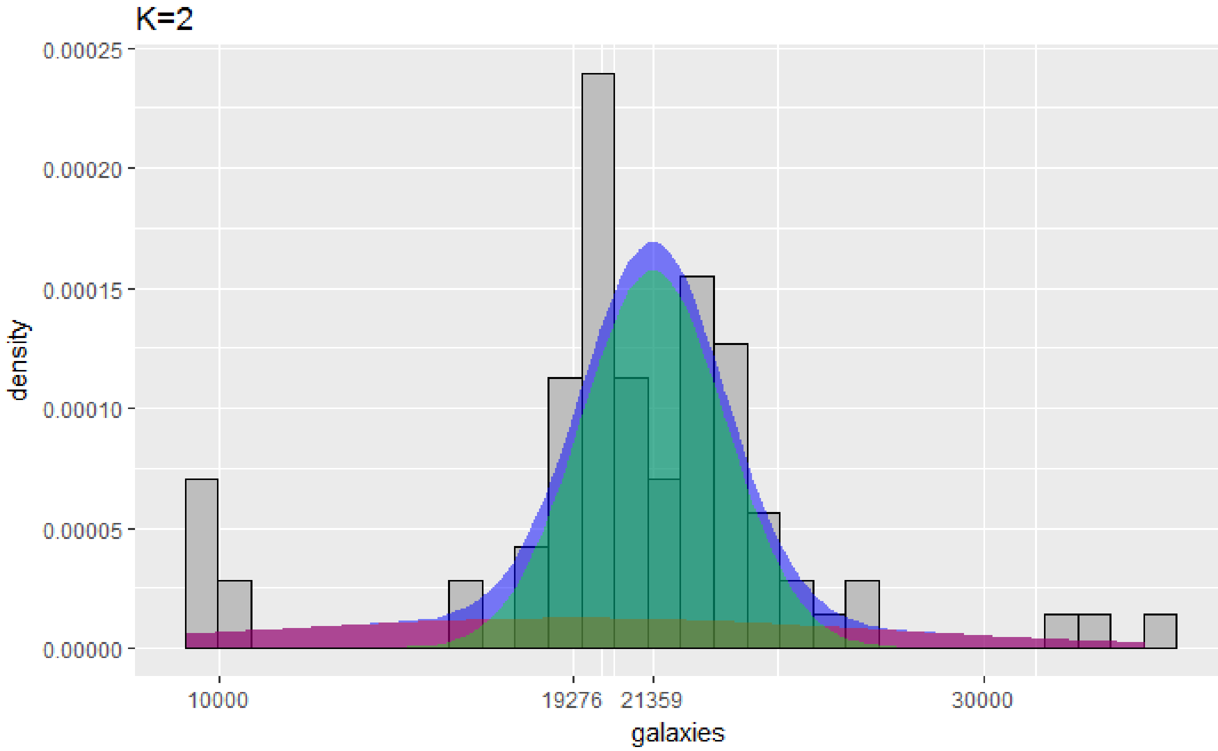
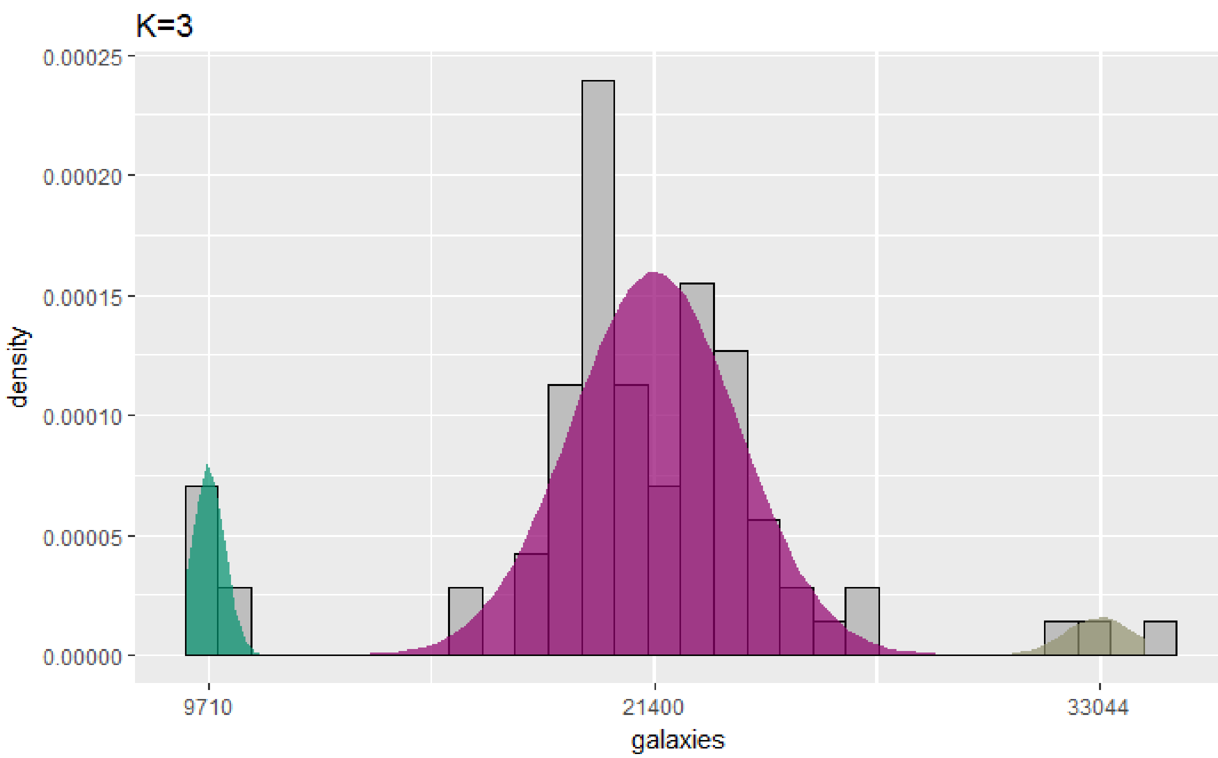
Exercise 1.2

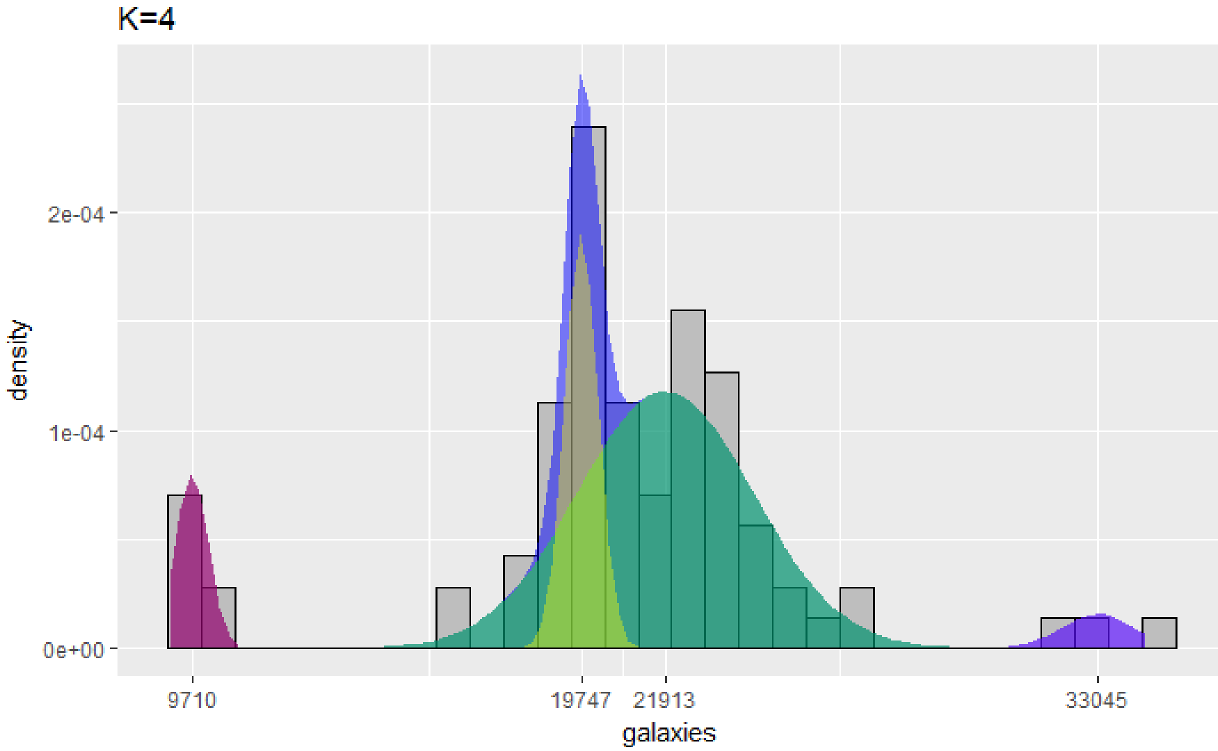
1.2a

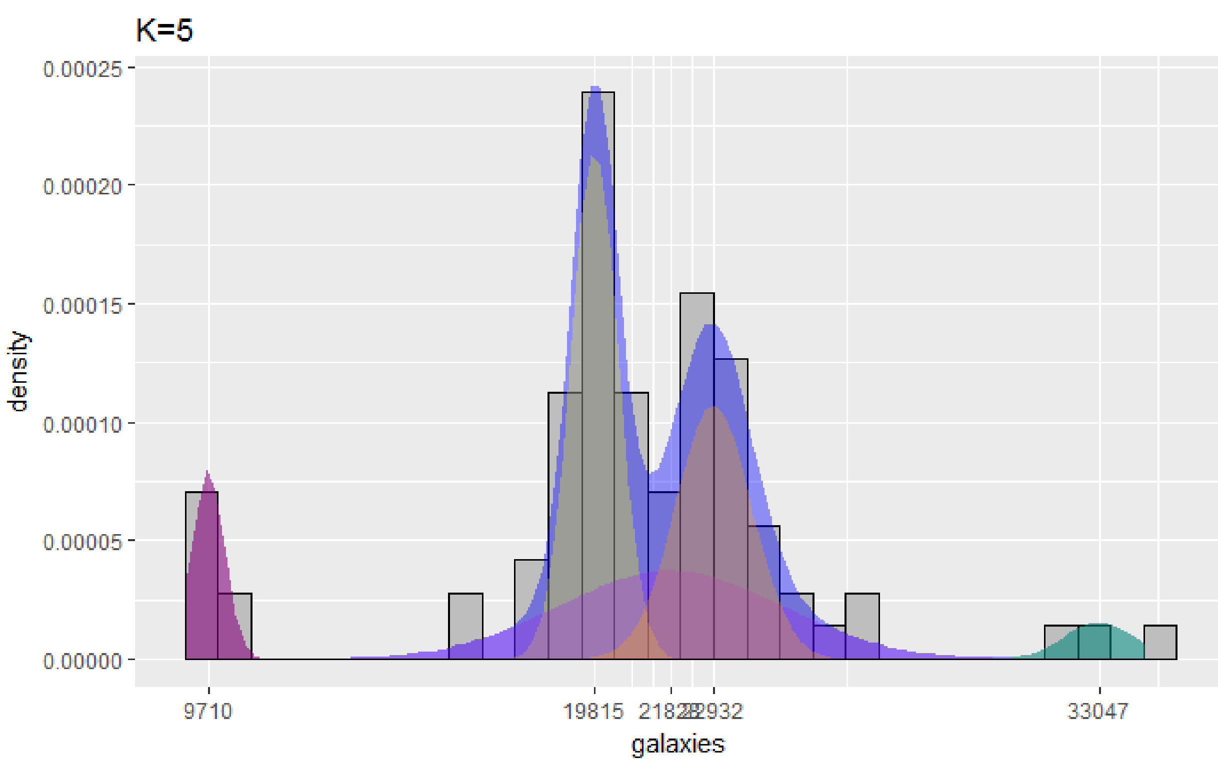
Fitting mixed Gaussians on the galaxy data set in R. Different colors (except for blue) indicate the different Gaussian components. The blue distribution equals the sum of the components.




K=2
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
density
10000
19276
21359
30000
galaxies


K=3
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
density
9710
21400
33044
galaxies


K=4
2e-04
1e-04
0e+00
density
9710
19747
21913
33045
galaxies


K=5
0.00025
0.00020
0.00015
0.00010
0.00005
0.00000
density
9710
19815
33047
galaxies

Exercise 1.2b

$\mathcal{F}$ is the class of functions. Sample $X_1, \ldots, X_n \sim f_0 \in \mathcal{F}$. Then $X_1, \ldots, X_n$ are finite a.s.

Let $M \in \mathbb{R}_+$ be large. Our goal is to find an $f \in \mathcal{F}$ such that

$$\prod_{i=1}^{n} f(X_i) > M$$

Let $K = n$ be the number of Gaussian distributions in the mixture. Set $p_k = 1/n$ for all $k = 1, \ldots, n$. Then, for all $\sigma > 0$, letting $\sigma_1 = \ldots = \sigma_n = \sigma$, the function

$$f_\sigma = \frac{1}{n\sigma} \sum_{k=1}^{n} \phi\left(\frac{x - X_k}{\sigma}\right)$$

is an element of $\mathcal{F}$ a.s. Then

$$\begin{aligned}
\prod_{j=1}^{n} f_\sigma(X_j) &= \prod_{j=1}^{n} \left[\frac{1}{n\sigma} \sum_{k=1}^{n} \phi\left(\frac{X_j - X_k}{\sigma}\right)\right] \\
&= (n\sigma)^{-n} \prod_{j=1}^{n} \left[\sum_{k=1}^{n} \phi\left(\frac{X_j - X_k}{\sigma}\right)\right] \\
&\geq (n\sigma)^{-n} \prod_{j=1}^{n} \phi\left(\frac{X_k - X_k}{\sigma}\right) \\
&\geq (n\sigma)^{-n} \prod_{j=1}^{n} \phi\left(0\right) \\
&= \left(n\sigma\sqrt{2\pi}\right)^{-n}
\end{aligned}$$

Letting

$$\sigma < M^{-1/n} \frac{1}{n\sqrt{2\pi}}$$

we see that

$$\prod_{j=1}^{n} f_\sigma(X_j) > M$$

Since for every $M \in \mathbb{R}_+$ we can find a f_σ such that this holds, we conclude that

$$\max_{f \in \mathcal{F}} \prod_{j=1}^{n} f(X_j) = \infty$$

a.s.