Homework 1 Responses

Chris Crabtree

Sept. 30, 2021

1

I spent a lot of time on questions 3,4, and 5 and could not finish this.

2

Similarly, I could not finish this question either.

3.a

Plots are given in figure 1. I used the log-likelihood function as my stopping criteria. For the these non-stochastic experiments I stopped iterating when the improvement in the log-likelihood was less than .001.
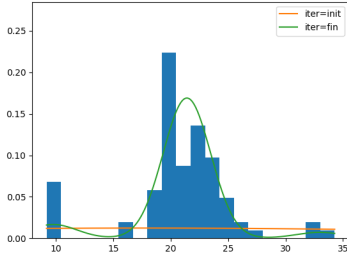
3.b

Tabulated AIC and BIC scored below:

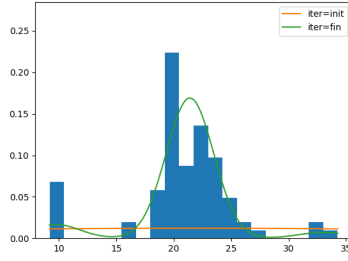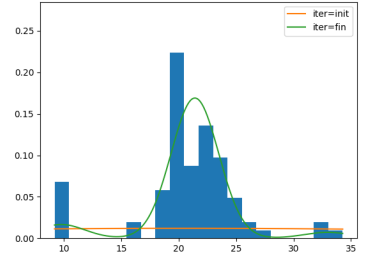| k	aic	bic	iters
4	433.4*	455.1*	88
6	450.7	482.0	40
8	449.4	490.4	267
11	440.0	495.4	173
15	448.6	523.2	179
20	468.7	567.4	476

Best values of each are marked with a *.

3.c

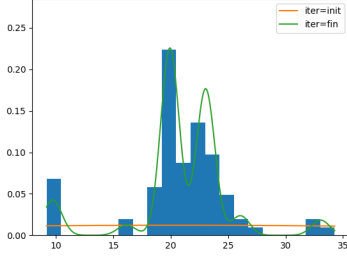Plots for the location-scale models are given in figure 2. For the location-scale models I limited the σ^2's to have a minimum of .001.
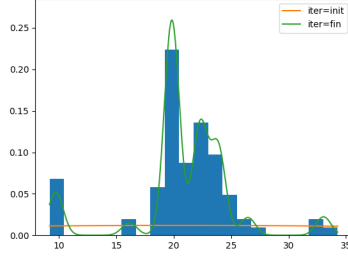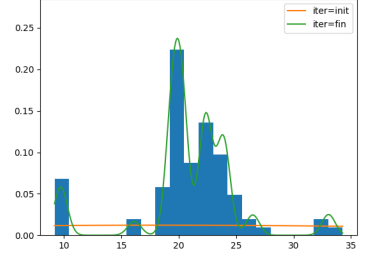
(a) K=4   (b) K=6   (c) K=8

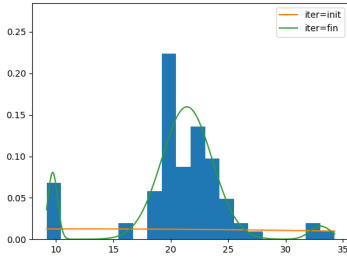(d) K=11   (e) K=15   (f) K=20
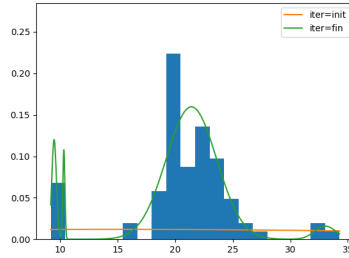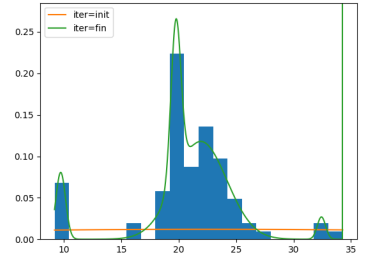
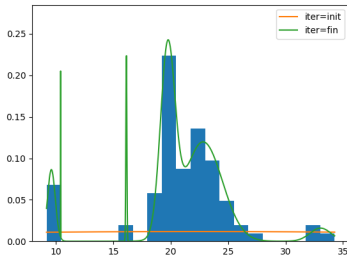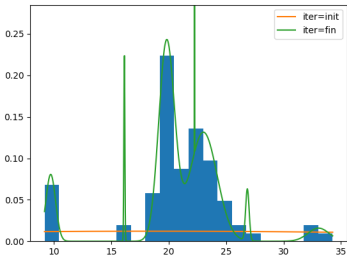Figure 1: Non-Stochastic 1-D Location Mixtures

(a) K=3   (b) K=4   (c) K=5
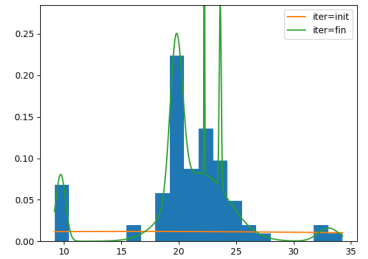
(d) K=6   (e) K=7   (f) K=8

Figure 2: Non-Stochastic 1-D Location-Scale Mixtures
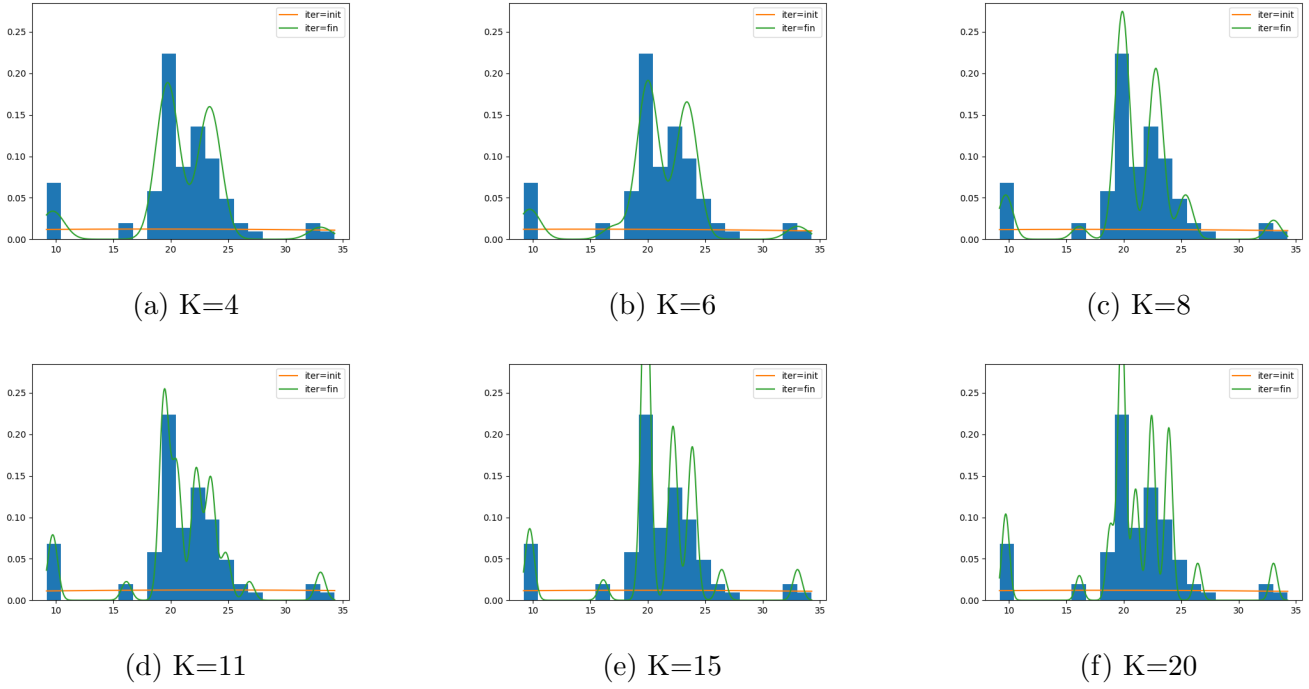
Figure 3: Stochastic 1-D Location Mixtures

3.d

Tabulated AIC and BIC scored below:

k	aic	bic	iters
3	424.4	446.0*	64
4	425.1	454.0	71
5	410.1*	446.2	123
6	425.9	469.3	549
7	410.5	461.1	522
8	412.0	469.7	499

Best values of each are marked with a *.

3.e

It seemed that the location only models produced more reliable results. Allowing the variance parameters to update allowed the excess centroids to increase the log-likelihood arbitrarily by decreasing the variance to zero. This is why I limited the minimum σ^2's.

It also appeared that initialization played a large role in the goodness of fit. The resulting mixture pdf's could vary substantially on repeated runs.

4.a

Plots given in figure 3. For the these *stochastic* experiments I stopped iterating when the improvement in the log-likelihood was less than .1. I lowered this threshold because I found that the stochasticity seemed to make the algorithm take thousands of iterations to converge in some cases.
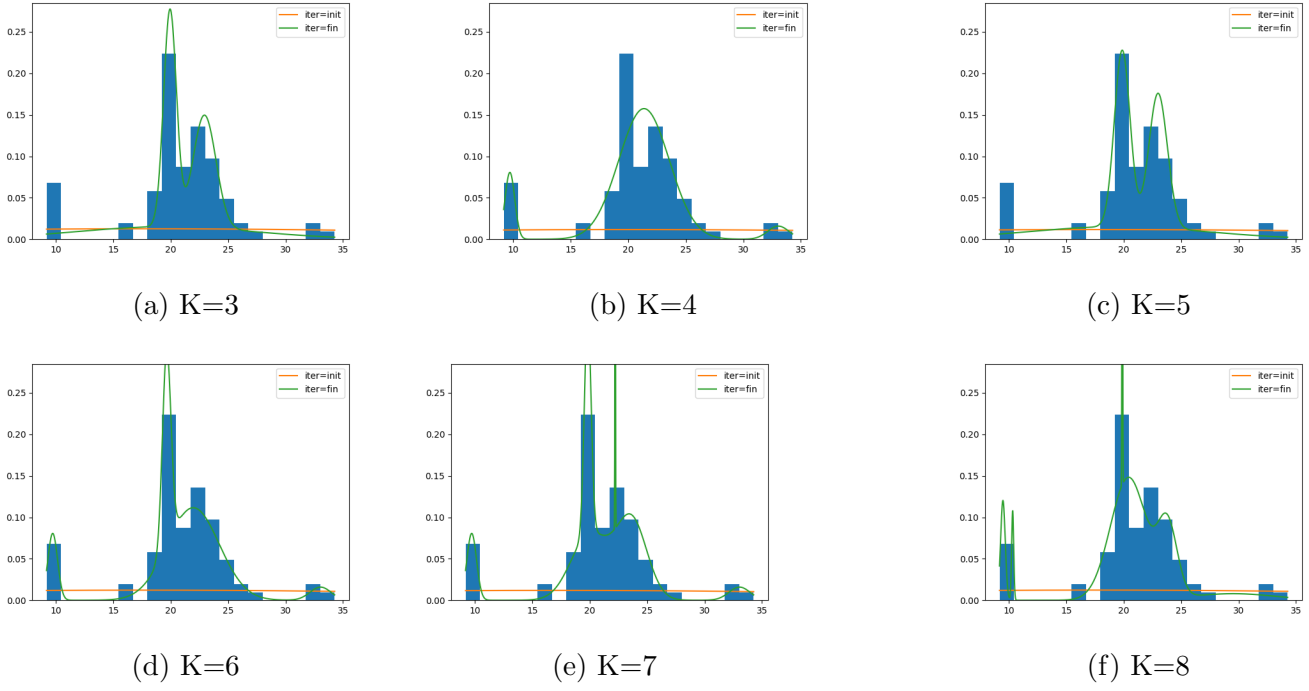
Figure 4: Stochastic 1-D Location-Scale Mixtures

4.b

Tabulated AIC and BIC scored below:

k	aic	bic	iters
4	478.7	500.4	29
6	435.2*	466.5*	53
8	443.2	484.1	98
11	455.2	510.6	78
15	448.6	523.2	117
20	475.6	574.3	77

Best values of each are marked with a *.

4.c

4.d

Tabulated AIC and BIC scored below:

k	aic	bic	iters
3	443.9	465.5	110
4	373.8	402.7	49
5	397.3	433.4	28
6	319.2*	362.5*	78
7	434.6	485.2	43
8	334.8	392.6	25

Best values of each are marked with a *.

4.e

The stochastic EM models did not seem to perform as well as the normal EM versions. The location only models tended to favor a high value of the sole $sigma^2$ parameter, but the location-scale models did tend to produce less centroids that became singularities.

5

I sadly did not have time to complete this.