Lab 3

Gustav Sternelöv

3 maj 2016

Assignment 1 - Normal model, mixture of normal model with semi-conjugate prior.

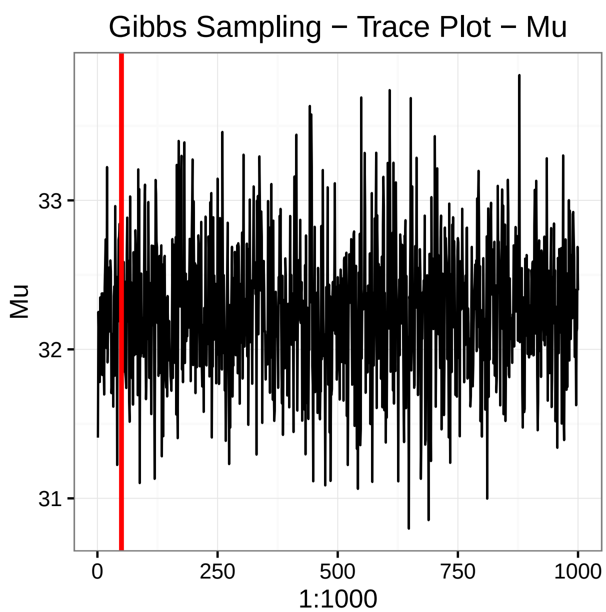
a) Normal model.

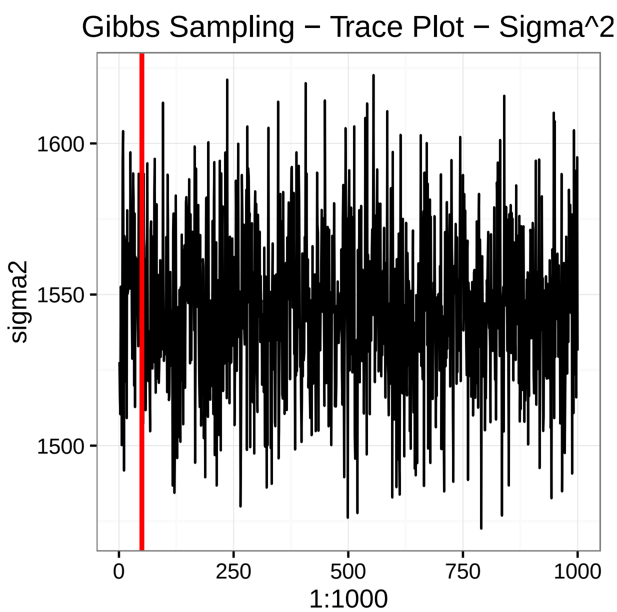
i)

The code used to implement the Gibbs sampler that simulates from the joint posterior can be seen in the appendix *R-code*.

ii)

The Gibbs sampler from *i)* is tested and it is of interest to investigate the convergence of the chains and if the sampler is efficient. One way to check this is by looking at trace plots and auto-correlation plots.

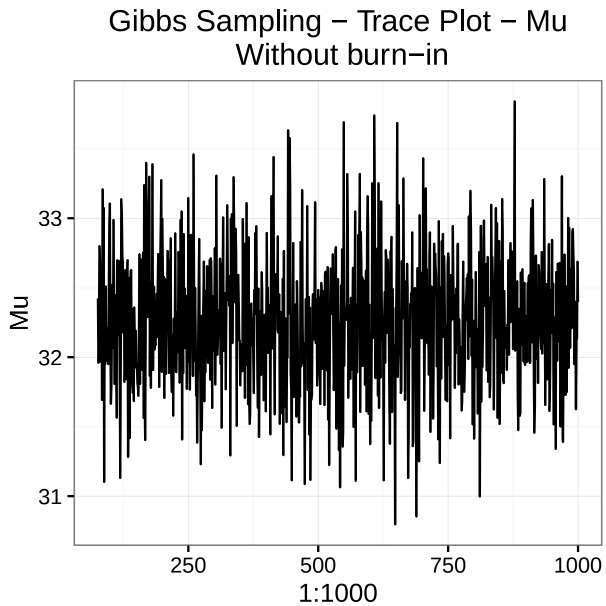


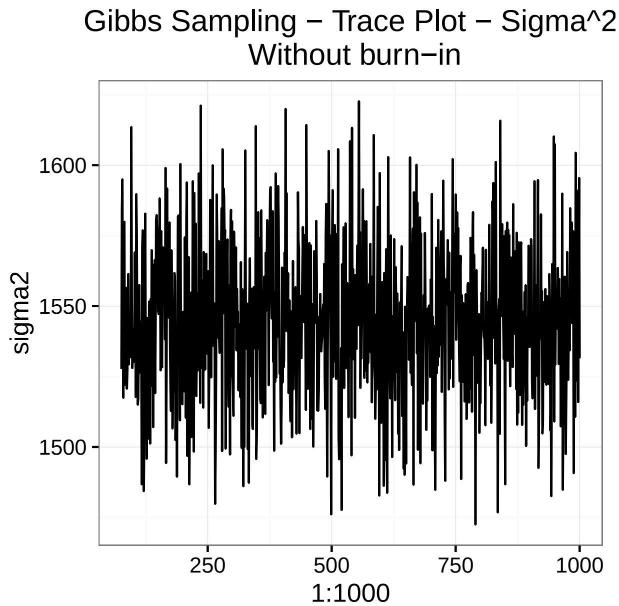


By the look of the plots above it seems like the chain has converged for both μ and σ^2. The chains converges quickly and if there is a burn-in period, it is thought to be short. Out of the 1000 iterations, perhaps 75 iterations in both cases can be classified as belonging to the burn-in period. Even though it is hard to see a specific burn-in period it is reasonable to make the assumption that some proportion of the first iterations not have converged and should be discarded.

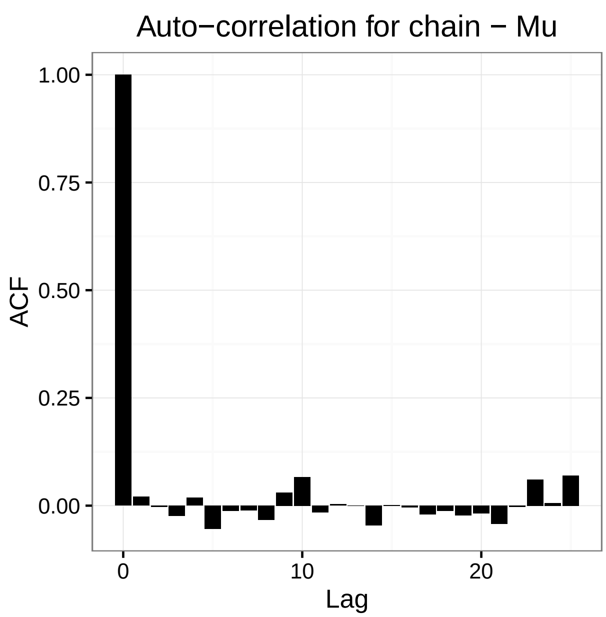
The chains are thought to have converged since they have settled rather well and have a good mixing. By mixing we mean that the chains not are following any clear pattern, the values looks to be random and quite uncorrelated to the upcoming values.

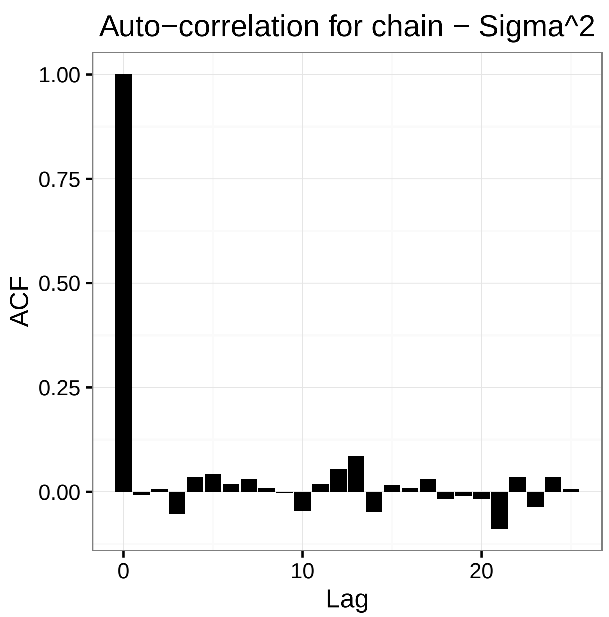
How the chains looks with the burn-in period discarded is shown by the trace plots below.





In the comments above about the trace plots we mentioned that the values seem to rather uncorrelated, the correlation between one value and the next upcoming values seem to be low. This is investigated more closely by looking at the autocorrelation of the chains with the burn-in removed.





As we thought, the generated values from the Gibbs sampler do not have a strong auto-correlation. The conlcusion from the visual examination of the Gibbs sampler then is that the chain both have converged and seem to be rather efficient.

b) Mixture normal model.

i)

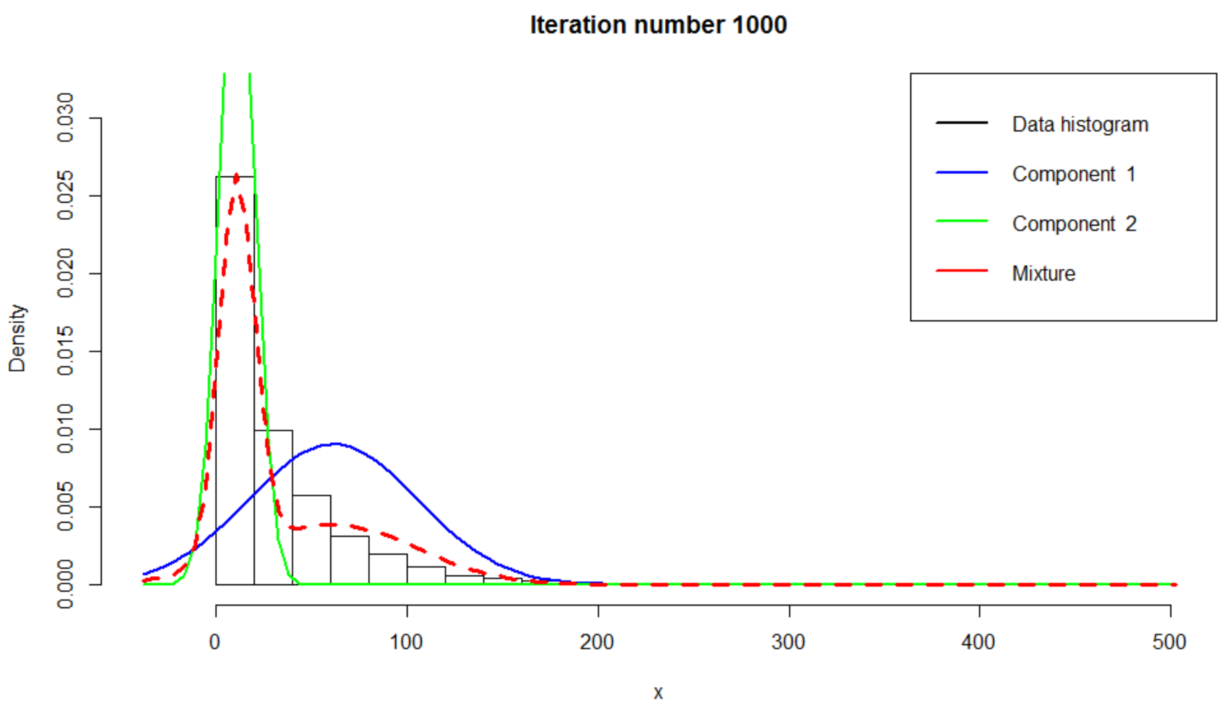
Mattias's code is used for this assignment and added to the appendix at the end of the report. The values for the prior hyperparameters are presented below.

We use the standard deviation for the observations to assign the prior standard deviation for μ.
The variance of the observations is used for σ_0^2.
The prior mean is set to 30, close to the mean for the observations (32.27).
For the prior on σ^2 is the degress of freedom set to 4.
α_0, the prior for the beta distribution, is set (10,10). This is a prior that is used for assigning the probability that a value belong to the distribution.

Some differents priors were tested and the values presented above seemed to be the most suitable ones.

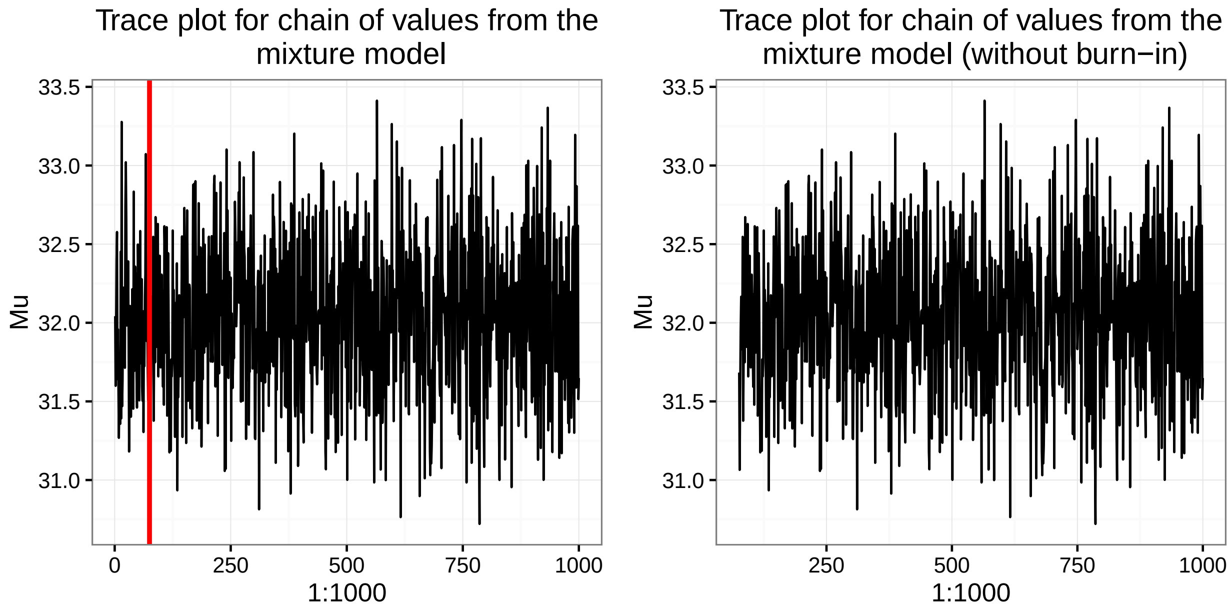
ii)

The resulting distribution given for both the normals and the mixture of the normals after 1000 iterations is shown by the plot below.



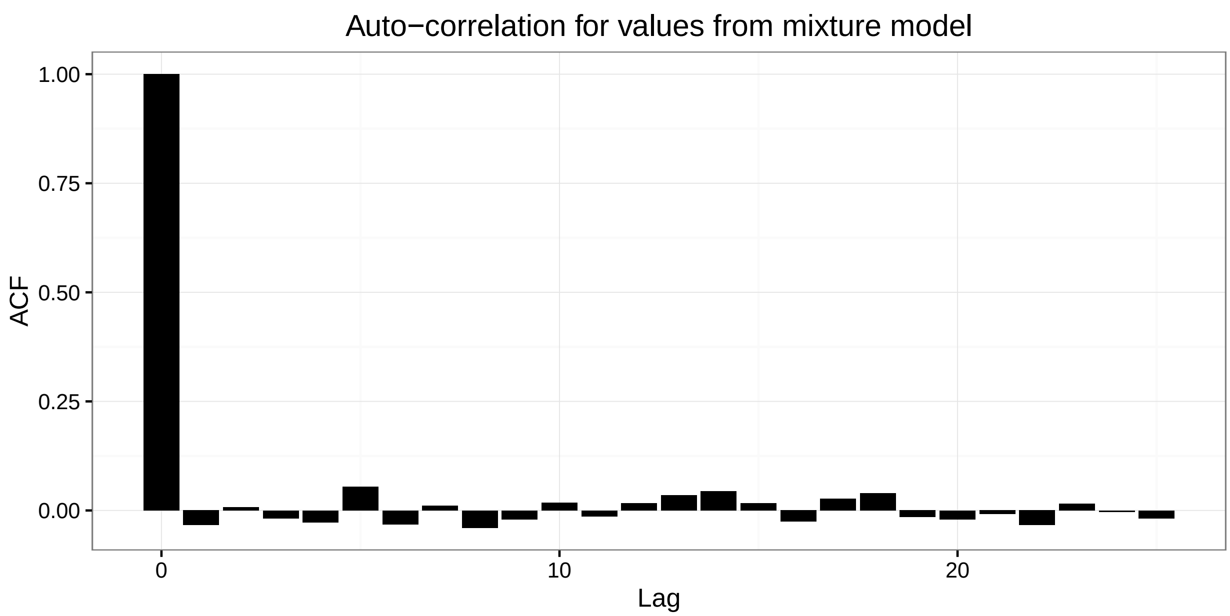
It can be seen the the mixture of the two normals quite well fits the observed data.

The convergence and efficiency of the Gibbs sampler is analyzed by looking at a trace plot for the chain of μ values given by the mixture of the normals.



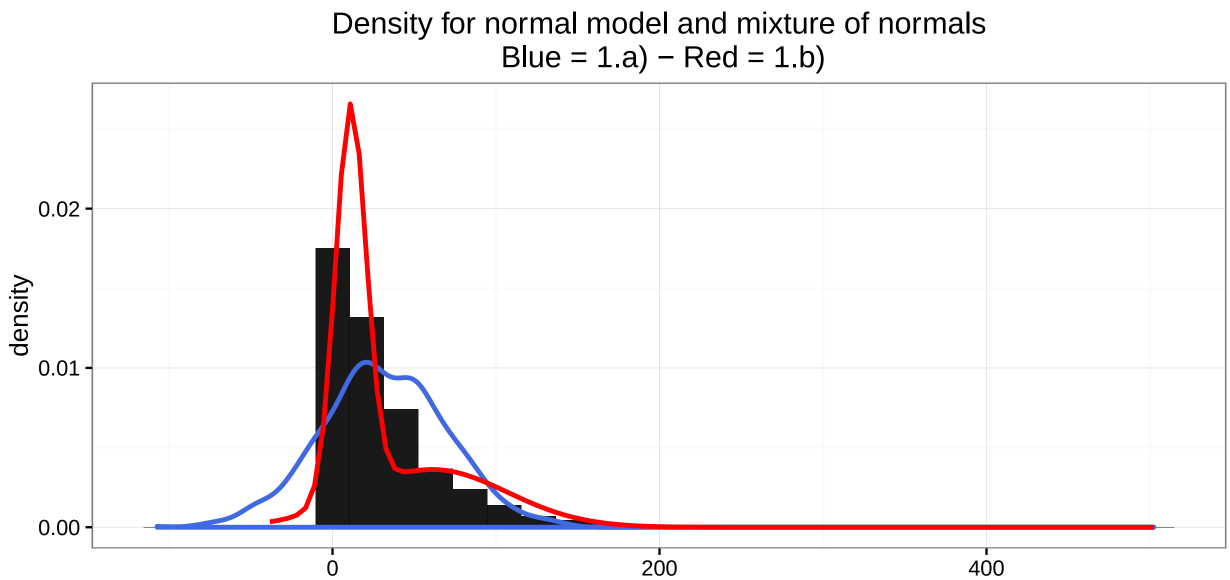
Again, it seems like the chain have converged rather rapidly. It settles quickly and generated follows does not get stuck nor follow any clear pattern. Looking at the burn-in period it is concluded that it probably is rather short. Hence, just the first 75 observations are discarded because of the burn-in period.

The autocorrelation is also visualised in order to investigate the convergence and efficiency of the chain.



The correlation between values from iterations close to each other is low which speaks in favor of the efficiency of the sampler.

c) Graphical comparison.



With no doubt, the mixture of normals model from *b)* is better in terms of fitting.

Assignment 2 - Binary regression models

a)

The code *OptimizeSpamR* is used in the spam data set and the results given by the code are presented below.

```
## [1] "The posterior mode is:"

##           our          over        remove      internet          free
##  0.2732466085  0.6815867686  1.2486100444  0.4539457721  0.6307512633
##           hpl            X.          X..1     CapRunMax   CapRunTotal
## -0.7684908652  0.1926618541  3.1970606303  0.0052354006  0.0004274572
##         const            hp        george         X1999            re
## -0.7398524569 -0.8940685712 -4.0724676136 -0.3214764568 -0.4075869443
##           edu
## -0.8881672861

## [1] "The approximate posterior standard deviation is:"

##  [1] 3.746767e-02 9.982157e-02 1.170905e-01 7.163540e-02 5.643741e-02
##  [6] 1.830635e-01 2.536963e-02 2.416595e-01 6.629637e-04 5.319823e-05
## [11] 4.360858e-02 1.100609e-01 5.197431e-01 8.927327e-02 6.797135e-02
## [16] 1.128443e-01
```

b-c)

A data augmentation Gibbs sampler is implemented and the code can be seen in the appendix at the end of the report.

The implemented Gibbs sampler is used for analyzing the spam data and the following results are received:

d)

The confusion matrices for the results in *a)* versus the results in *c)*: