

Sim Study Results Investigation

Jonathan Klus

2024-09-27

```
# homemade functions
# source("./Multivariate_DPMM_unknownvar_DEV.R")
# source("./Multivariate_DPMM_unknownvar_DEE.R")
#source("./Multivariate_DPMM_unknownvar_UVV.R")
source("./posterior_helper_fxns.R")

## Warning: package 'tidyr' was built under R version 4.2.2
## Warning: package 'dplyr' was built under R version 4.2.2
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
source("./post_processing_inf.R")

## Warning: package 'gridExtra' was built under R version 4.2.2
##
## Attaching package: 'gridExtra'
## The following object is masked from 'package:dplyr':
##
##   combine
##
## Attaching package: 'LaplacesDemon'
## The following objects are masked from 'package:gtools':
##
##   ddirichlet, logit, rdirichlet
## Warning: package 'stringr' was built under R version 4.2.2
## Package 'mclust' version 6.0.1
## Type 'citation("mclust")' for citing this R package in publications.
# load R libraries
library(ggplot2)
library(dplyr)
library(stringr)
library(scatterplot3d)
```

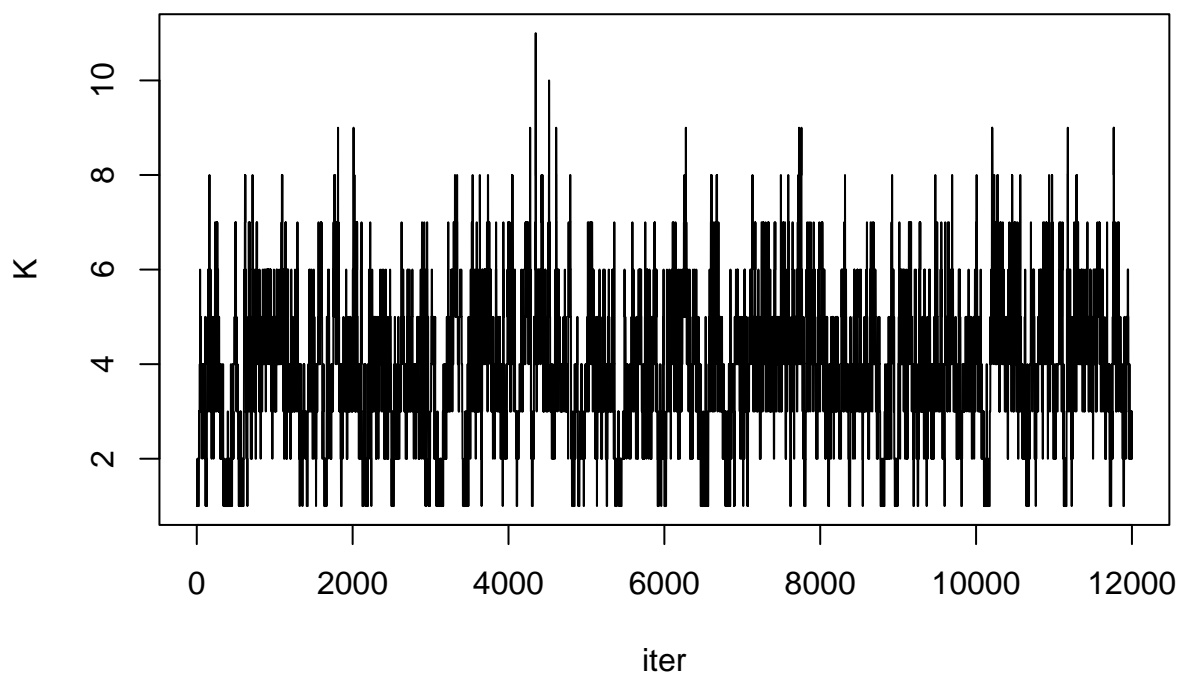
```
##### DEFINE HELPER FUNCTIONS#####
# running locally
# data_path = "//GENE.bst.rochester.edu/Projects/JKSTproj/BlueHive_Sim_Results/"

# running on server
data_path = "/projects/jklus/JKSTproj/BlueHive_Sim_Results/SummaryProposal"
```

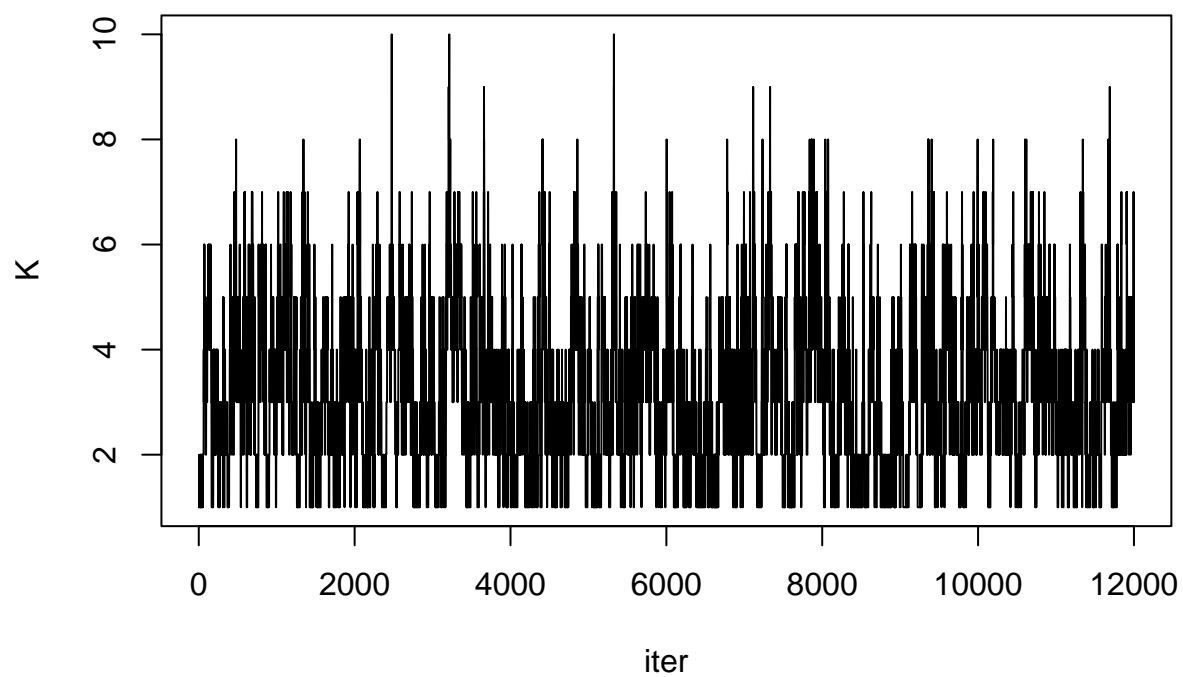
with SM

```
# n=30
sim_dir = "MODSUM_conjDEE_3close_n30_withSM_sim_results_2024_07_28" # found MAP(K)=16, truth is 3
result_wd = paste0(data_path, "/", sim_dir)
result_file = list.files(result_wd)
for(i in 1:10){
  output = readRDS(paste0(result_wd, "/", result_file[i]))
  # output$settings
  # dim(output$group_assign)
  iter_k = sapply(X = 1:nrow(output$group_assign),
                 FUN = function(x){
                   length(unique(output$group_assign[x,]))
                 })
  plot(x = 1:length(iter_k), y = iter_k, type = "l", xlab = "iter", ylab = "K",
       main = paste0("Traceplot of K, ", "Dataset ", i))
}
```

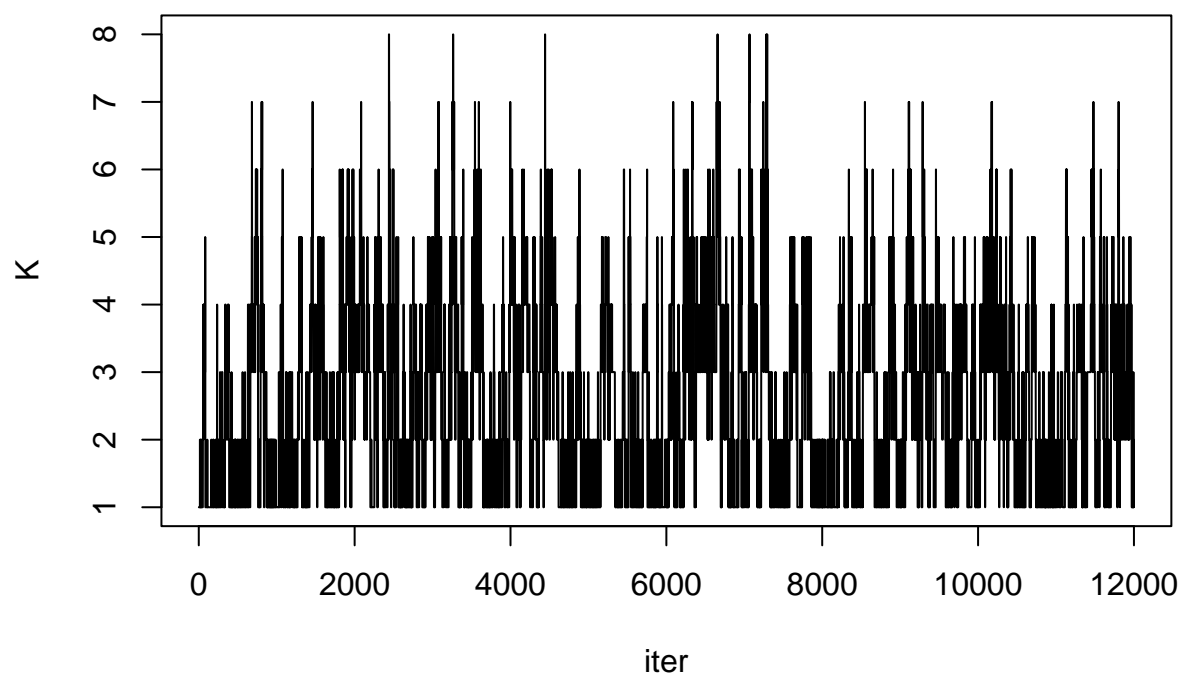
Traceplot of K, Dataset 1



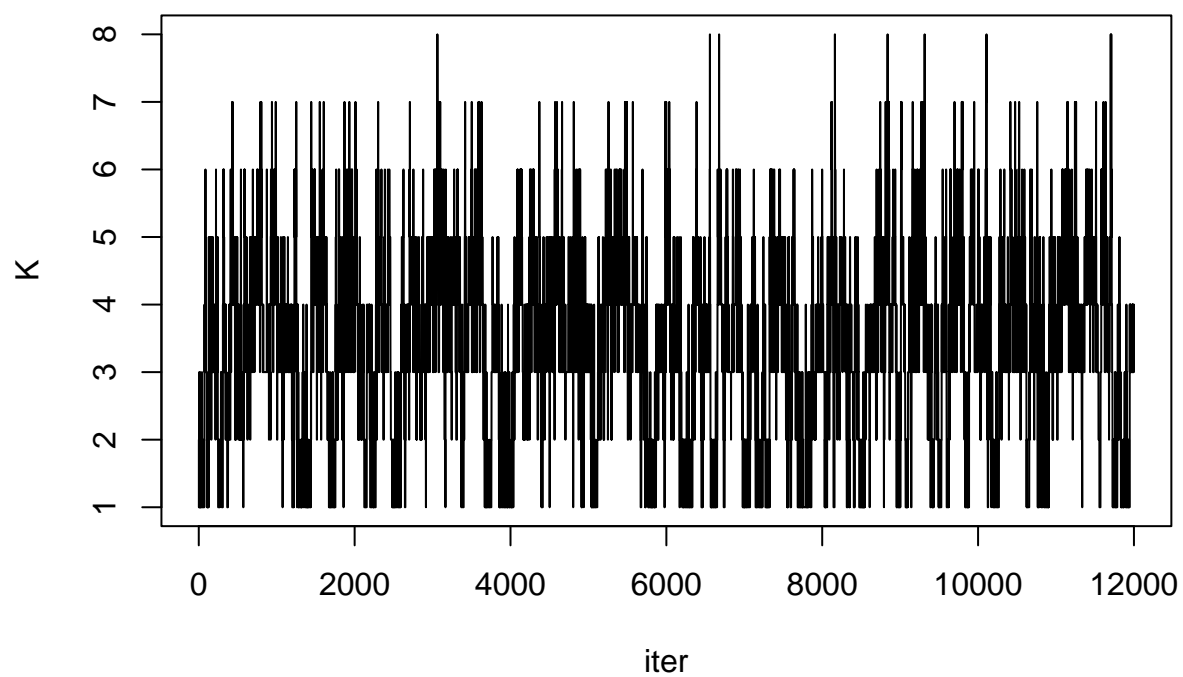
Traceplot of K, Dataset 2



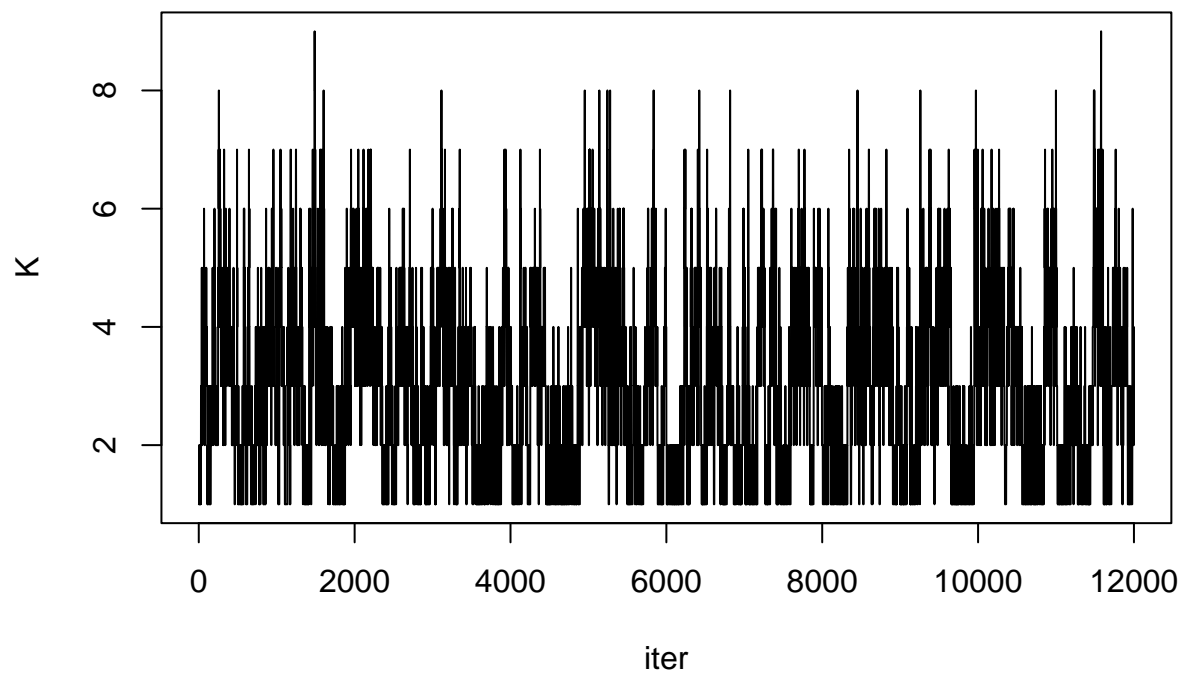
Traceplot of K, Dataset 3



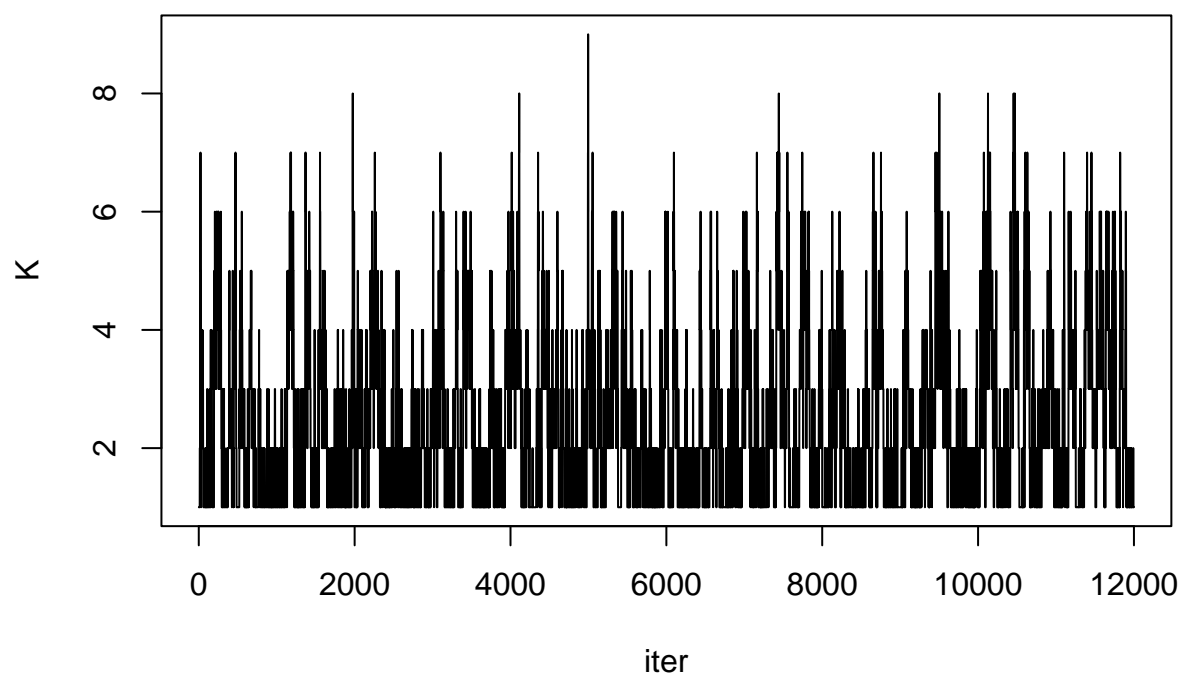
Traceplot of K, Dataset 4



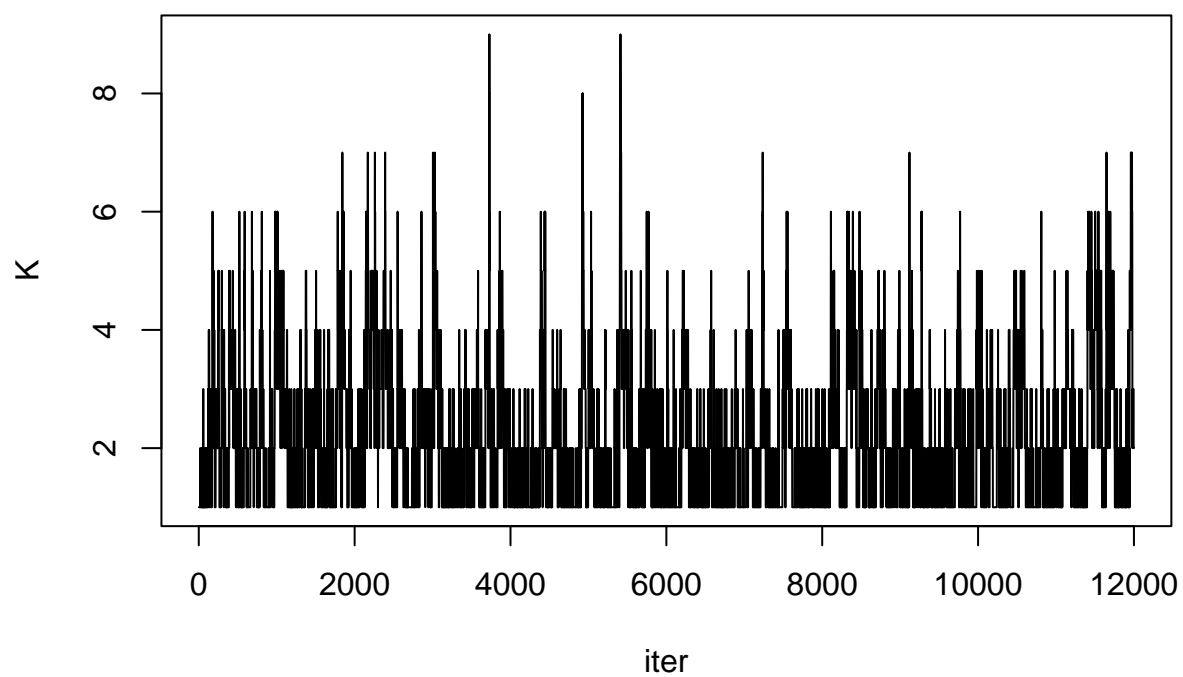
Traceplot of K, Dataset 5



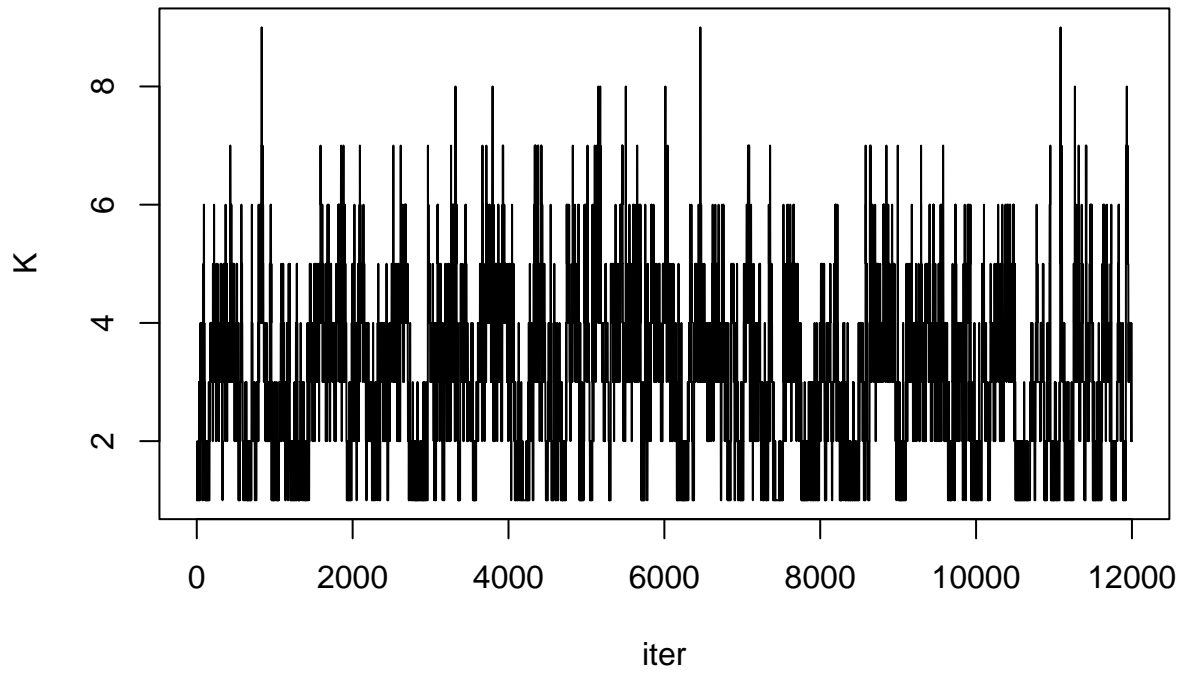
Traceplot of K, Dataset 6



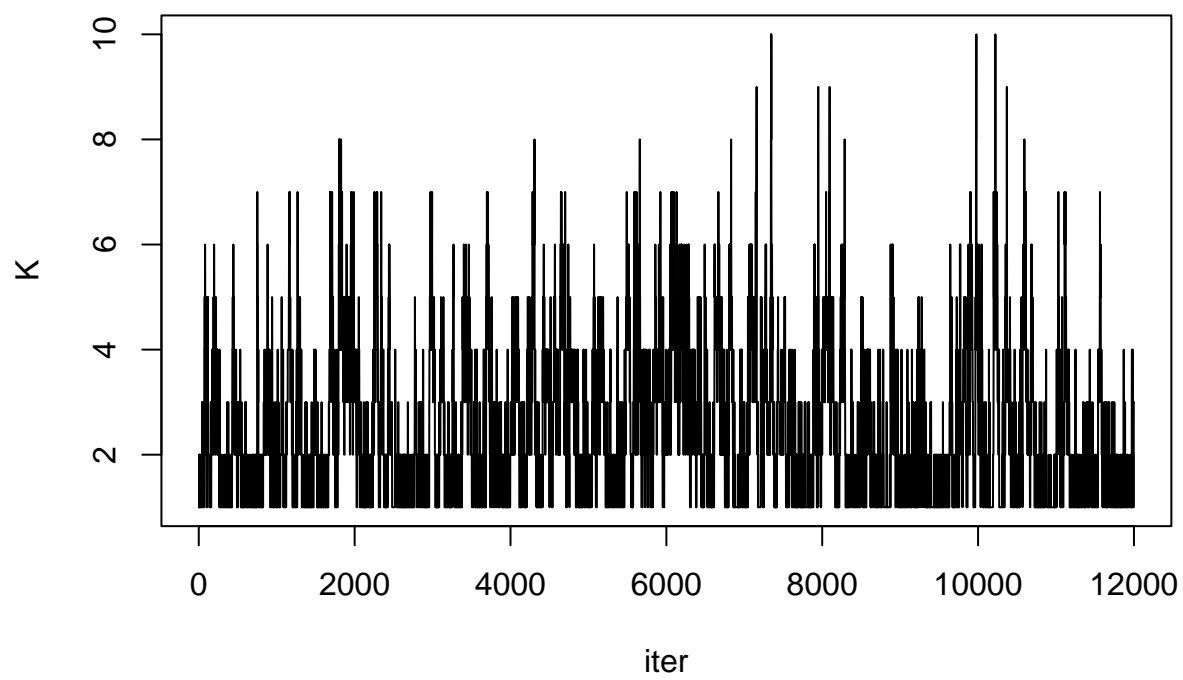
Traceplot of K, Dataset 7



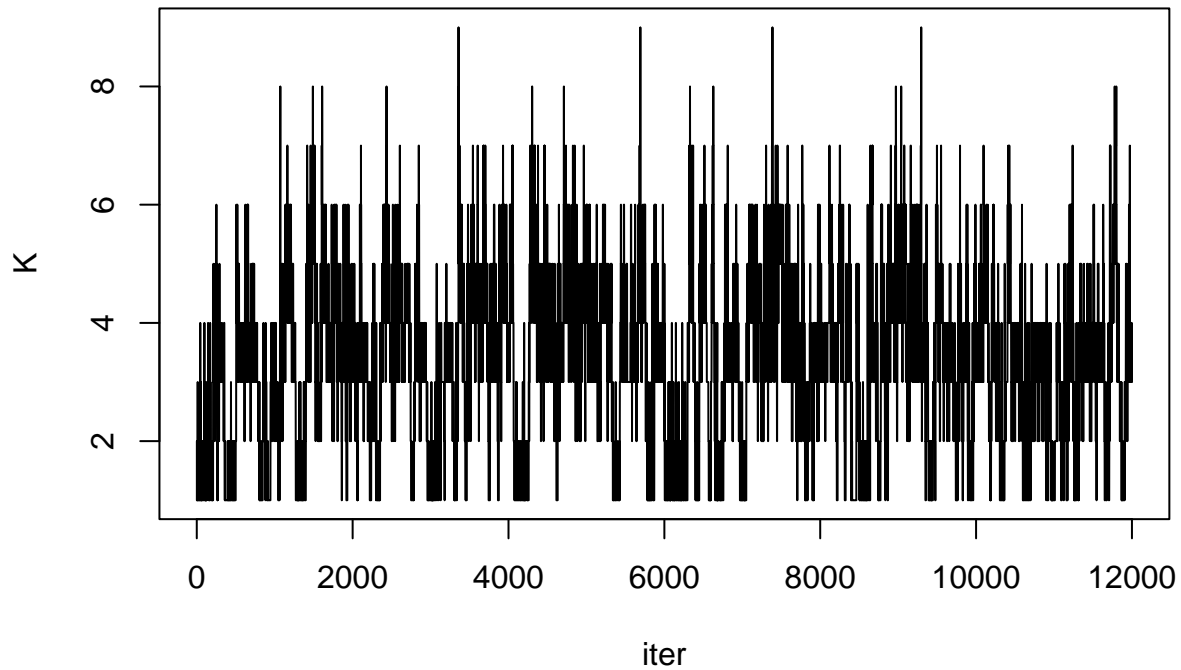
Traceplot of K, Dataset 8



Traceplot of K, Dataset 9

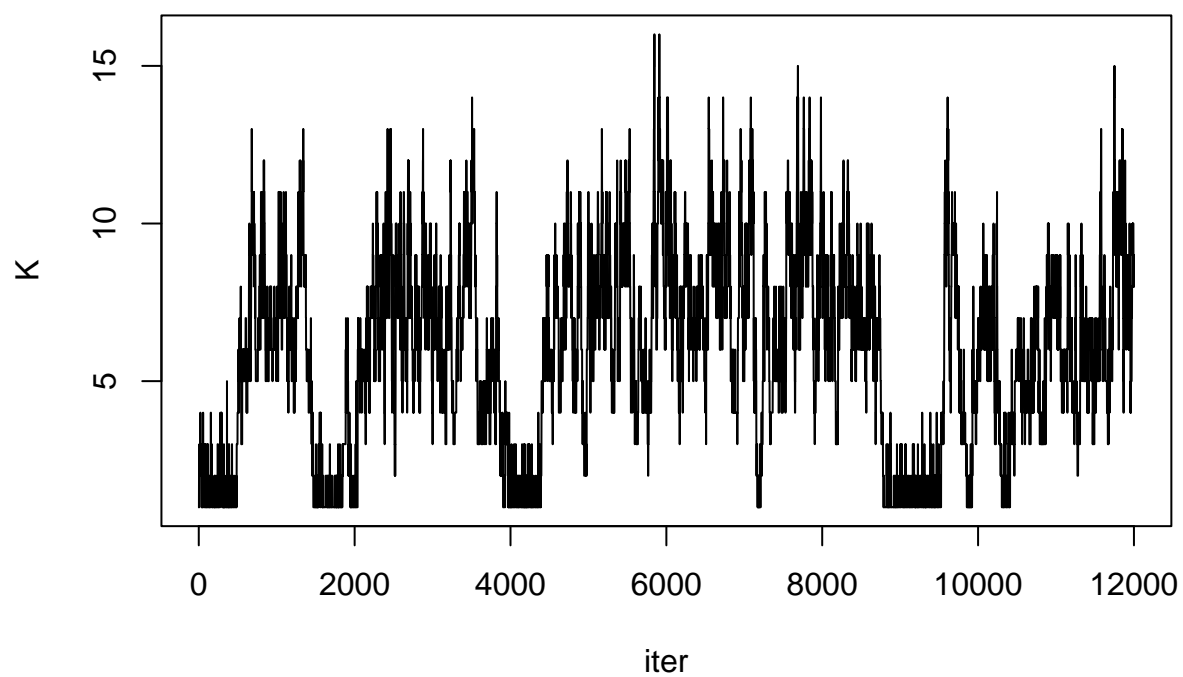


Traceplot of K, Dataset 10

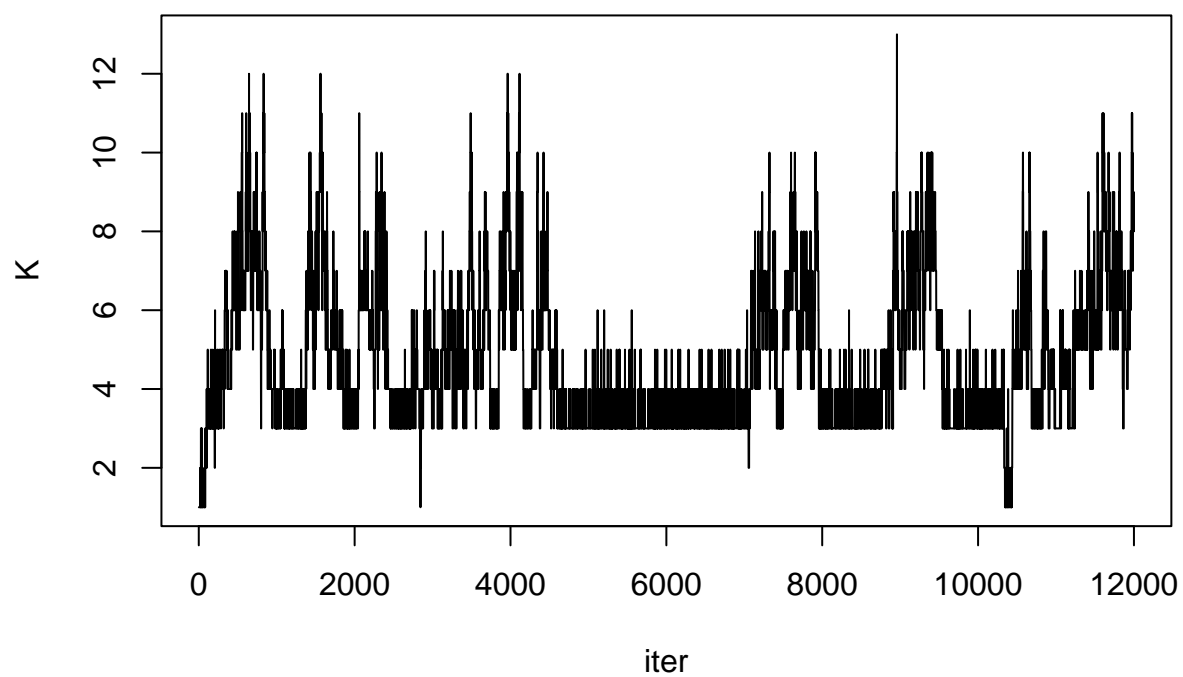


```
# n=100
sim_dir = "MODSUM_conjDEE_3close_n100_withSM_sim_results_2024_07_28" # found MAP(K)=16, truth is 3
result_wd = paste0(data_path, "/", sim_dir)
result_file = list.files(result_wd)
for(i in 1:10){
  output = readRDS(paste0(result_wd, "/", result_file[i]))
  output$settings
  dim(output$group_assign)
  iter_k = sapply(X = 1:nrow(output$group_assign),
    FUN = function(x){
      length(unique(output$group_assign[x,]))
    })
  plot(x = 1:length(iter_k), y = iter_k, type = "l", xlab = "iter", ylab = "K",
    main = paste0("Traceplot of K, ", "Dataset ", i))
}
```

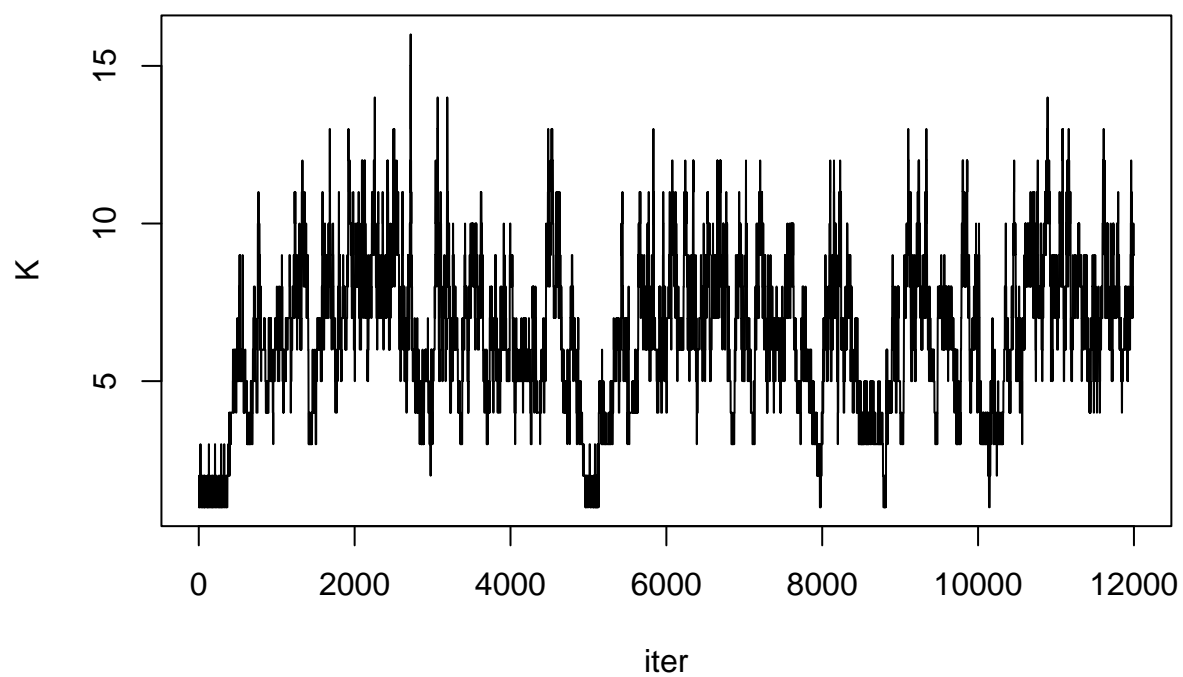
Traceplot of K, Dataset 1



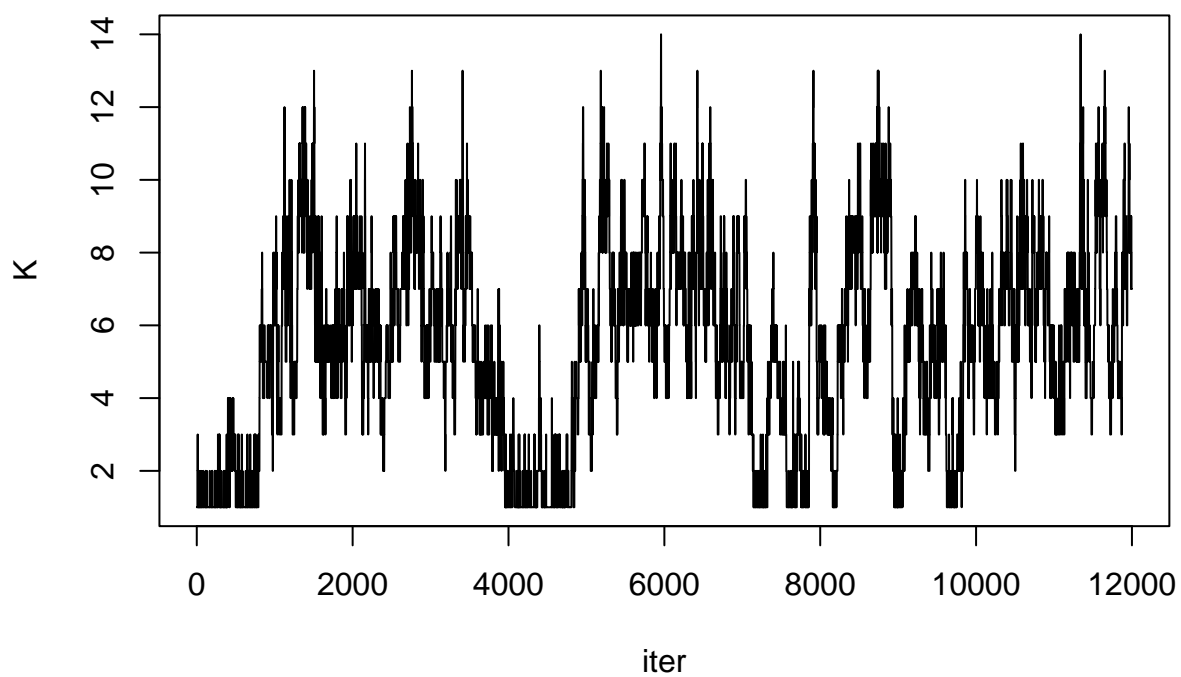
Traceplot of K, Dataset 2



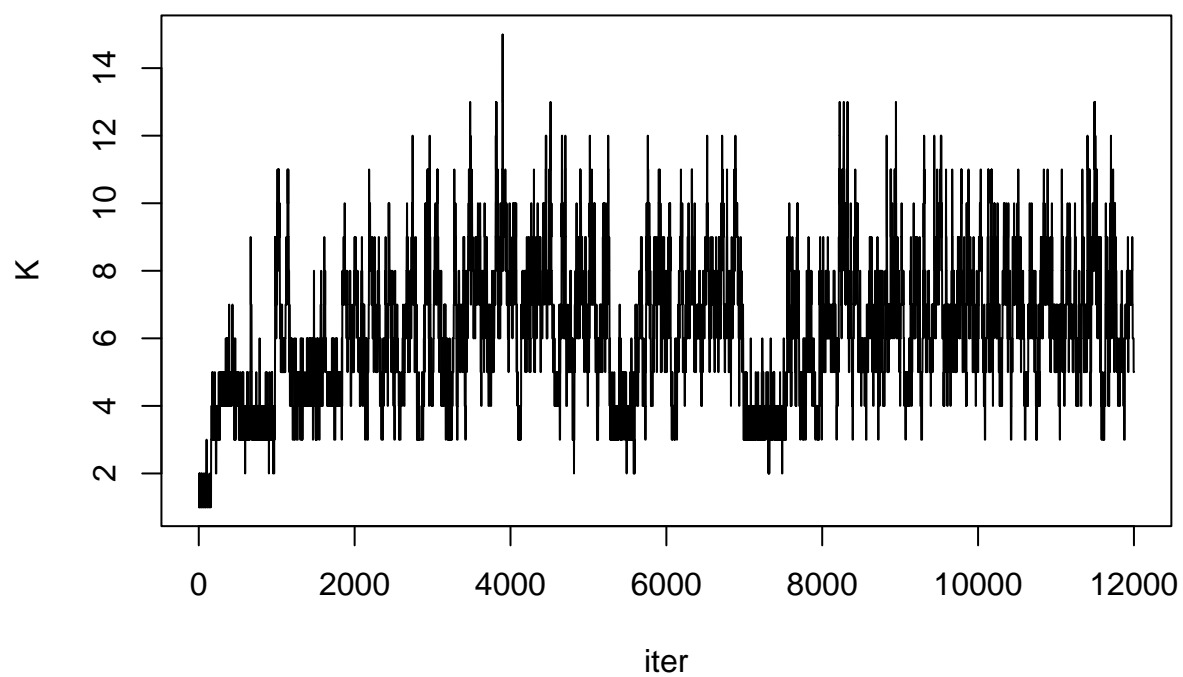
Traceplot of K, Dataset 3



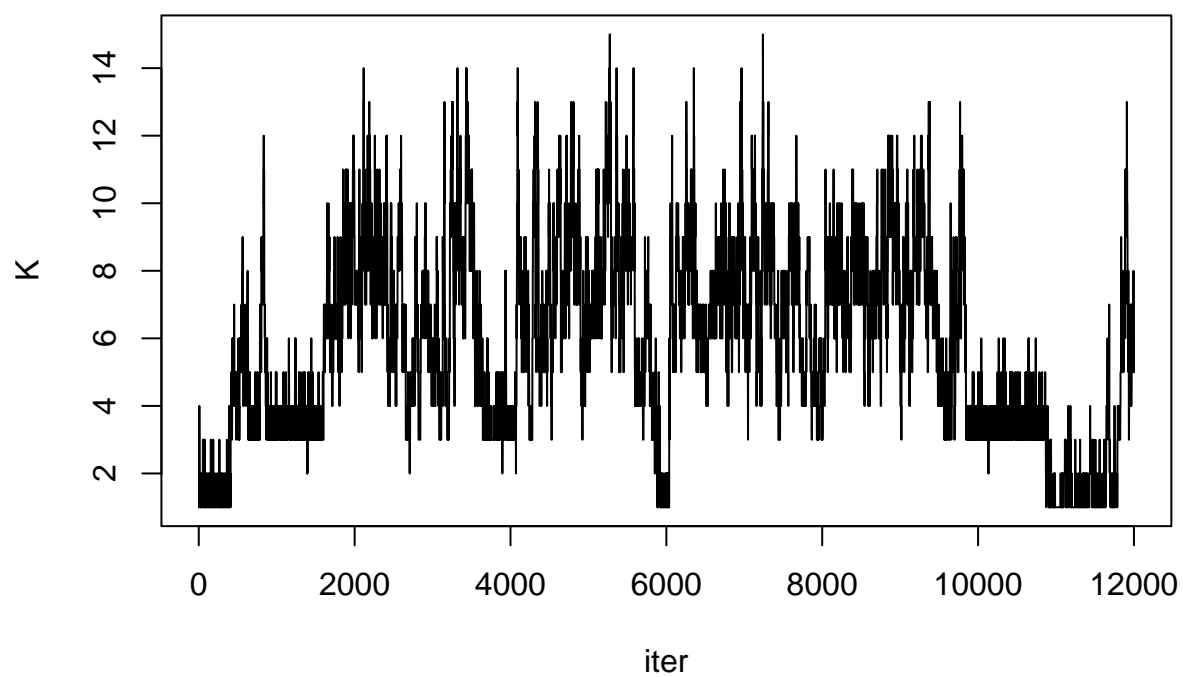
Traceplot of K, Dataset 4



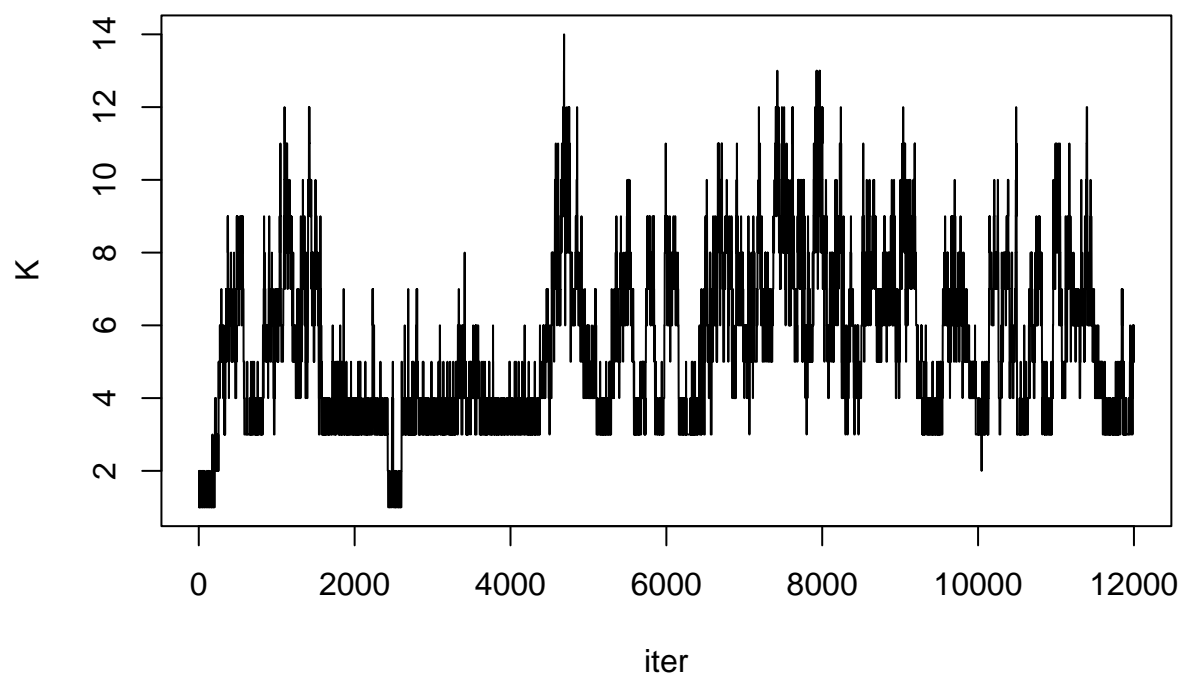
Traceplot of K, Dataset 5



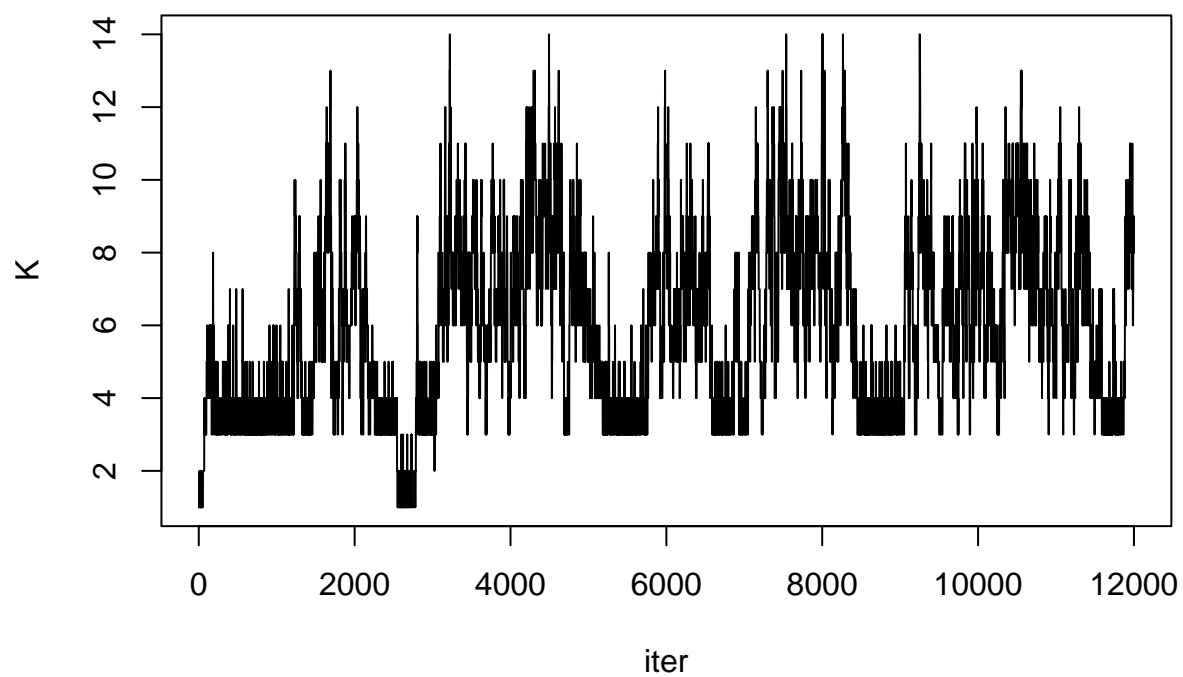
Traceplot of K, Dataset 6



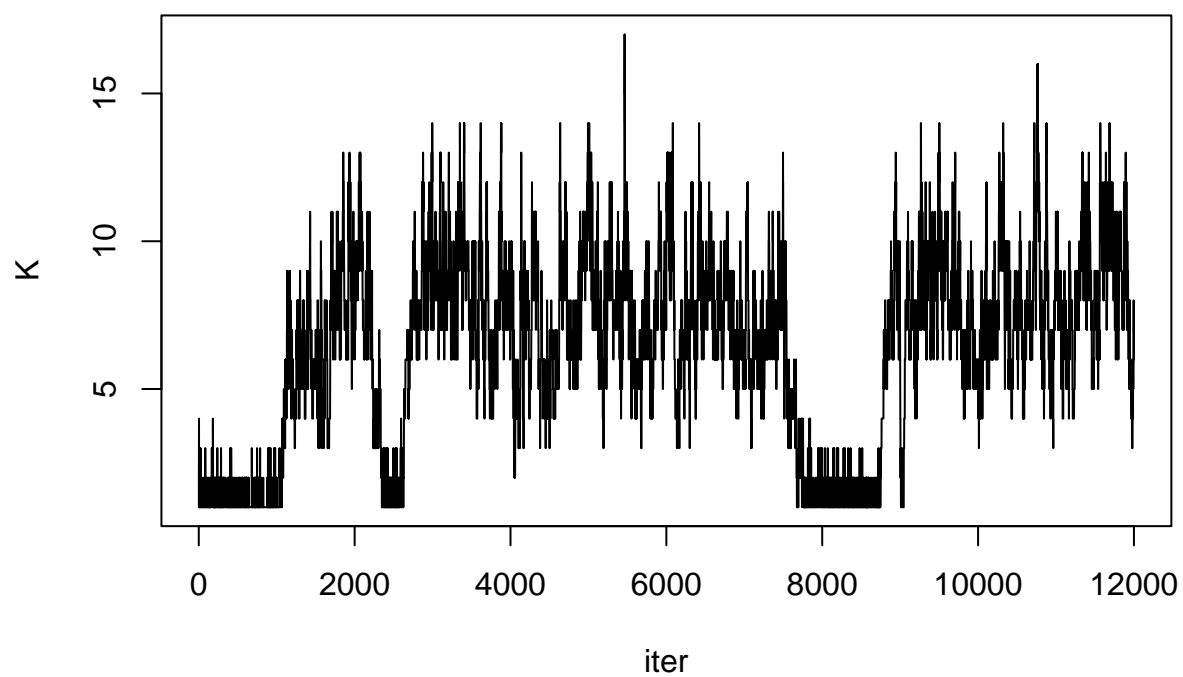
Traceplot of K, Dataset 7



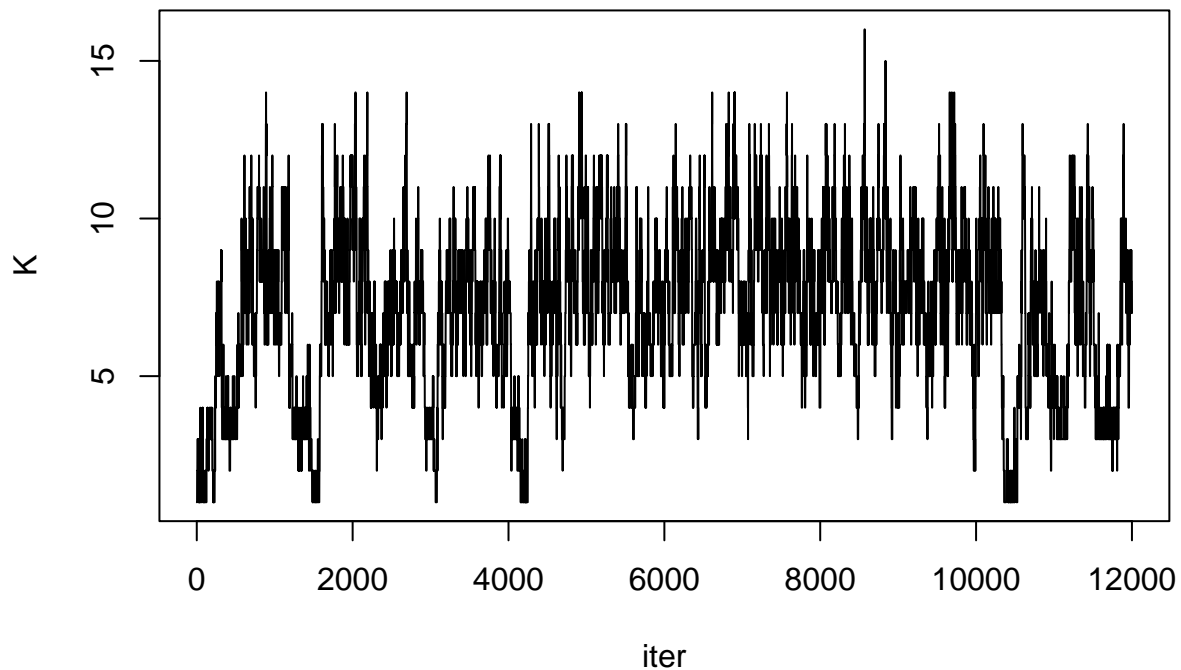
Traceplot of K, Dataset 8



Traceplot of K, Dataset 9



Traceplot of K, Dataset 10

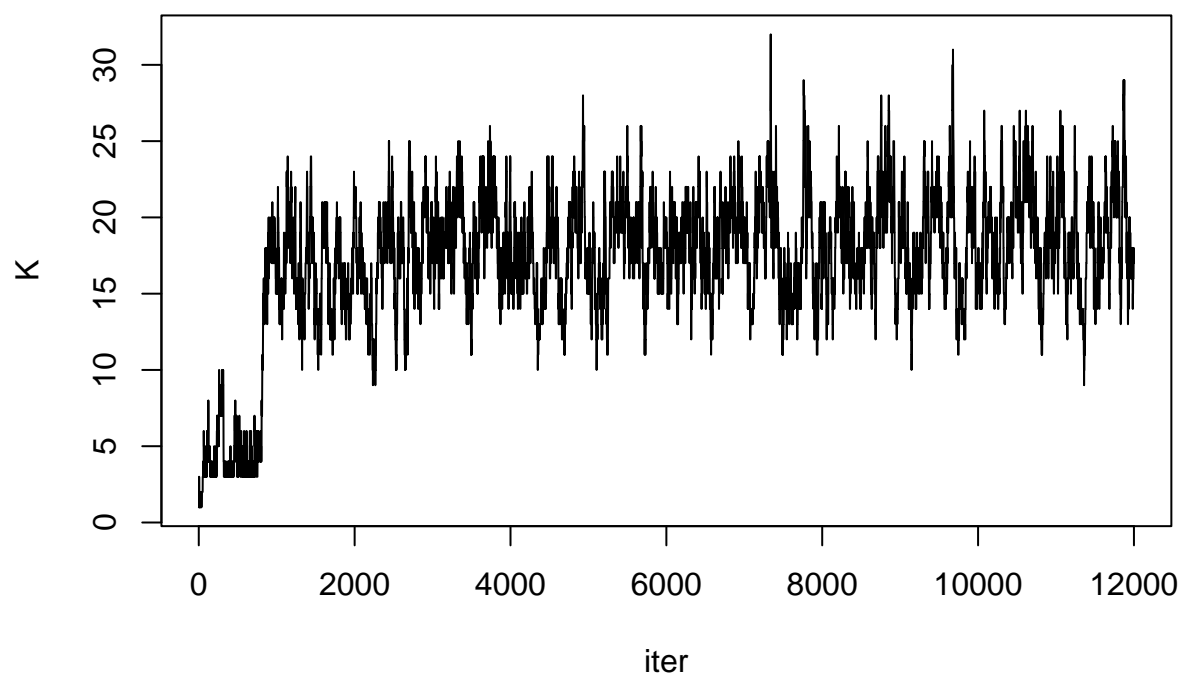


```
# n=300
sim_dir = "MODSUM_conjDEE_3close_n300_withSM_sim_results_2024_07_14" # found MAP(K)=16, truth is 3
result_wd = paste0(data_path, "/", sim_dir)
result_file = list.files(result_wd)

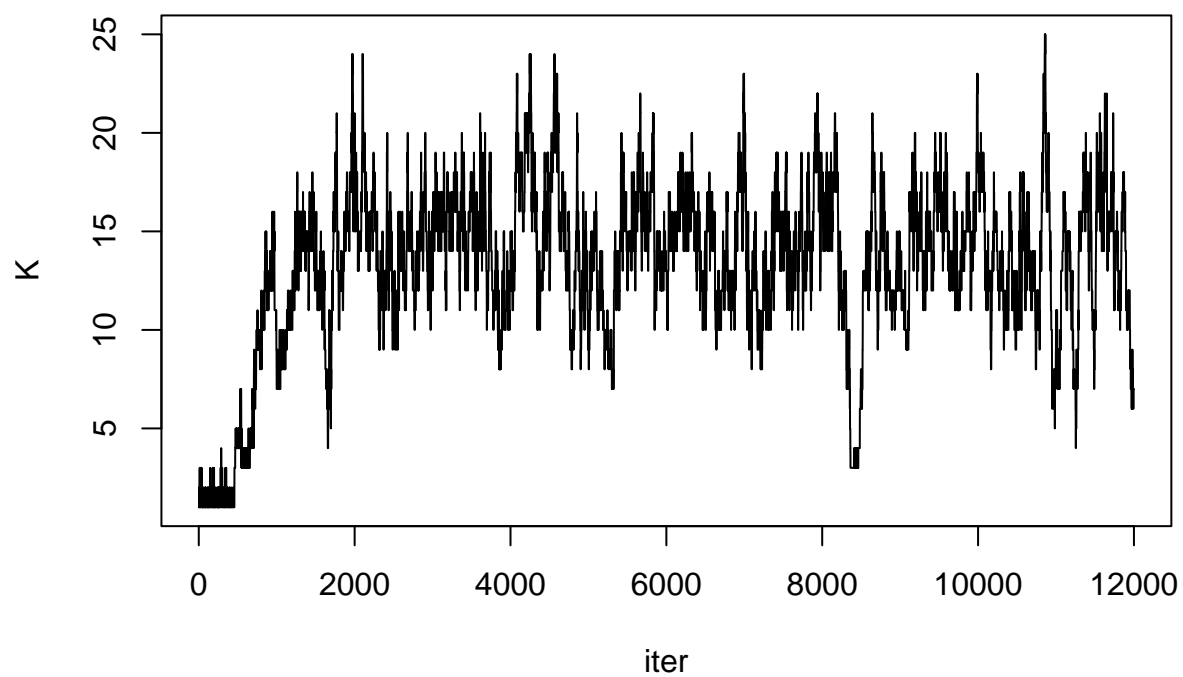
for(i in 1:10){

  output = readRDS(paste0(result_wd, "/", result_file[i]))
  output$settings
  dim(output$group_assign)
  iter_k = sapply(X = 1:nrow(output$group_assign),
                  FUN = function(x){
                    length(unique(output$group_assign[x,]))
                  })
  plot(x = 1:length(iter_k), y = iter_k, type = "l", xlab = "iter", ylab = "K",
       main = paste0("Traceplot of K, ", "Dataset ", i))
}
```

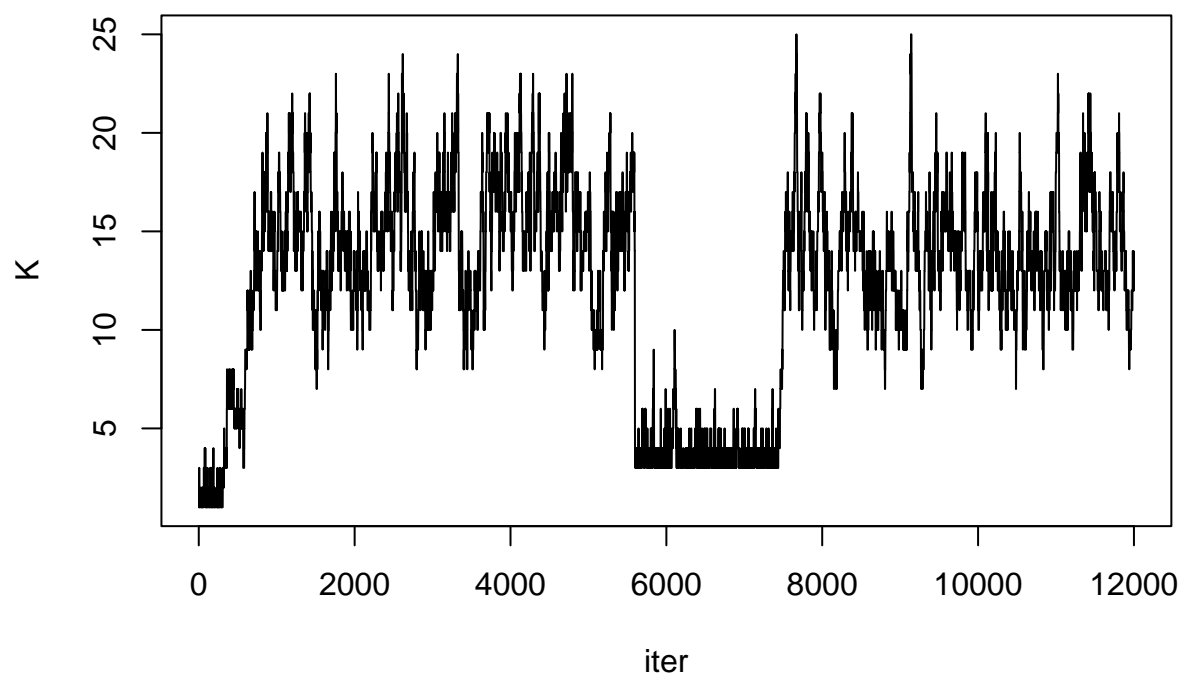
Traceplot of K, Dataset 1



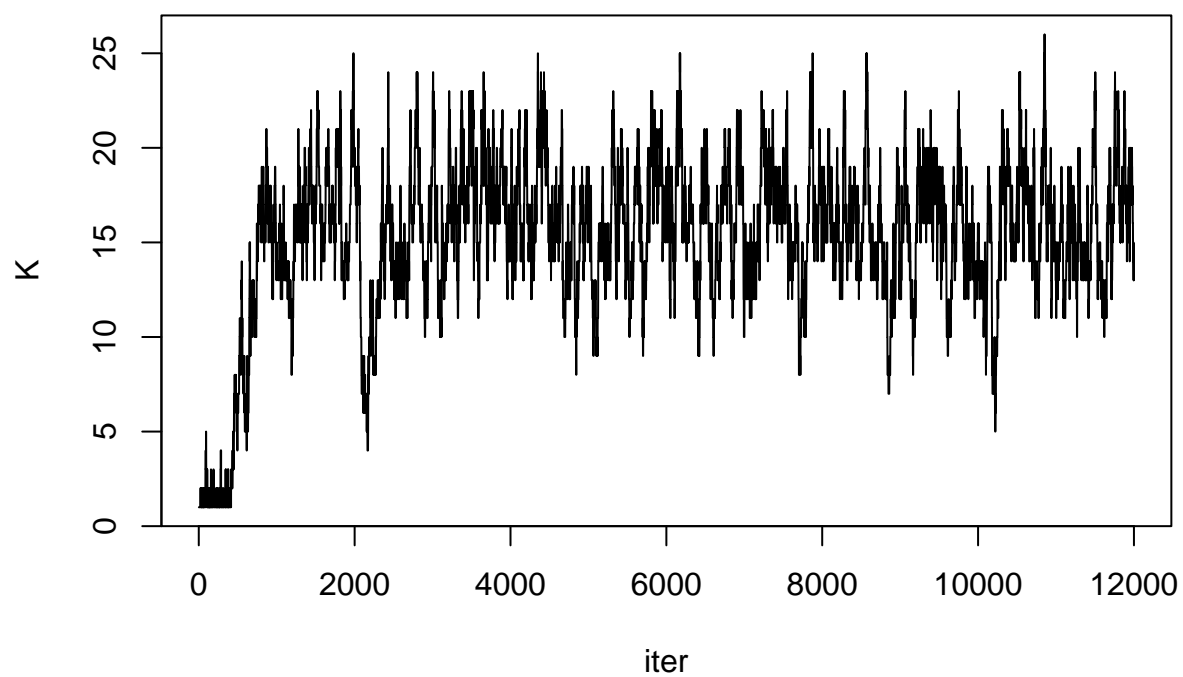
Traceplot of K, Dataset 2



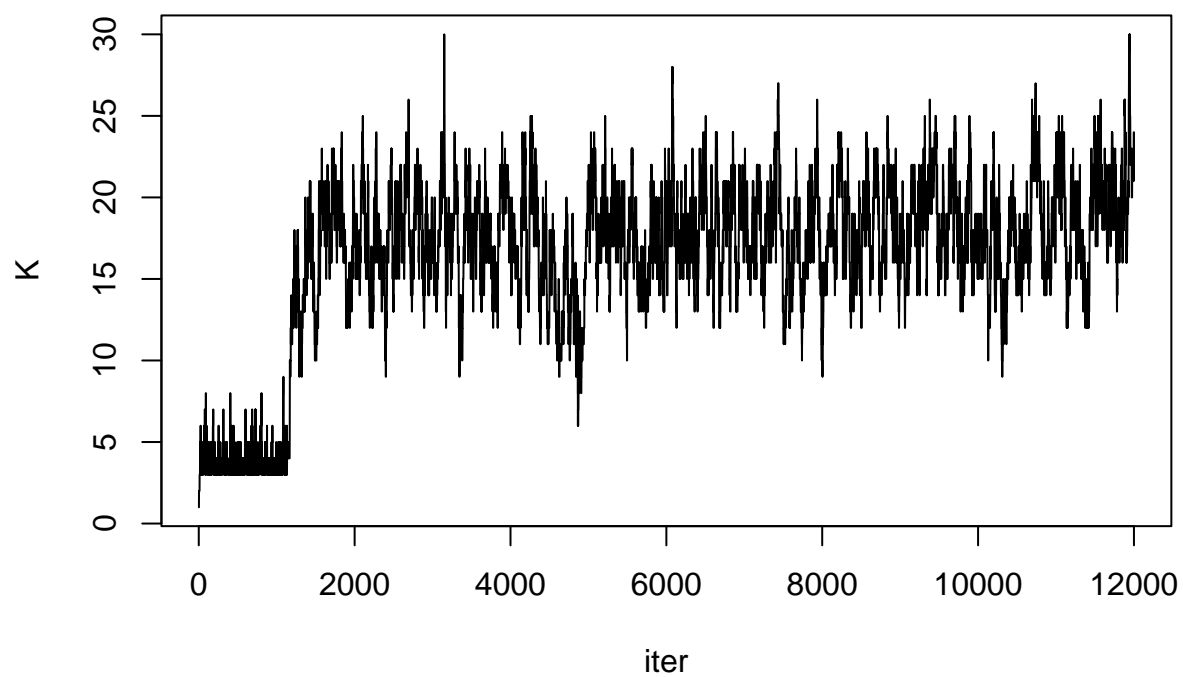
Traceplot of K, Dataset 3



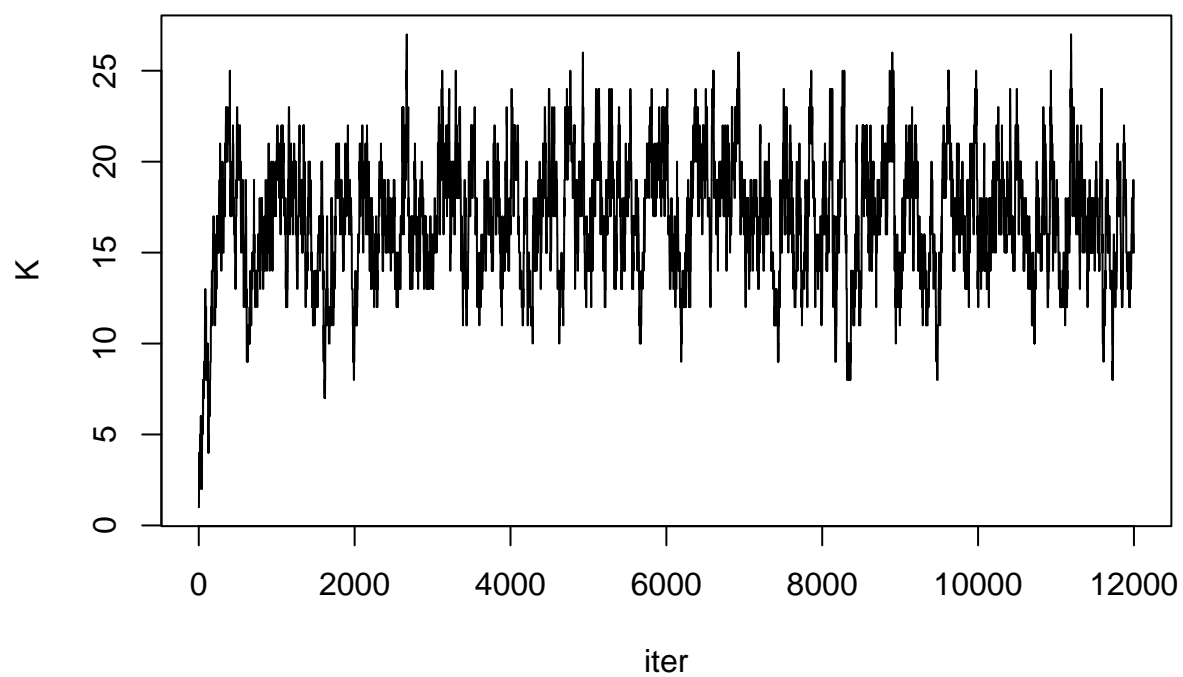
Traceplot of K, Dataset 4



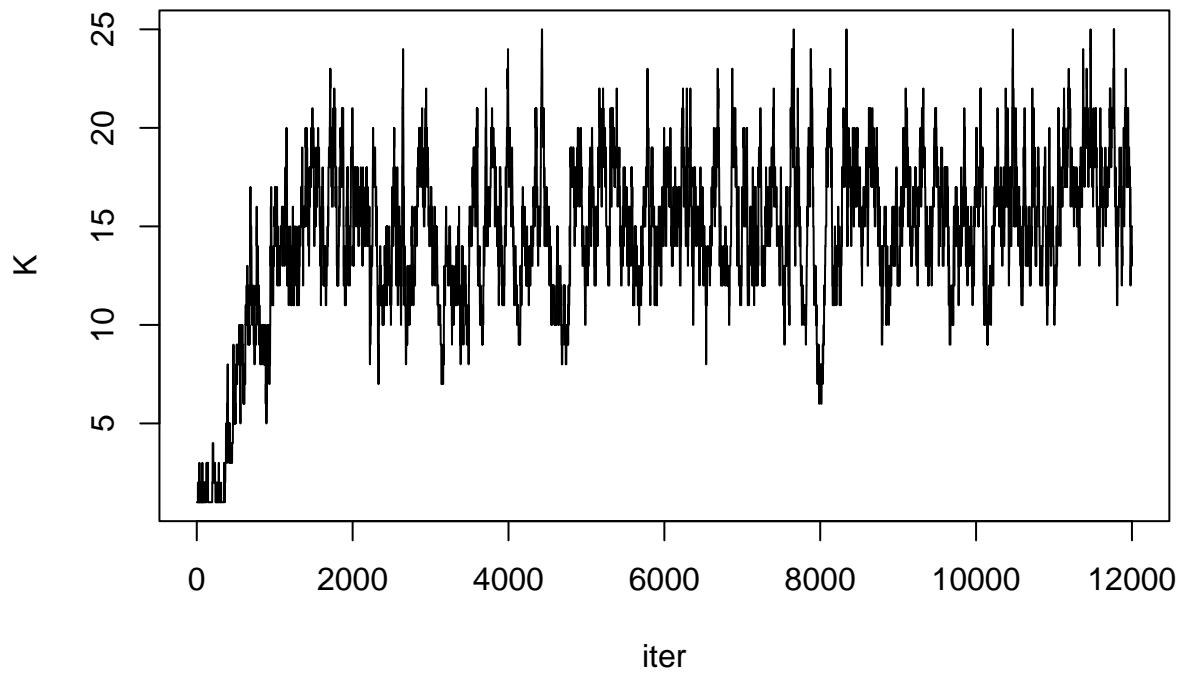
Traceplot of K, Dataset 5



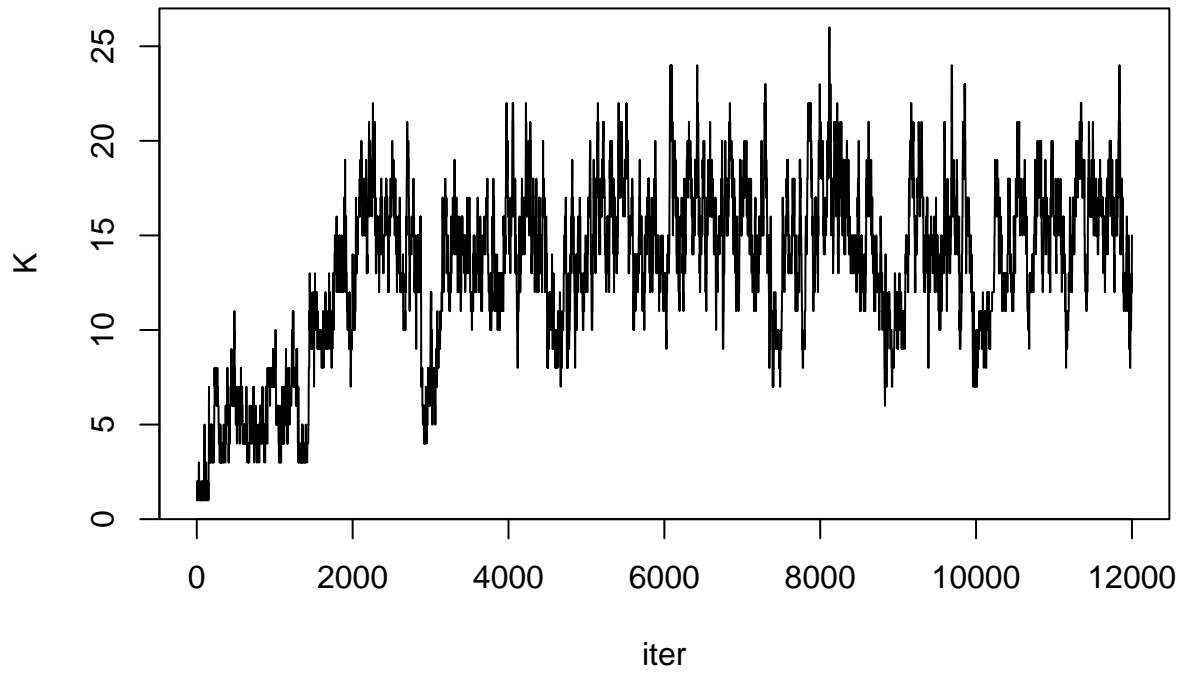
Traceplot of K, Dataset 6



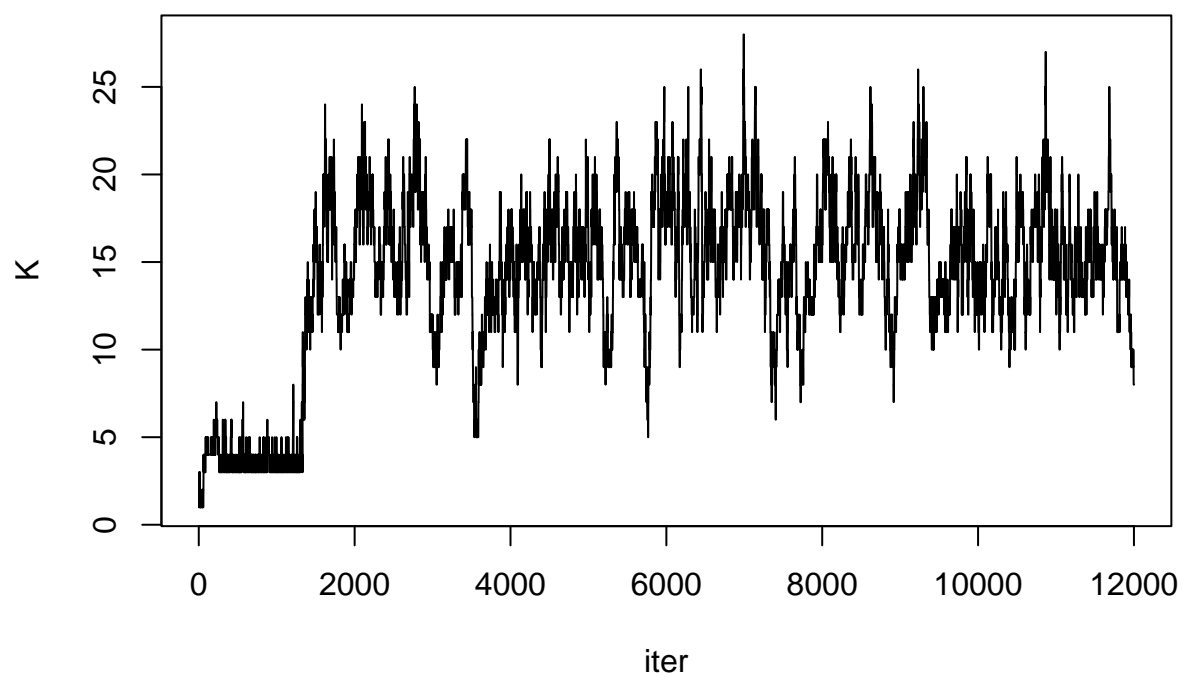
Traceplot of K, Dataset 7



Traceplot of K, Dataset 8



Traceplot of K, Dataset 9



Traceplot of K, Dataset 10

