simulation_week8

Tingyu Zhu

11/29/2020

```
library(tidyverse)
library(purrr)
library(ggplot2)
library(here)
library("KernSmooth")
devtools::load_all()
set.seed(1222)
```

Beaver's temprature KDE

Here, we try to apply our KDE method to a real dataset: Beaver1, which records the beaver's body temperature during a day.

We try to estimate its distribution by two types of kernel functions. The first type of bandwidths are selected using the "plug-in" method under corresponding kernel functions. And using the "Guassian optimal bandwidth" as the second type of bandwidth.

```
# ?beaver1

# three types of bandwidth
h_plug_gua <- dpik(beaver1$temp,kernel = "normal")
h_plug_bi <- dpik(beaver1$temp,kernel = "biweight")
h_opt <- 1.06*sd(beaver1$temp)*length(beaver1$temp)^(-0.2)

# generate parameter combinatons
params_big <- list(
  kernel_type = c("normal","biweight"),
  bandwidth_type = c("gua_plug"=h_plug_gua,"bi_plug"=h_plug_bi,"gua_opt"=h_opt)
  )

est_big <- cross_df(params_big)

# generate grid points corresponding to the range of data
grid = seq(min(beaver1$temp)-0.5, max(beaver1$temp)+0.5, length.out=512)

# generate the estimated desities and save the corresponding grid points
est_big <- est_big %>%
  mutate(
    f_ests = map2(.x=kernel_type, .y=bandwidth_type,
                  ~KDE_est(beaver1$temp,ker=.x,h=.y,grid=grid)$f_est),
    grid =  map2(.x=kernel_type, .y=bandwidth_type,
```

```
                  ~KDE_est(beaver1$temp,ker=.x,h=.y,grid=grid)$grid)
  )

# transform the values of bandwidth to the character names.
est_big$bandwidth_type <- ifelse(abs(est_big$bandwidth_type - 0.053) < 0.001,
                                 "plugin_gua",
                                 ifelse(abs(est_big$bandwidth_type -0.138)<0.001,
                                 "plugin_bi", "opt_gua"))
# save the results
write_rds(est_big, here("results", "beaver.rds"))
```

Analysis

Analysis of the beaver's data is included in the `pre_plots` files.