

Summary__week6

Tingyu Zhu

12/6/2020

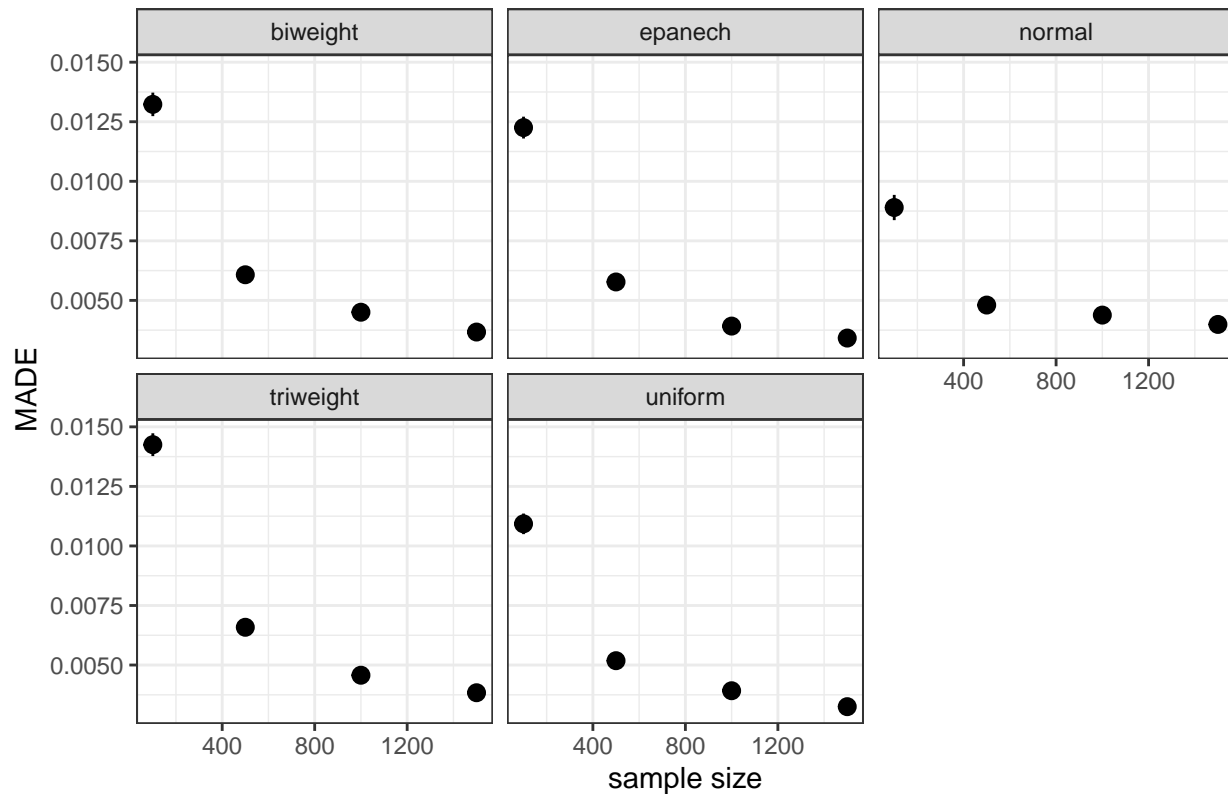
```
library(tidyverse)
library(here)
library(ggplot2)
library(purrr)
devtools::load_all()
```

Different sample sizes & kernels

```
made_comapre <- read_rds(here("results", "week6-sim-plot.rds"))
#png(here("plots", "KDE_small_size.png"),width=1600, height=900)

made_comapre %>%
  ggplot(aes(x = n, y = made_mean)) +
    geom_pointrange(aes(
      ymin = made_mean - 1.96*made_sd/sqrt(100),
      ymax = made_mean + 1.96*made_sd/sqrt(100))) +
    geom_point() +
    facet_wrap(~ kernel_type)+
    theme_bw() +ggtitle("MADEs of different kernels & sample sizes")+
    xlab("sample size")+
    ylab("MADE")+
    theme( legend.title = element_text(size = 14),
           legend.text = element_text(size = 14),
           plot.title = element_text(size=14))
```

MADEs of different kernels & sample sizes

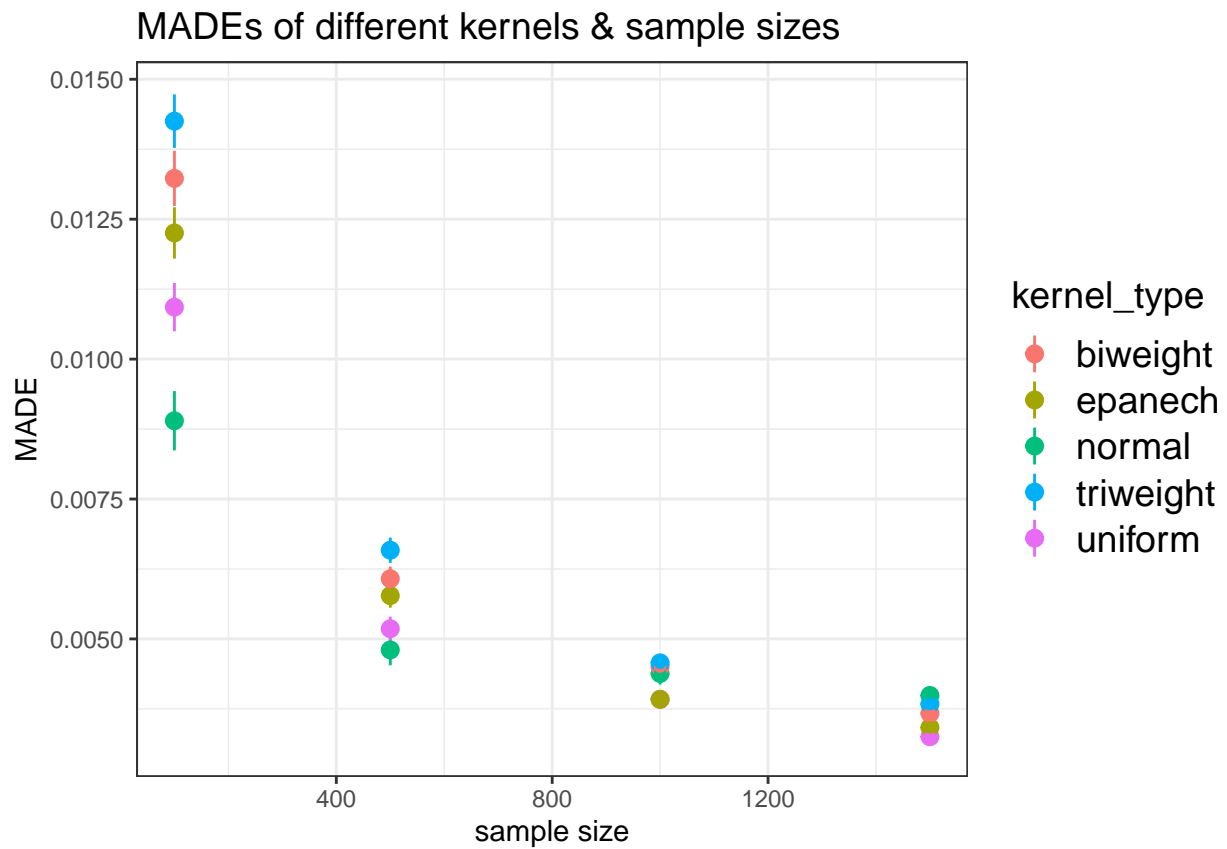


```
#dev.off()
```

Results shows that larger sample size will lead to smaller MADEs (more accurate estimation results).

```
#png(here("plots", "KDE_small_kers.png"),width=1600, height=900)
```

```
made_comapre %>%
  ggplot(aes(x = n, y = made_mean, color=kernel_type)) +
    geom_pointrange(aes(
      ymin = made_mean - 1.96*made_sd/sqrt(100),
      ymax = made_mean + 1.96*made_sd/sqrt(100))) +
    geom_point() +
    theme_bw() + ggtitle("MADEs of different kernels & sample sizes") +
    xlab("sample size") +
    ylab("MADE") +
    theme( legend.title = element_text(size = 14),
           legend.text = element_text(size = 14),
           plot.title = element_text(size=14))
```



```
#dev.off()
```

And with the same sample size and the same bandwidth, the MADEs are quite similar to each other. Thus, the choice of kernel functions is not very crucial.