Week 3 report

sigmadt

November 23, 2022

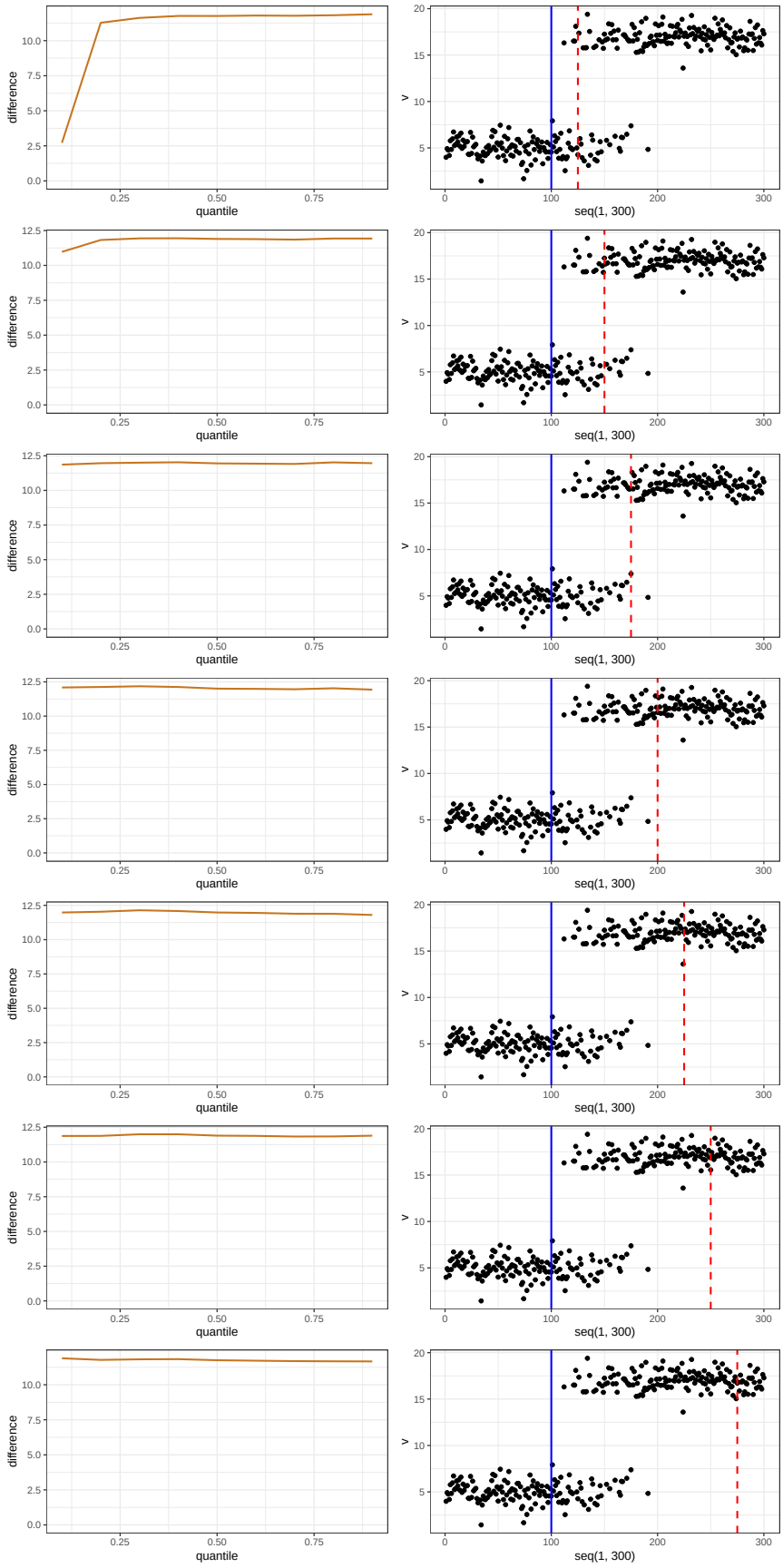
Diploma discussion: baseline for algorithm

Case: Normal disrribution

Plot shift function for partition with fixed step

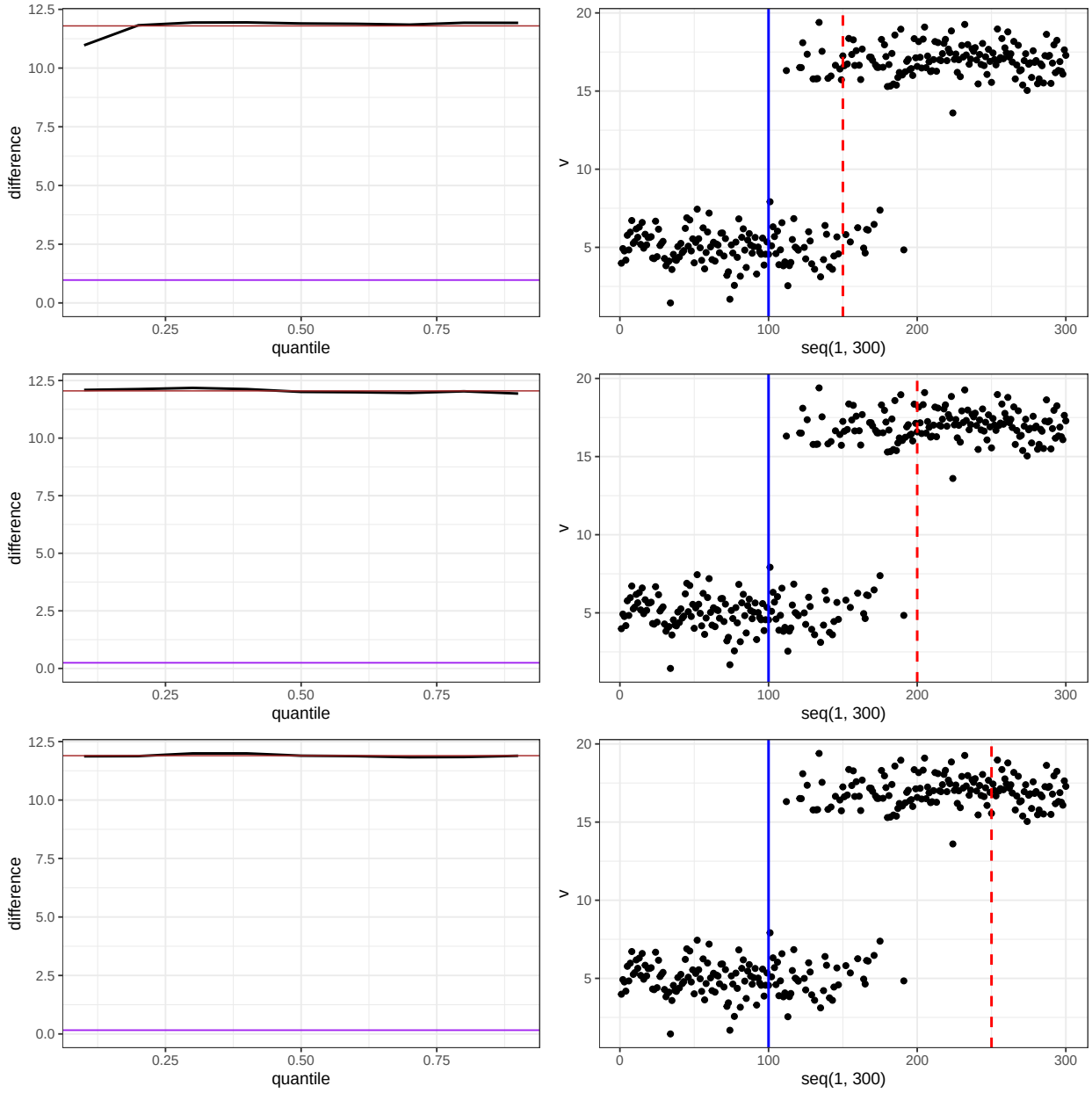
Let's suppose we have an oracle that give us true change point.

Idea is simple: we should make a partition of sample, compare sample to the left of CP and to the right of partition point using shift function. At some point we could see stabilization of difference of group of quantile, that's the sign we need to stop traversing. The partition point between the moments when difference stop changing is our second point two construct *interval from given CP*.


difference
quantile
v
seq(1, 300)

Finding cost function

Let's try to plot graph for `step=50`:



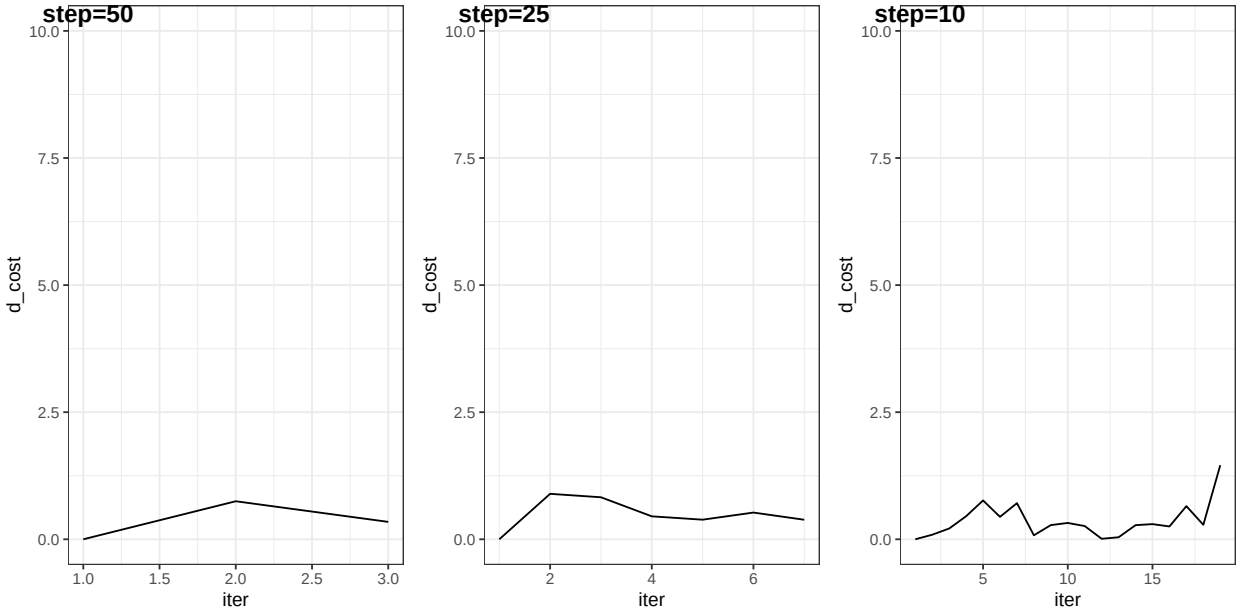
It is clear that mean doesn't give us information, but maxmin difference does. Let's construct our baseline using this cost function:

$$c(s_1, s_2) = |\max_i \Delta q_i - \min_j \Delta q_j|,$$

where Δq_i – difference of Harrell-Davis estimation between i-th quantile between s_1 and s_2.

Next thing is to determine where to stop. Depending on pictures we would like to stop on second iteration since third iteration keeps cost function the same. So, we can check growth $\frac{c_{curr} - c_{prev}}{c_{prev}}$ of $c(s_1, s_2)$ after each iteration and if it is $< \varepsilon$ we will keep the best answer.

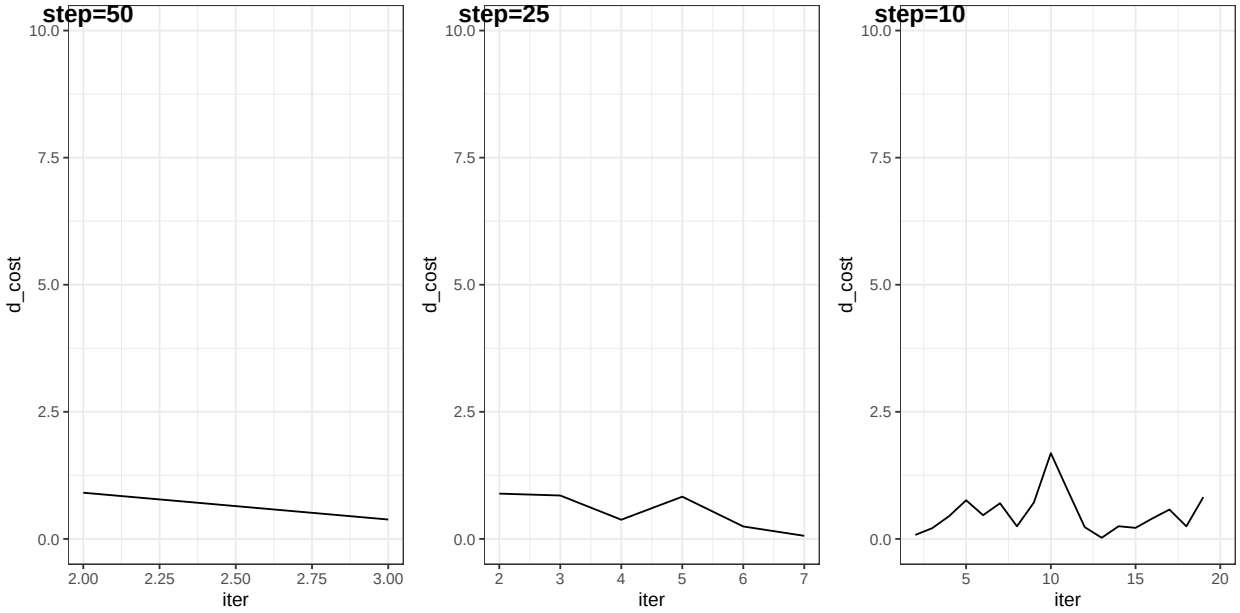
Picture of adjacent changes:



Imporving cost function I

It is clear that for this specific case we need to track only first 3-5 quantile. Let's recalculate and see the results:

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
## Removed 1 row(s) containing missing values (geom_path).
## Removed 1 row(s) containing missing values (geom_path).
```



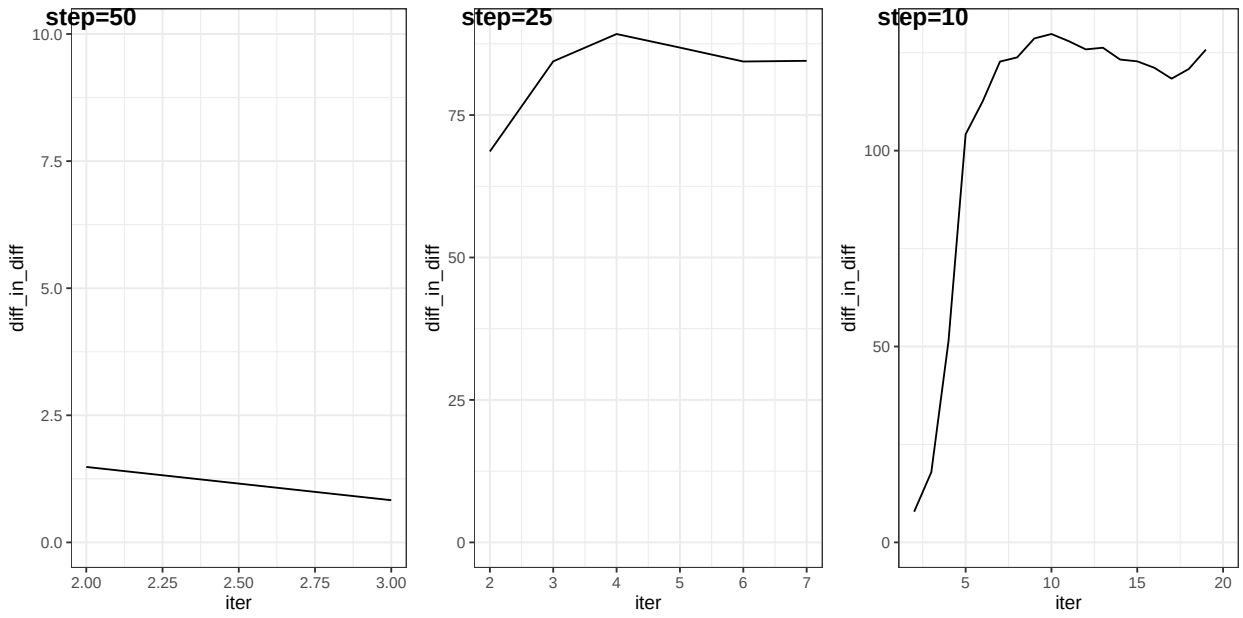
Diffrenet approach

Let's consider this:

$$c(s_1, s_2) = \sum (\Delta[\Delta q_i])^2,$$

Plot:

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
## Removed 1 row(s) containing missing values (geom_path).
## Removed 1 row(s) containing missing values (geom_path).
```



.