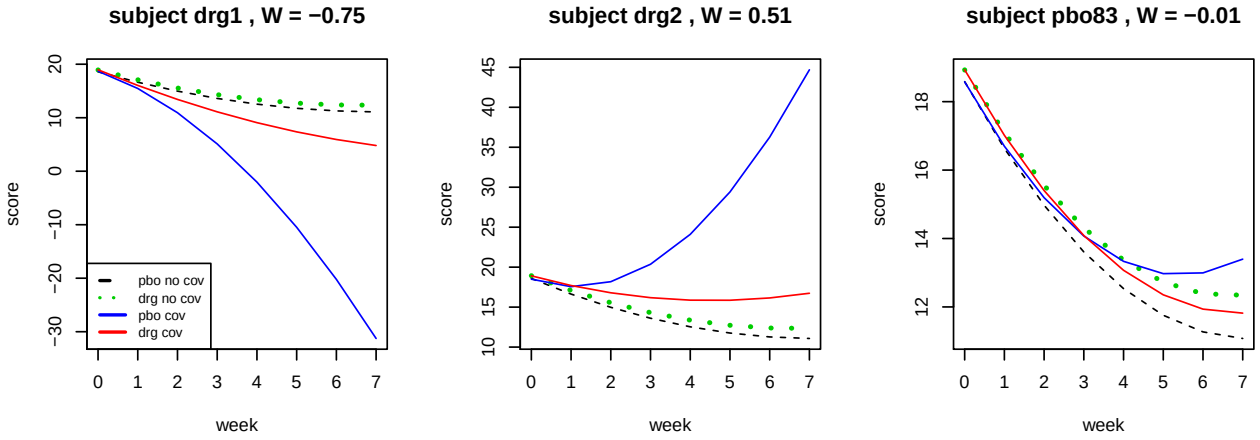
measurement of performance

2020-01-26

We have calculated the purity and the α and studied their estimation performance by simulations. Furthermore, we would like to check whether we can predict a patient group assignment with the help of covariates.



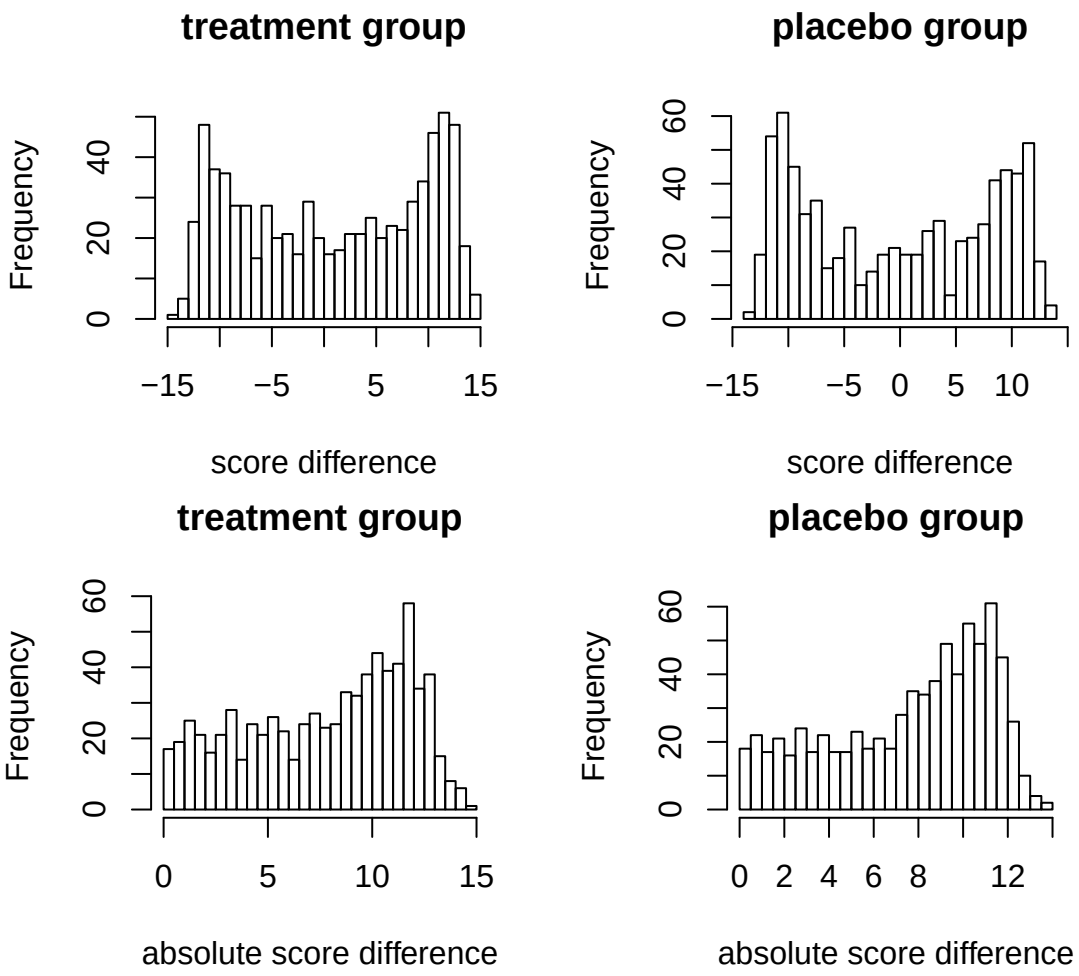
For the above plots, the effect of the treatment depends on the covariates. Without the consideration of covariates, the trajectories of the coefficients are the black line and the green line, separately. When the covariate $|W| > 0$, the differences between the two trajectories get larger. However, when W is close to 0, the difference does not change a lot.

How could we measure the difference between the two trajectories? Firstly, we would like to use the score difference in the last week.

We first generate a large dataset with n = 1000 in each group. The dataset is then randomly separated into a training set and a testing set, with $n1 = 1500, n2 = 500$, respectively. The training set is used to calculate the score difference for the treatment group and the placebo group.

- The optim function is applied to find the $\hat{\alpha}$ that maximizes the purity.
- The $\hat{W} = \hat{\alpha}'x$ is calculated.
- The LME models are fitted for the treatment group and the placebo group.
- The $\hat{\beta}_1, \hat{\Gamma}_1, \hat{D}_1, \hat{\beta}_2, \hat{\Gamma}_2, \hat{D}_2$ are estimated.
- The coefficient for each subject can be calculated with $z_{i1} = \hat{\beta}_1 + \hat{\Gamma}_1 w_i$,$z_{i2} = \hat{\beta}_2 + \hat{\Gamma}_2 w_i$
- The quadratic trajectory is fitted by using the coefficient calculated above.
- The score difference is calculated at week 8.

The histogram of the score difference for treatment group and for placebo group



The mean value and the mean absolute values of the s_1 are

```
round(mean(rules1),3)
```

```
## [1] 0.872
```

```
round(mean(abs(rules1)),3)
```

```
## [1] 7.813
```

The mean value and the mean absolute values of the s_2 are

```
round(mean(rules2),3)
```

```
## [1] -0.121
```

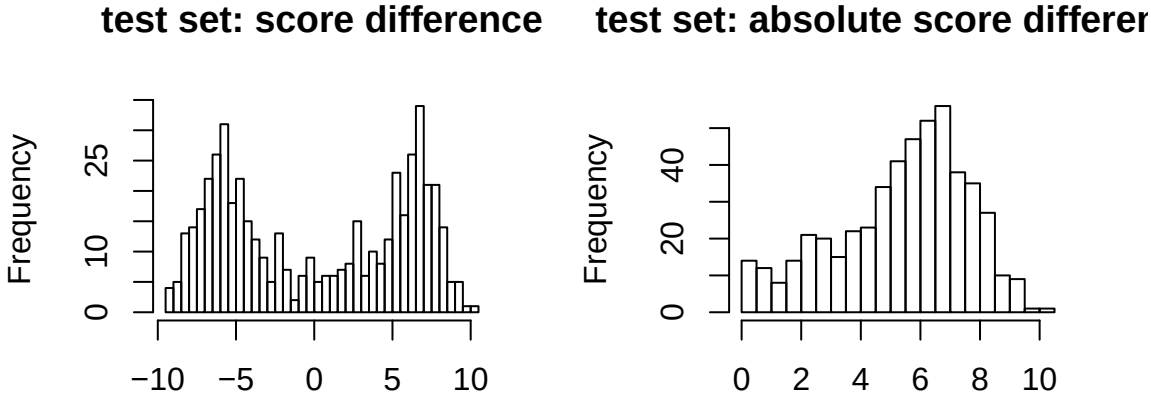
```
round(mean(abs(rules2)),3)
```

```
## [1] 7.647
```

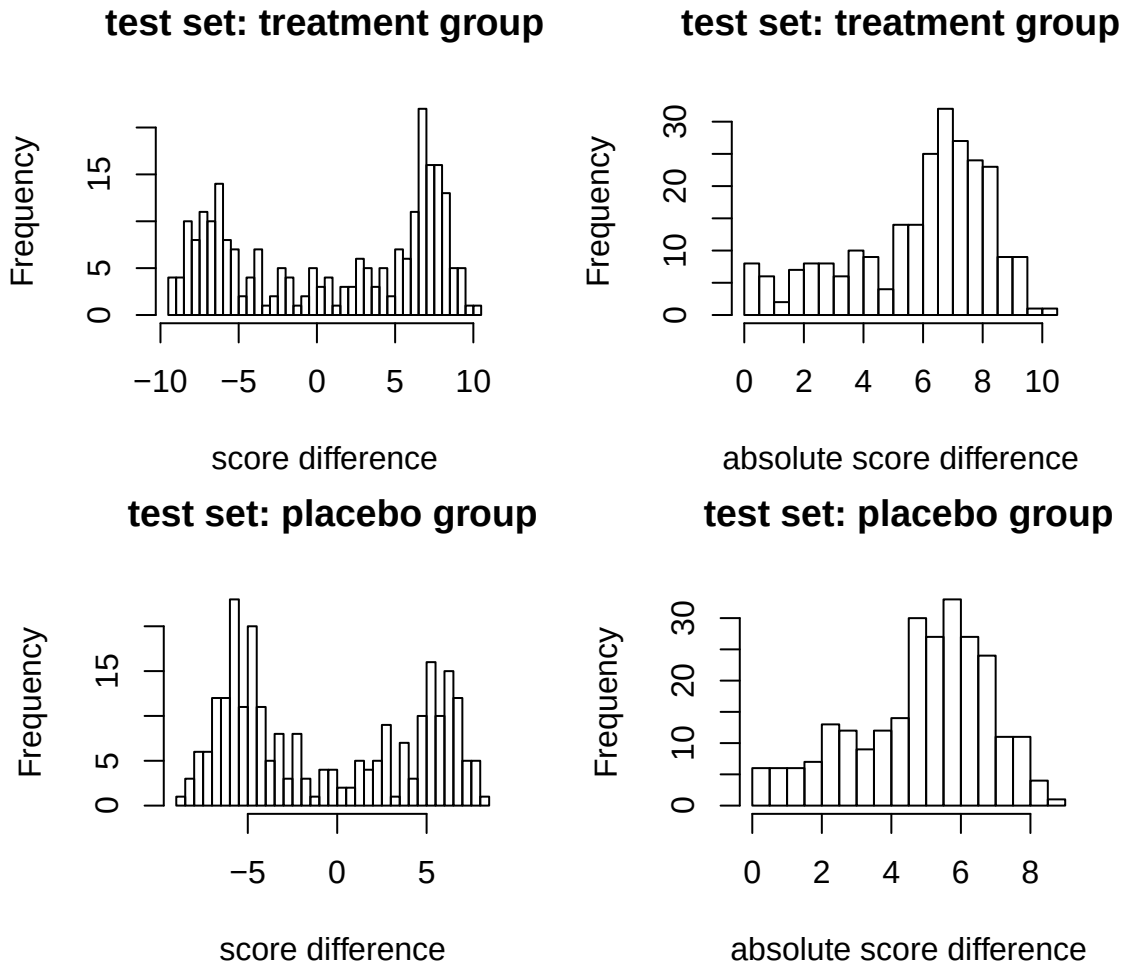
For the test dataset,

- Fit the LME and calculate the score differences by following the same procedures as the training dataset.
- The mean value of the score differences in treatment group is s_1, the mean value of the score difference in placebo group is s_2
- If the score difference is more close to s_1, the subject is assigned as treatment group; if the score difference is more close to s_2, the subject is assigned as placebo group.
- Calculate the AUC between the above assignment and the actual assignment.

The score differences in the testing dataset



The score differences in the testing dataset with the true assignment:



If we use the score difference to decide the assignment, the AUC is

```
##
## Call:
## roc.default(response = trt_est, predictor = trt_test)
```

```
##
## Data: trt_test in 245 controls (trt_est 1) < 255 cases (trt_est 2).
## Area under the curve: 0.5559
```

If we use the absolute score difference to decide the assignment, the AUC is

```
##
## Call:
## roc.default(response = trt_est, predictor = trt_test)
##
## Data: trt_test in 65 controls (trt_est 1) < 435 cases (trt_est 2).
## Area under the curve: 0.7024
```

The score differences in the testing dataset with the estimated assignment: