Bucci Teo
10621873

Exercise 1

I have two paired populations for which I have to compare the mean of two covariates.
I have to add two new covariates defined as the difference of the original ones, done pairwise and work on this latter.

I want to perform the following test
H0: delta == delta.0 vs H1: delta != delta.0 with delta.0=c(0,0)

I set alpha = 0.05.

Then I compute the T^2 test statistics by

$$T^2 = n * (d.mean-delta.0) \%*\% S^{-1} \%*\% (d.mean-delta.0)$$

where d.mean is the sample mean of the difference and S the covariance matrix. It is equal to

$$T^2 = 11.0284$$

Finally I compute the proper quantile of the Fisher distribution multiplied by (n-1)*p/(n-p)

cfr.fisher = (n-1)*p/(n-p)*qf(1-alpha,p,n-p)

where p is the number of covariates (2). It is equal to:

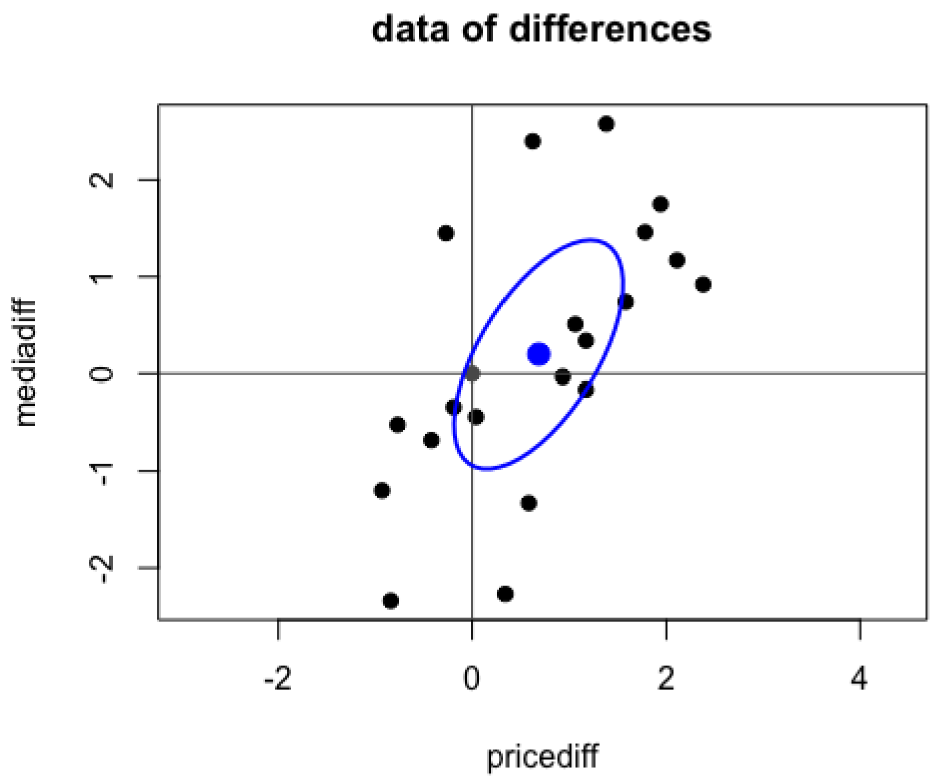
cfr.fisher = 14.28286

By comparing the two, the test statistics is not greater than cfr.fisher, thus the test says that there is evidence to not reject H0, and the means are not different between the two stores.

The p-value of the test is indeed: 0.1011722 which is greater than 0.05.

I check that delta (my new dataframe of differences) belongs to a multivariate Gaussian distribution. To do so I perform a Shapiro test (H0: “data is Gaussian”, H1: $H0^C$), obtaining a p-value of 0.7508, high enough at any reasonable level to not reject H0. The assumptions are fulfilled.

The plot of the confidence region is provided here



I see that (0,0) is inside it, which confirms the test results.

I compute the four simultaneous Bonferroni intervals:

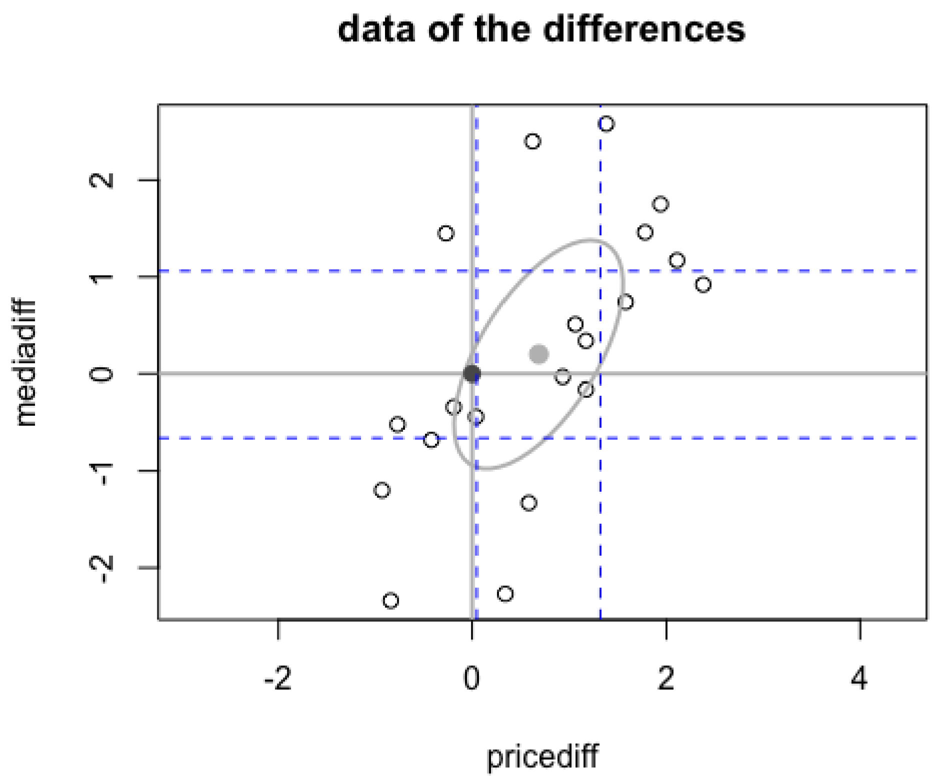
For the mean

```
                 inf center      sup
pricediff  0.0480422 0.6830 1.317958
mediadiff -0.6597456 0.2005 1.060746
```

For the variance

```
                inf   center      sup
pricediff 0.5322847 1.059612 2.841861
mediadiff 0.9770104 1.944921 5.216247
```

The Bonferroni intervals are plotted here



I see that for pricediff they don't contain 0, while for mediadiff they are inside it.
The fact that intervals are simultaneous is so that they are stricter than the rectangle built outside the ellipse, and in this case they are so short that for one covariate it falls inside the ellipse but outside the simultaneous one, even though in the initial test we didn't reject.