TFofLDS

Szymon Czop

10 12 2019

Assignment 1

In this exercise will have a closer look at the Holm's procedure. We will check the Family Wise Error Rate and the power of the procedure for a different number of the non-zero entry in vector, a different covariance between each random variable in vector and different value in the non-zero entry.In all plots, we will compare it to the behavior of the very classic Bonferroni's procedure.

In the whole report by null hypothesis, we understand $H_0 : \theta_i = 0 \; for \; i \in \{1....n\}$

The Family wise error rate (FWER): $FWER = P(V \geqslant 1)$ where V is number of false positives (Type I error)

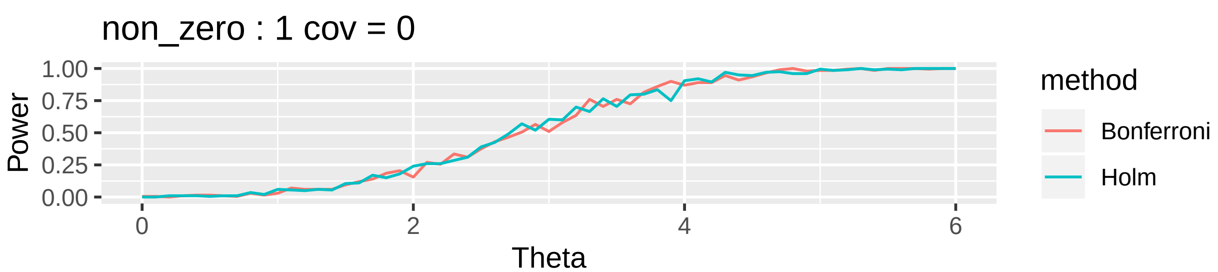
Explanation about the legend of the plot:

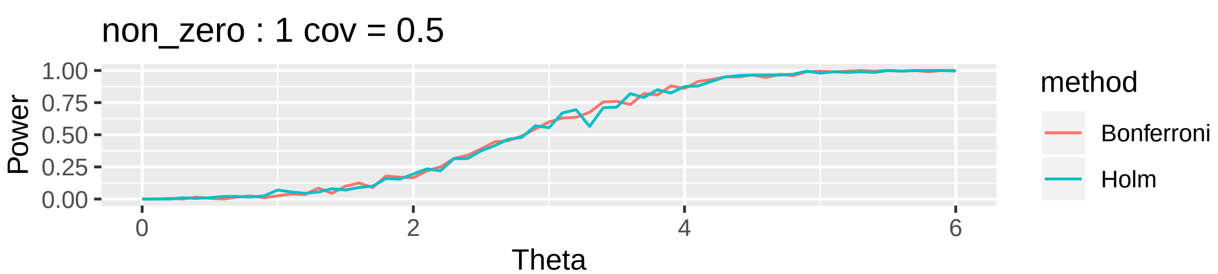
non_zero: number of non-zero entries in the mean vector

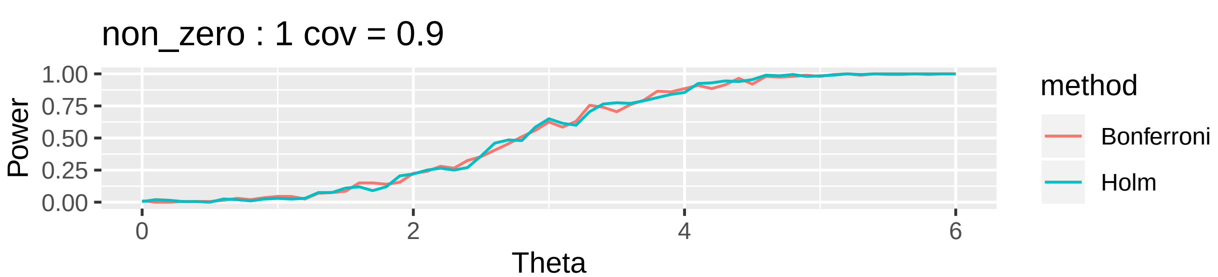
cov: covariance between each variable (variance for all of them is equal to 1)

Power

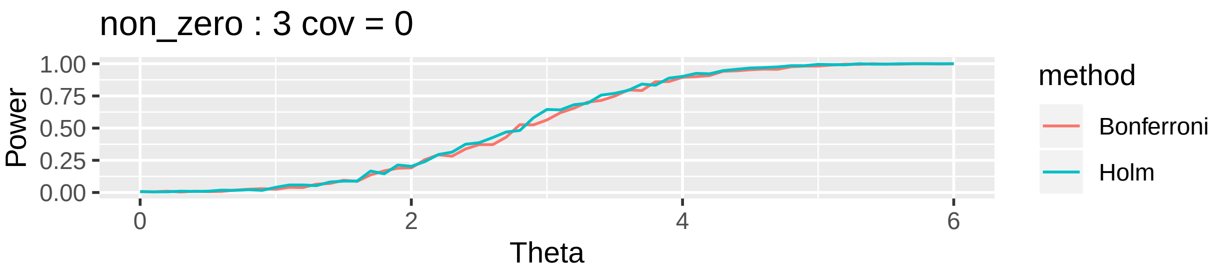
```
## Loading required package: gridExtra
```

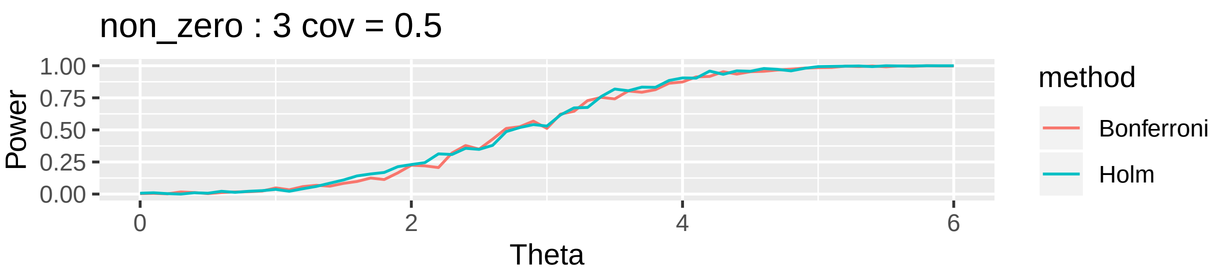


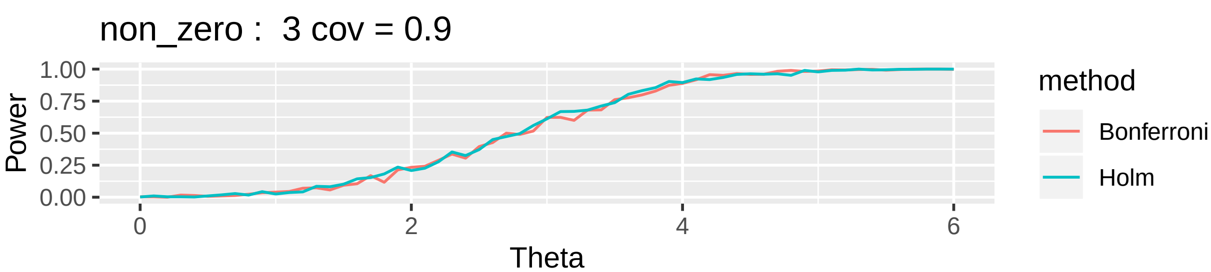




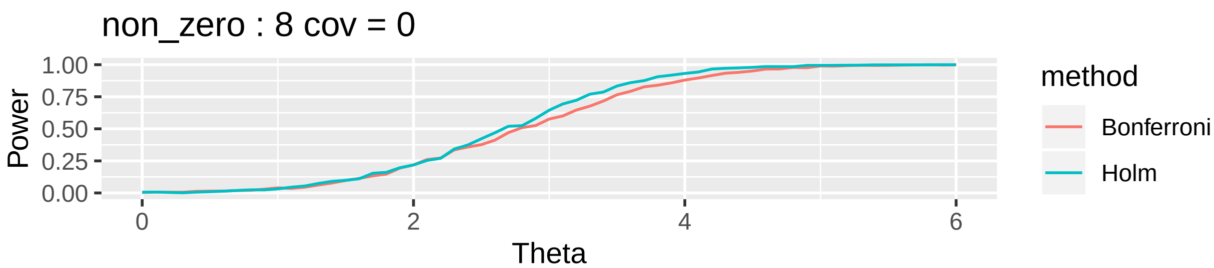
It is really hard to find out any outstanding difference in those two methods when the number of non-null components is equal to 1. Both procedures seem to work very similarly and have power at the level of 75%, when $\theta = 4$

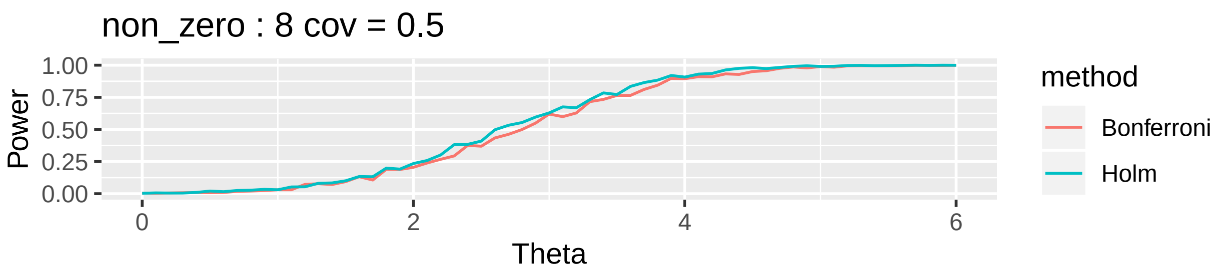


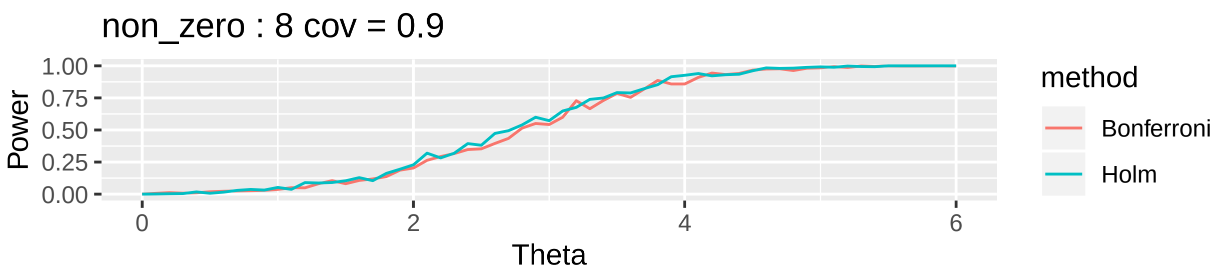




When we increase the number of non-null entries to 3, Holm's procedure seems to give us a slightly bigger power, but this plot is still not showing which one of them is less or more strict when it comes to rejecting the null hypothesis.Power is increasing faster and for $\theta = 4$ we have power on the level of 90%





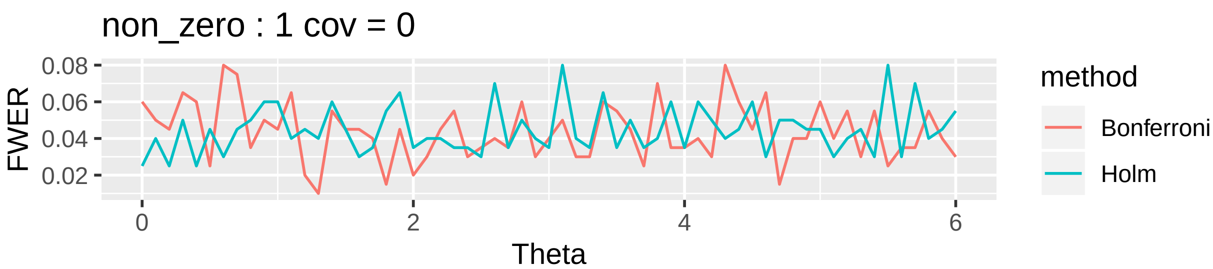


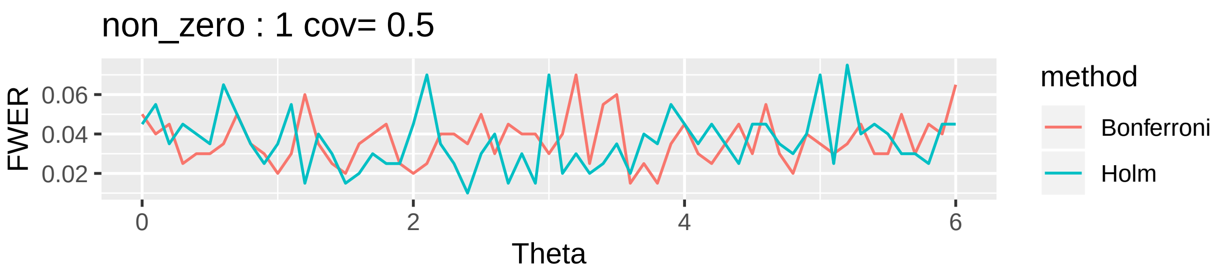
For 8 (of 10) non-zero entries we see that Holms procedure is increasing faster in the case when a covariance between each variable is eqaul to zero. When we change the covariance this advantage is blurred and again those two procedures are almost at the same level. In this case (8 non-zero entries) red line rarely is above the blue which indicates Holme's procedure. For $\theta = 4$, Holmes procedure gives us more than 90% power.

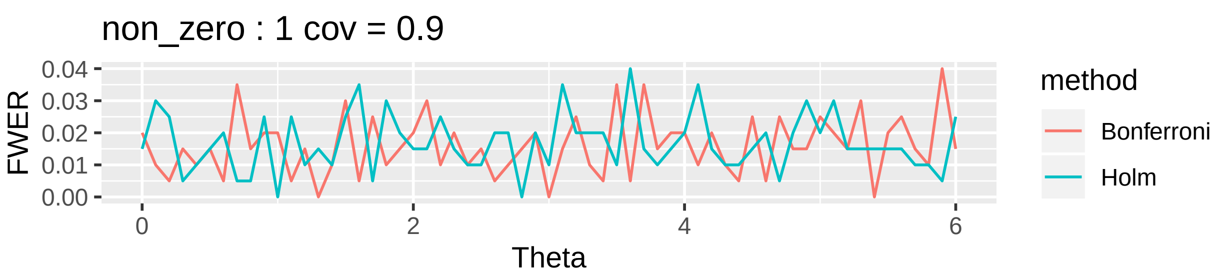
Summary

From the lecture, we know that Bonferroni's procedure is more strict than the one made on Holm's assumptions. The real difference that is noticeable at the first look of the aye occurs in the case when we have much more non-zero entries than zero ones. The power of Holme's procedure increases much faster and gives us a bigger probability to reject the null hypothesis.

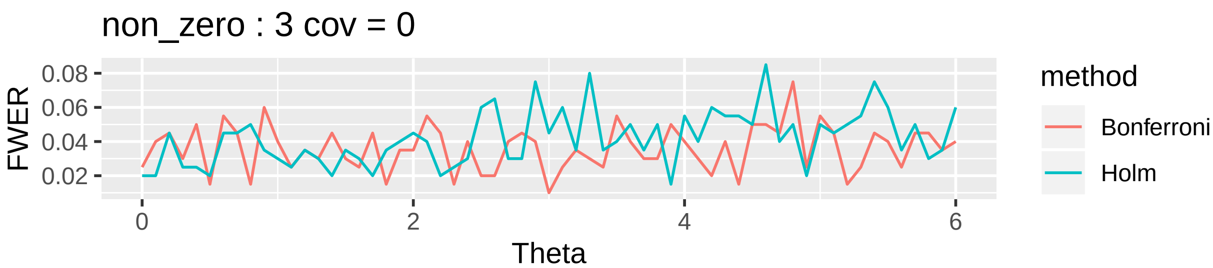
FWER

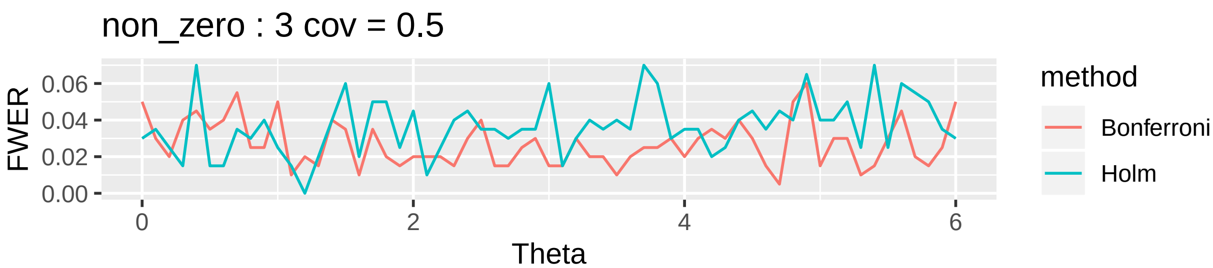


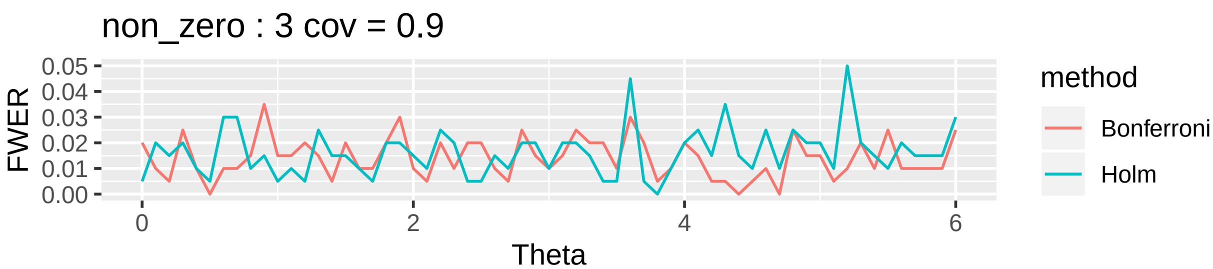




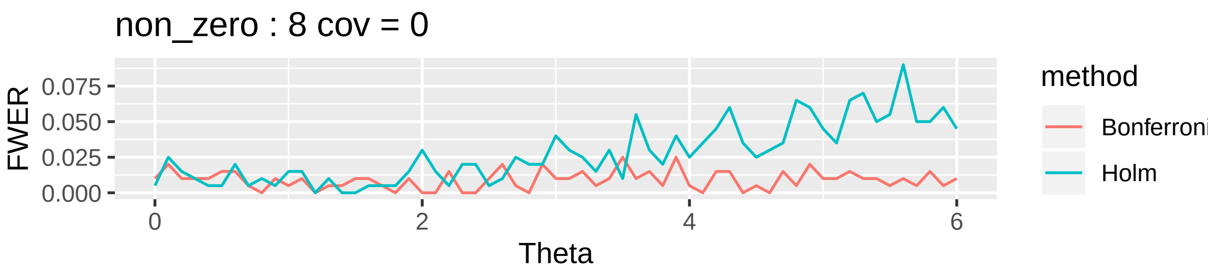
When covariance is equal to zero both procedures work similarly. The biggest error on the level 8% is reached two times in Holme's procedure and one time for Bonferroni's. Bonferroni's procedure has had a much lower rate (sometimes close to zero) in many cases. When covariance increases error rate is going down

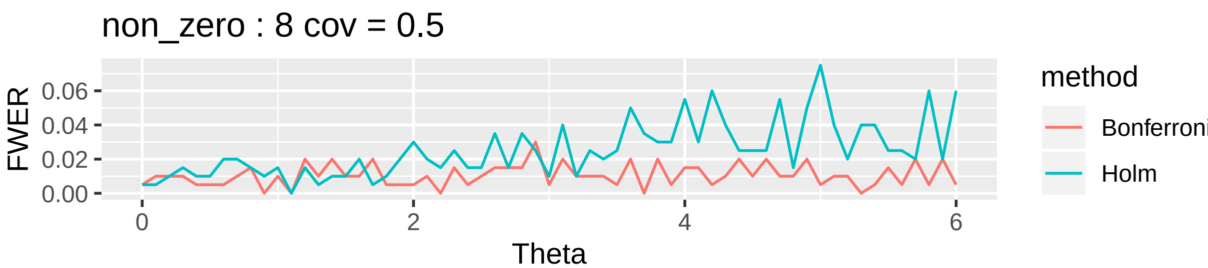


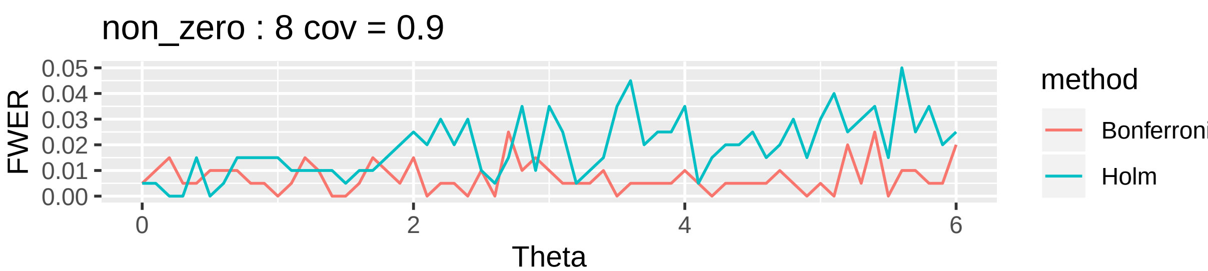




For the smaller value of non-zero theta, Holme's is behaving similar to Bonferroni (maybe even a bit better) but for bigger values like 3-6, it's easy to see that blue line is almost always above the red one which indicated bigger error rate.







Again the biggest difference in the performance of both procedures is again clearly visible for the case when non-null entries are 8. Holm's procedure has much bigger error rate that is increasing for bigger entries. Because the power of Holm's procedure, in this case, is bigger it may be seen as somehow natural to have greater power for this case but at the same time higher error rate.

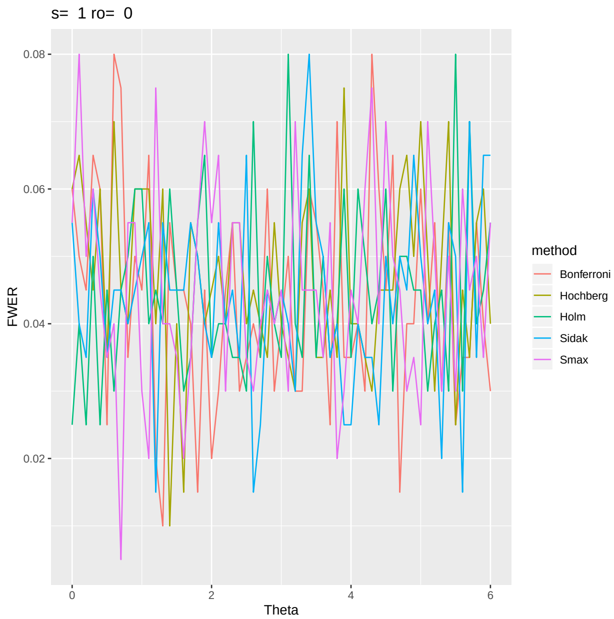
summary

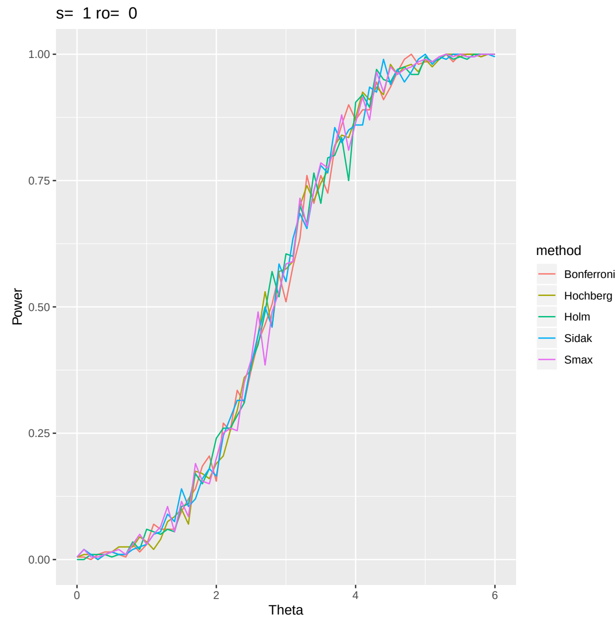
It's visible that Bonferroni's procedure is more strict and in overall look controls FWER in a bit bigger extend

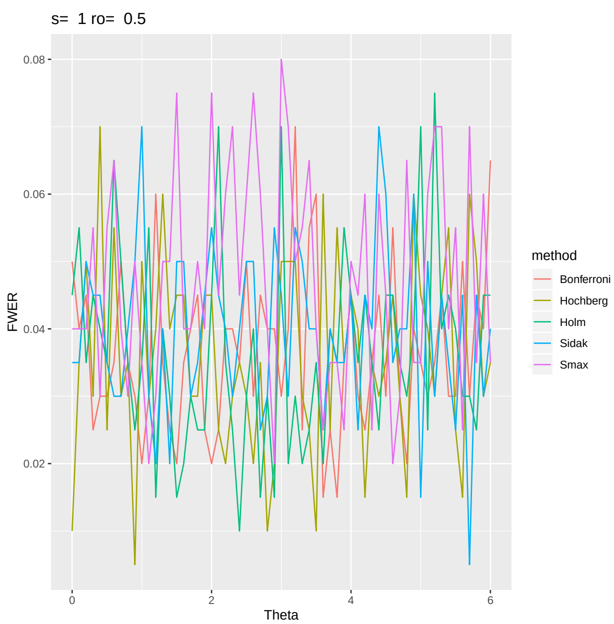
Assignment 2

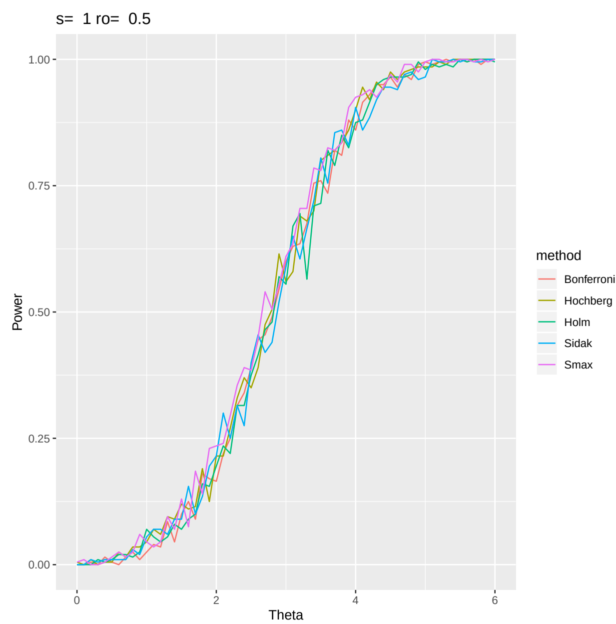
For this assignment, for the sake of simplicity, I will sum up my insights after every three pairs of plots. Each pair is representing FWER and Power for a fixed value of non-null entries and different covariances predefined on the list.

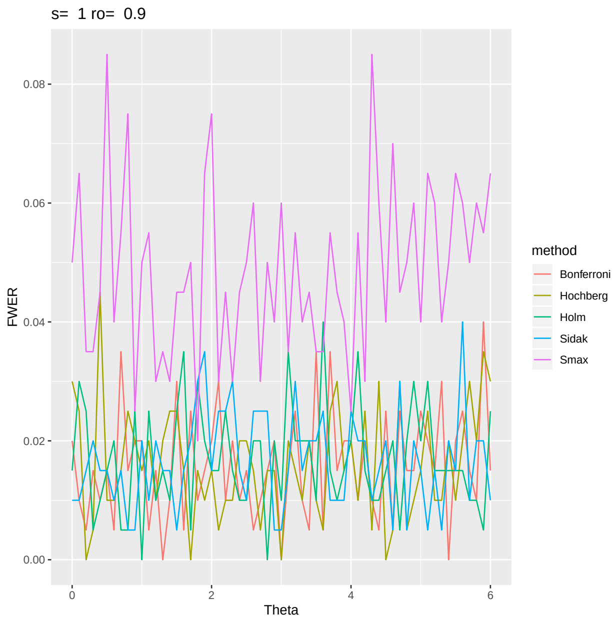
Non-zero entries: 1

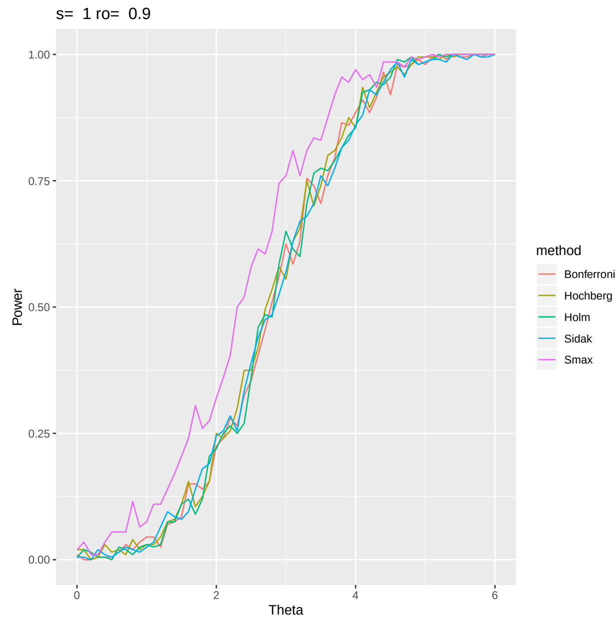






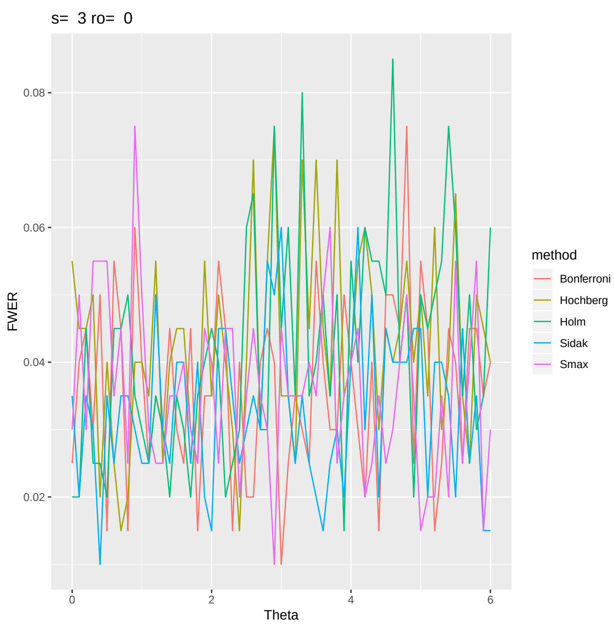


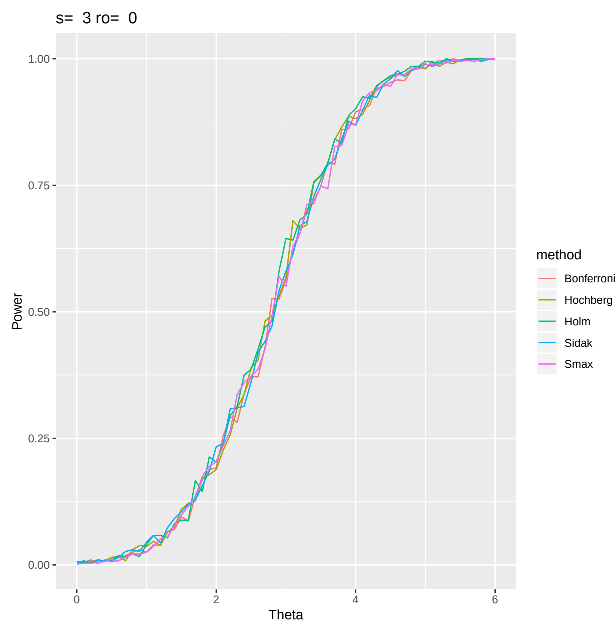


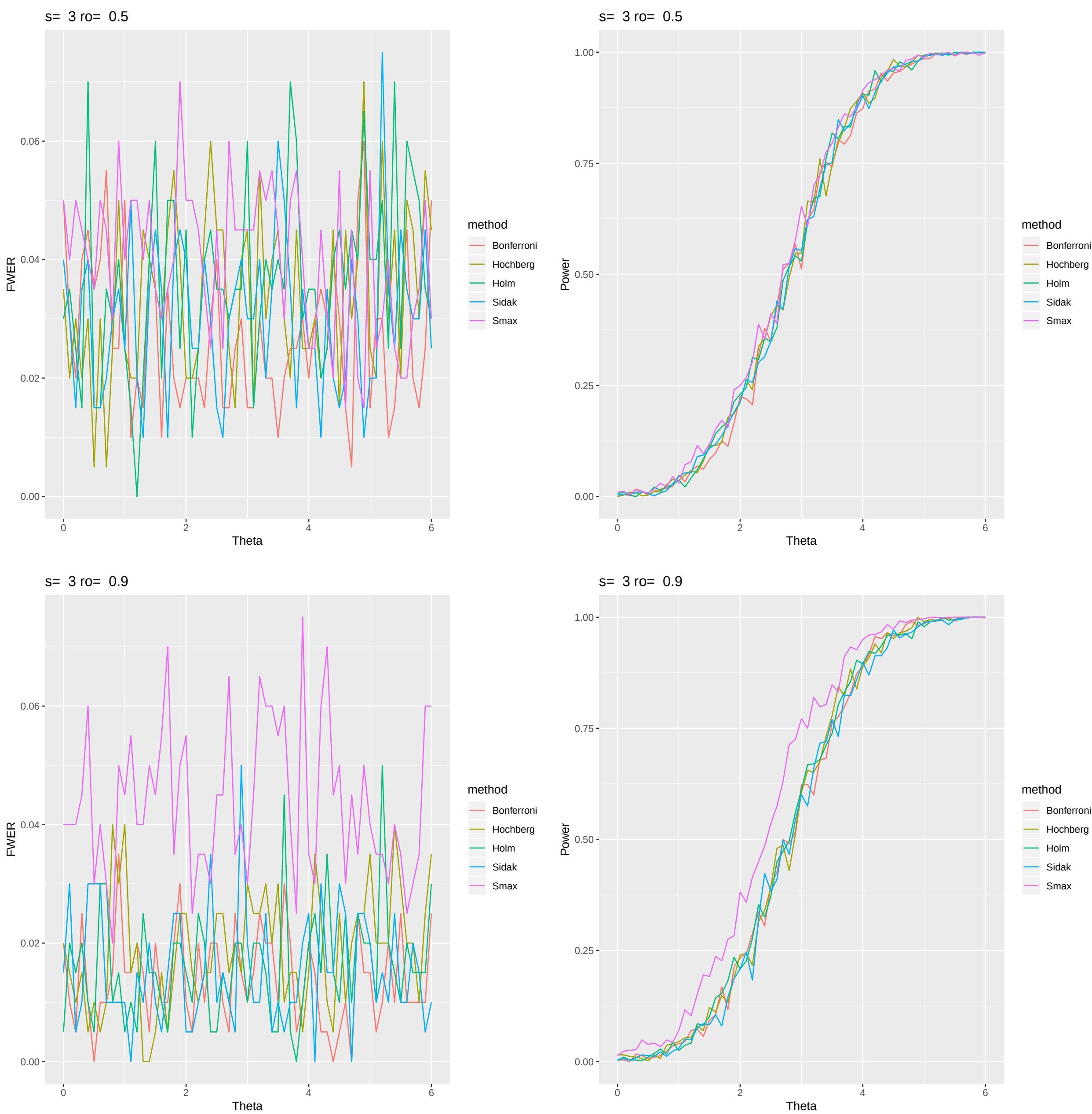


The procedure that stands out for one non-null entry and covariance on the level of 0.9 is Smax. It has the biggest power, clearly bigger than any other procedure taken into account, but at the same time has the biggest FWER. Now it's up to us if we want to pick (under condition made in simulations about the value and covariance) procedure that has low FWER and lower power or have the biggest power possible with a bigger probability.

Non-zero entries: 2

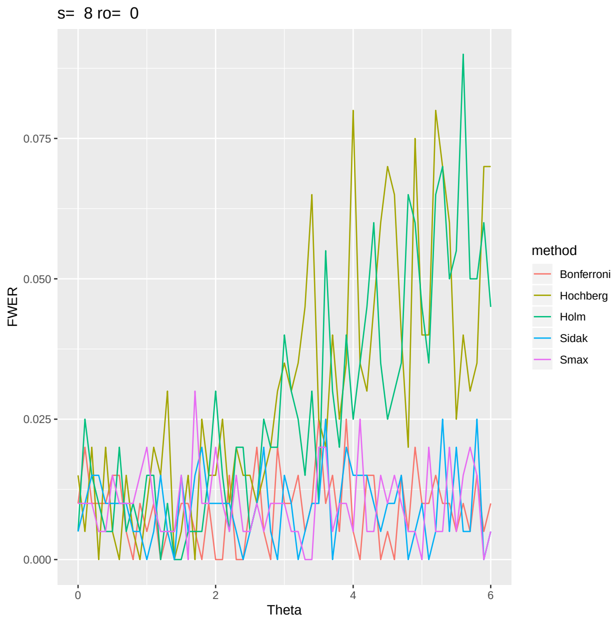


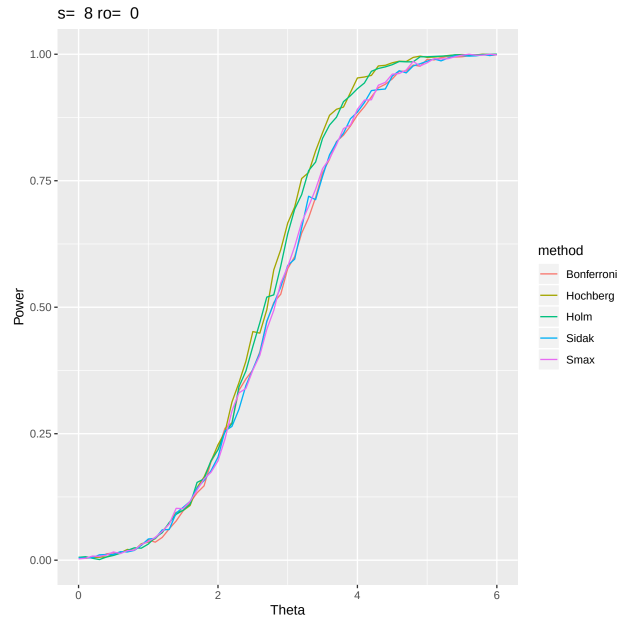


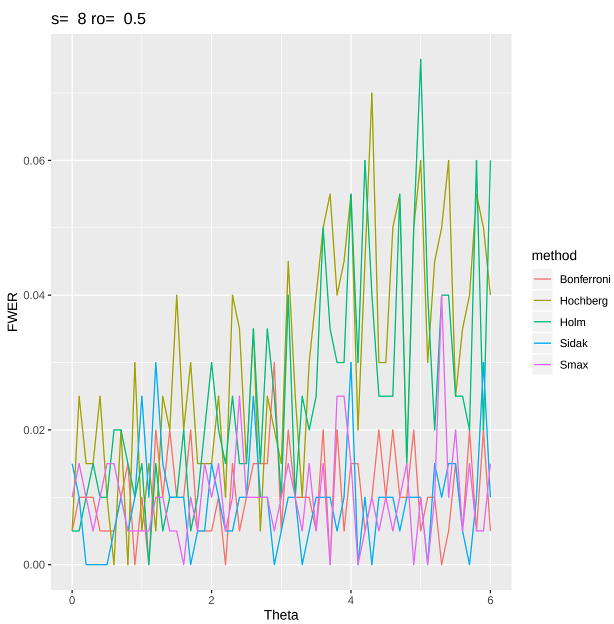


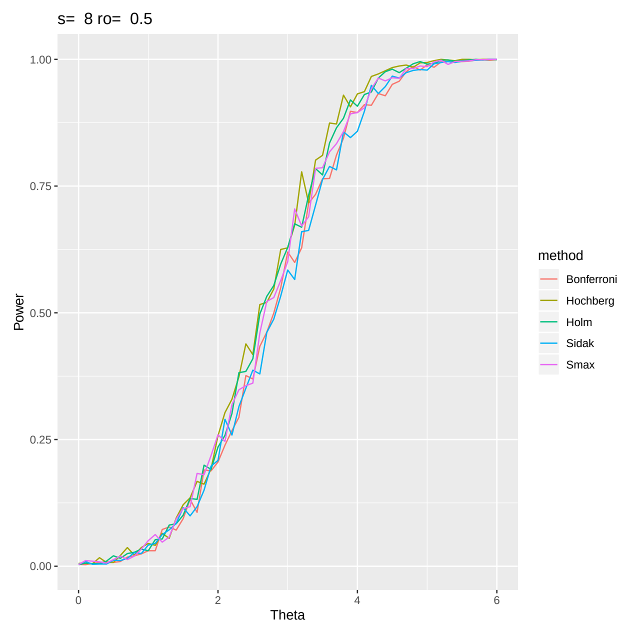
The situation showed above is very similar to the previous one, again almost all procedures give us the same power an FWER. Only at the last pair of plots, we see that Smax procedure breaks off from the rest and gives bigger power but at the same time higher FWER.

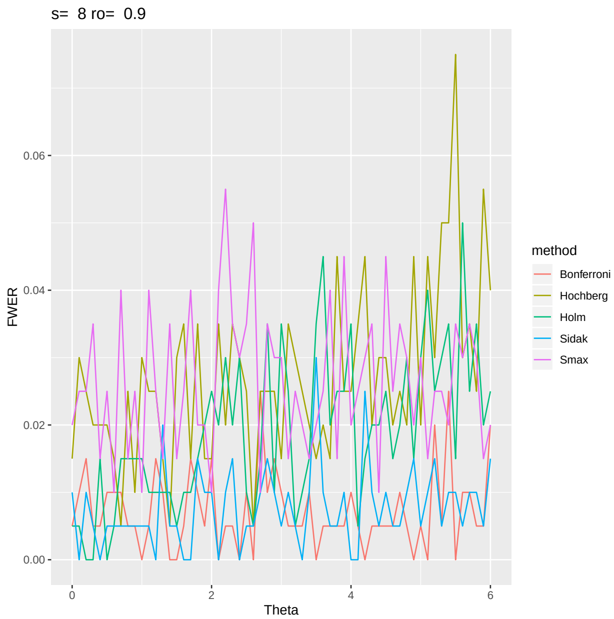
Non-zero entries: 3

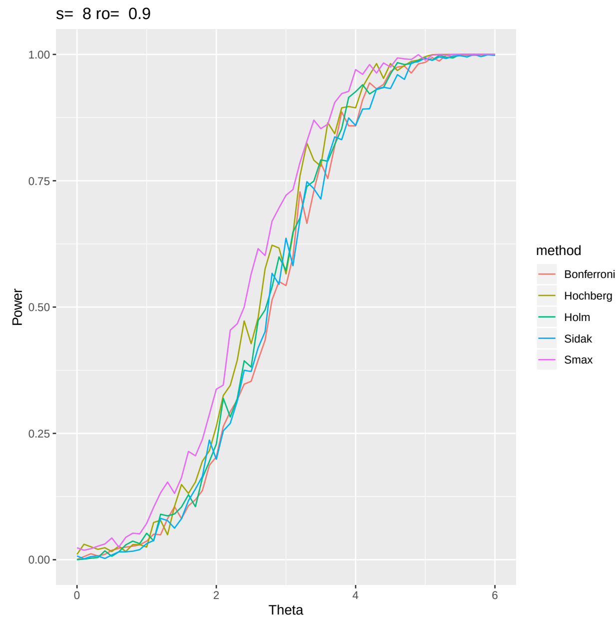












This is the most interesting case!

For 8 non-null entries and zero covariance Hochberg and Holm procedure give us the biggest power. If our non-null entries are bigger then 2 power of both above-mentioned procedures is starting to increase much faster than in other procedures. Again exactly in the same point, FWER for those two procedures that start on letter H is increasing much more clearly. While other procedures for some points have FWER equal to zero for Hochber and Holm we are close to 0.08.

When covariance is at level 0.5 power of each procedure is almost on the same level. Again Hochberg and Holm are the most powerful ones but also with the biggest error rate.

For the biggest covariance again Smax has the biggest power but Hochberg is just behind it. What is interesting FWER for Smax is the biggest when non-zero values are between 0,1 - 3. After this it's FWER is similiar to the other procedures. In the same time Hochberg is incresing from 3 to 6 and it can be seen on the last plot of FWER.

This example clearly show us that in some special cases, type of testing procedure may have significant inpact on power and FWER of our testing.

Made by Szymon Czop