Workshop 3: Computing the clustering of a galaxy catalog (part 2)

Today we will continue working in the computation of the clustering of our catalog.

1. If you finished the computation of DD, then you can now easily compute the term RR:
    A. You need to create an array with the same length as the array "DD". Call this array "RR"
    B. For the computation of RR you have to repeat exactly what you did for DD but this time instead of using the position of your galaxies, you use the positions of the random catalog that you created. If you were not successfully creating the random catalog, you can download it from https://home.strw.leidenuniv.nl/~garcia/astrotwincolo/Workshop2/Results/

2. Make a log-log plot of theta versus RR.

3. Now we have to normalize the arrays DD and RR by the number of pairs. The number of (non repeated) pairs is N*(N-1)/2, with N the number of sources in the catalog, then you only need to divide DD by Nd*(Nd-1)/2 (with Nd the number of galaxies in the catalog) and RR by Nr*(Nr-1)/2 (with Nr the number of random sources). Note that if you counted the pairs twice (i.e. you compute the distance between the galaxy A and B and then between the galaxy B and A), then the number of pairs in that case is N^2, and you have to use that factor of normalization.

4. Use the more basic estimator to compute the correlation function: w(theta) = DDn/RRn -1, where DDn and RRn are the normalized DD and RR arrays.

5. Make a log-log plot of theta versus w(theta) and check that this has a power law shape. What can you see in the plot?

6. Now, we will add error bars to the measurement. For now, just assume Poisson error bars. This means that the error in w is [sqrt(DD)/ (Nd*(Nd-1)/2)]/RRn (Note that DD here is the not-normalized array).

7. How big are the error bars? why?

8. Finally, run your code but now using the second galaxy catalog. Note that the RR term will be the same, then you don't have to compute it again.

9. Overplot the correlation function for this sample together with the other, and give answer to the following questions:
    A. Are the correlation functions different?

B. Which is higher?
C. Which population of galaxies is more clustered?
D. What does it implies?
E. Were you expecting this result?
F. How are the error bars compared between the samples.