Write a code to compute the correlation function $\xi(r)$ for a set of points. Use the following estimator: $\xi(r) = \left(\frac{N_R}{N_D}\right)^2 \frac{DD(r)}{RR(r)} - 1$, where N_D and N_R are the number of data and random points, respectively, and $DD(r)$ and $RR(r)$ are the pair counts as a function of separation r for data-data and random-random pairs, respectively. Compute $\xi(r)$ in 15 logarithmic bins of separation in the range $0.1 - 20h^{-1}\mathrm{Mpc}$ (i.e., the first bin starts at $\log r = -1$ and the last bin ends at $\log r = 1.301$).

Part 1

Compute $\xi(r)$ for three SDSS mock galaxy samples:

- Galaxies with *r*-band absolute magnitudes brighter than -21 (in real space)
- Galaxies with *r*-band absolute magnitudes brighter than -20 (in real space)
- Galaxies with *r*-band absolute magnitudes brighter than -20 (in redshift space)

These galaxy samples are in data files that you can download from the course website:
`SDSS_Mr21_rspace.dat`, `SDSS_Mr20_rspace.dat`, `SDSS_Mr20_zspace.dat`
The file containing random points is:
`SDSS_random.dat`
These files contain galaxy positions in spherical coordinates: (RA, DEC, z), where RA and DEC are the galaxies' coordinates on the sky in degrees, and z is the redshift. All the samples are limited to the redshift range $0.02 \le z \le 0.06$ and to the sky coverage of the SDSS.

Make the following plots:

a. $\log \xi(r)$ vs. $\log r$ for the two real-space galaxy samples of different luminosity: $M_r < -20$ and -21. *(Show both curves on the same plot)*

b. $\log \xi(r)$ vs. $\log r$ for the real-space compared to the redshift-space galaxy samples with $M_r < -20$. *(Show both curves on the same plot)*

Discuss the difference between the correlation functions in plot **b.** in light of what you have learned about redshift-space distortions.

Part 2

Now use your code to compute $\xi(r)$ for dark matter (DM) in a N-body simulation. The DM data file is `DM.dat` and the DM random file is `DM_random.dat`
These files contain DM particle positions in Cartesian coordinates in $h^{-1}\mathrm{Mpc}$: (x, y, z). The size of the simulation cube is $141.3h^{-1}\mathrm{Mpc}$.

Compute the bias function for galaxies with $M_r < -20$ and -21 (in real space): $b(r) = \sqrt{\dfrac{\xi_{gal}(r)}{\xi_{DM}(r)}}$

Make the following plot:
c. $b(r)$ vs. $\log r$ for the two luminosities. *(Show both curves on the same plot)*

Discuss the shape of the bias function for the two galaxy samples.

At what scales does the bias become approximately scale-independent?

What is the approximate large-scale bias factor for galaxies with $M_r < -20$ and $M_r < -21$?

Part 3 – extra credit

Estimate Poisson errors for the correlation function. Re-make plot **a.** from part 1, including these errorbars.

What limits the range of scales that $\xi(r)$ can be measured on with these data-sets? Can it be measured on arbitrarily small or large scales? Why or why not?