The Press-Schechter mass function

To state the obvious: It is important to relate our theories to what we can observe. We have looked at linear perturbation theory, and we have considered a simple model for nonlinear structure formation, the spherical collapse model. But we cannot observe directly the process of structure formation. What we can see are the results, for example clusters of galaxies. We can count these objects and find their density, and also further details about their statistical distribution. One example we have encountered already is the galaxy luminosity function, which gives the number density of galaxies of a given luminosity.

The quantity we will consider in the following is the mass function \( n(M) \) of cosmic structures, defined by

\[
dN = n(M) dM, \tag{1}
\]

where \( dN \) is the number of structures per unit volume with mass between \( M \) and \( M + dM \). It should be fairly obvious how the mass function is measured in principle: Just select a volume in space and count the number of structures of a given mass contained within it. In practice there will, of course, be complications and I will mention a couple of them when we look at an application of the theory that now follows.

Having an analytical expression for the mass function would be good. In a classic paper published in 1974 Press and Schechter set out to provide us with that. We will now look at what they did.

Consider the density fluctuation field \( \delta(x) \) (I have omitted to make the time dependence explicit, but it is of course there.) We can filter the field on a length scale \( R \), which means throwing away information about the detailed behaviour of \( \delta(x) \) on scales smaller than \( R \). The filtering can be accomplished by convolving \( \delta \) with a window function \( W_R(x - x') \):

\[
\delta(x; R) = \int d^3x' \delta(x') W_R(x - x'). \tag{2}
\]

A popular choice of \( W_R \), and the one we will use, is the top hat function

\[
W(x - x') = \begin{cases} 
1, & \text{if } |x - x'| < R, \\
0, & \text{otherwise.}
\end{cases} \tag{3}
\]
The region of radius $R$ contains a mass

$$M = \frac{4\pi}{3} \rho_0 R^3,$$

where $\rho_0$ is the uniform background density. We introduce the notation $\delta(x; R) \equiv \delta_M$.

Press and Schechter assumed that the density field has a Gaussian probability distribution:

$$P(\delta_M) d\delta_M = \frac{1}{\sqrt{2\pi\sigma_M}} \exp \left( -\frac{\delta_M^2}{2\sigma_M^2} \right) d\delta_M,$$

where $\sigma_M$ is the variance of the density field filtered on a scale $R$ enclosing a mass $M$. The probability that at some point $\delta_M$ exceeds some critical value $\delta_c$ is now given by

$$P_{>\delta_c}(M) = \int_{\delta_c}^{\infty} P(\delta_M) d\delta_M.$$

This probability depends on the filter mass $M$, and also on the redshift $z$ through the variance

$$\sigma_M^2 = \sigma_R^2 = \frac{1}{2\pi^2} \int_0^{\infty} dk k^2 P_m(k, z) \tilde{W}_R^2(k),$$

where $P_m(k, z)$ is the matter power spectrum and $\tilde{W}_R(k)$ is the Fourier transform of the top hat filter function.

It is also the case that $P_{>\delta_c}$ is proportional to the number of cosmic structures characterized by a density perturbation $> \delta_c$, regardless of whether they are isolated or contained within denser structures which collapse with them. A value of $\delta_c$ of great interest to us is $\delta_c \approx 1.686$, since we have seen that this is the linear-theory density contrast which corresponds to virialized structures. To find the number of regions with mass $M$ which are isolated, in other words surrounded by underdense regions, we must subtract $P_{>\delta_c}(M + dM)$. This ignores to so-called cloud-in-cloud problem: At a given instant some object, which is nonlinear on a scale $M$, can later be contained within another object on a larger mass scale. In essence we assume that the only objects which exist on a given mass scale are those which have just collapsed.

Another problem is that we cannot treat underdense regions properly (those with $\delta < 0$), which means that half the mass is unaccounted for.
Presumably these underdensities will accrete onto overdensities. Press and Schechter fixed this problem by brute force: They multiplied their mass function by a factor of 2.

We can now set up an expression for the mass function \( n(M) \). It is proportional to the difference of the probabilities already mentioned, but to convert probabilities to a quantity with units of per volume, we need to multiply by the background density divided by the mass scale \( M \). In addition we have the artificial factor of 2 to account for underdense regions:

\[
\frac{d}{dM} \ln \left( \frac{M}{M_*} \right) = -2 \rho_0 \int_0^{\infty} \frac{dP_{> \delta_c}(M + dM)}{dM} \rho_0(M + dM) dM.
\]

We can use the fundamental theorem of calculus

\[
\frac{d}{dx} \int_a^x f(t) dt = f(x) = -\frac{d}{dx} \int_x^a f(t) dt,
\]

to evaluate the derivative, along with the substitution \( x = \frac{\delta_c}{\sqrt{2} \sigma_M} \):

\[
\frac{d}{d\sigma_M} P_{> \delta_c} = \frac{d}{d\sigma_M} \int_{\delta_c}^\infty \frac{1}{\sqrt{2\pi \sigma_M}} \exp \left( -\frac{\delta_c^2}{2\sigma_M^2} \right) d\delta_M = \frac{1}{\sqrt{\pi}} \int_{\delta_c/\sqrt{2\sigma_M}}^\infty e^{-x^2} dx = -\frac{1}{\sqrt{\pi}} e^{-\delta_c^2/2\sigma_M^2} \frac{d}{d\sigma_M} \delta_c = \frac{1}{\sqrt{2\pi} \sigma_M^2} e^{-\delta_c^2/2\sigma_M^2}.
\]

Finally we have

\[
n(M) dM = -\frac{\sqrt{2}}{\pi} \frac{d\sigma_M}{dM} \frac{\rho_0 \delta_c}{M^2 \sigma_M^2} \exp \left( -\frac{\delta_c^2}{2\sigma_M^2} \right) dM.
\]

This mass function, with \( \delta_c = 1.686 \), gives us the number density of collapsed objects per unit mass. The normal procedure is to evaluate \( n(M) \) by calculating \( \sigma_M \) and its derivative from the linear theory matter power spectrum.
An application of Press-Schechter theory to galaxy clusters

Determining $n(M)$ from galaxy clusters has been a popular job for cosmologists. One thing that allows one to do is to find the normalization of the matter power spectrum through the parameter $\sigma$, that is, $\sigma_R$ evaluated for $R = 8 \ h^{-1} \text{Mpc}$ at $z = 0$. But in order to do this, one must be able to determine the mass of galaxy clusters. Since the dominant contribution to the mass of a cluster is in the form of dark matter, this is a non-trivial task. At least three different approaches may be taken:

1. Determine the line-of-sight velocity dispersion, assume isotropy, and use the virial theorem.

2. Determine the temperature of the X-ray emitting hot gas in the cluster. The temperature is related to the depth of the gravitational potential well in the cluster, so there should be a correspondence between temperature and mass. The problem here is to model the gas accurately.

3. Gravitational lensing. This is in principle the least ambiguous and most assumption free method, but it is time-consuming. Increasingly, though, this is the method of choice.

We will consider a simplified version of how the cluster abundance can be used to find $\sigma_8$, assuming for simplicity that the background universe is described by the Einstein-de Sitter model. The number density of clusters at temperature corresponding to $k_B T = 7 \text{ keV}$ at the present epoch has been measured to be

$$n(7 \text{ keV}, z = 0) = 2.0_{-1.0}^{+2.0} \times 10^7 h^3 \text{Mpc}^{-3} \text{keV}^{-1}. \quad (12)$$

In simulations of structure formation in an EdS background one finds that a cluster with X-ray temperature $T = 7.5 \text{ keV}$ has a mass within a radius $1.5 h^{-1} \text{Mpc}$ (this is often called the Abell radius) of

$$M_{\text{Abell}} = (1.1 \pm 0.2) \times 10^{15} h^{-1} M_\odot. \quad (13)$$

So the simulation gives us useful, but not immediately applicable information. To compare with the theoretical mass function, we need to know what virial
mass a temperature of 7 keV corresponds to. Fortunately the simulation can also give us the density profile of the cluster, allowing us to convert the Abell mass to the mass within the virial radius. And a scaling relation like

$$\frac{k_B T}{0.07 \text{ keV}} = \left( \frac{M}{10^{12} h^{-1} M_\odot} \right)^{2/3} (1 + z_{\text{vir}})$$  \hspace{1cm} (14)$$
can be used to scale the mass from 7.5 to 7 keV. The end result turns out to be

$$M_{\text{vir}} = (1.2 \pm 0.3) \times 10^{15} h^{-1} M_\odot.$$  \hspace{1cm} (15)$$

With the top-hat filter this corresponds to a sphere of radius

$$R = \left( \frac{3M_{\text{vir}}}{4\pi \rho_{c0}} \right)^{1/3} \approx 10 h^{-1} \text{ Mpc},$$  \hspace{1cm} (16)$$
where the critical density at the present epoch is $\rho_{c0} = 2.78 \times 10^{11} h^2 M_\odot \text{ Mpc}^{-3}$.

We now need to relate $dM$ to $d(k_B T)$ using equation (14) with $z_{\text{vir}} = 0$:

$$\frac{d(k_B T)}{0.07 \text{ keV}} = \frac{2}{3} \frac{1}{M} \left( \frac{M}{10^{12} h^{-1} M_\odot} \right)^{2/3} dM$$  \hspace{1cm} (17)$$
which can be rewritten as

$$d(k_B T) = \frac{2}{3} \frac{k_B T}{M} dM.$$  \hspace{1cm} (18)$$

The next thing we need is the derivative of $\sigma_M$. In accurate work we would have to evaluate the derivative numerically from the definition of $\sigma_M$ in terms of the matter power spectrum. Here we will settle for something simpler. A decent fit to $\sigma_M = \sigma_R$ for the EdS model with the scalar spectral index $n_s = 1$ is

$$\sigma_R = \sigma_8 \left( \frac{R}{8 h^{-1} \text{ Mpc}} \right)^{-0.8}.$$  \hspace{1cm} (19)$$
This and the relation $M = 4\pi \rho_{c0} R^3 / 3$ between mass and radius allows us to calculate the derivative we need:

$$\frac{d\sigma_M}{dM} = \frac{d\sigma_R}{dR} \frac{dR}{dM}$$
\[
\begin{align*}
&= -0.8 \frac{\sigma_R}{R} \left( \frac{dM}{dR} \right)^{-1} \\
&= -0.8 \frac{\sigma_R}{R} \left( \frac{3M}{R} \right)^{-1} \\
&= -0.8 \frac{\sigma_R}{3M}.
\end{align*}
\]

(20)

Inserting this in expression (11) for the mass function gives

\[
n(M) dM = \sqrt{\frac{2}{\pi}} \rho_0 \frac{0.8}{3} \frac{\delta_c}{\sigma_R} \exp \left( -\frac{\delta_c^2}{2\sigma_R^2} \right) dM.
\]

(21)

To get something we can compare with the observed cluster abundance we use equation (18) and

\[
n(M) = \frac{dN}{dM}
\]

(22)

to go from \(n(M)\) to \(n(T)\):

\[
\frac{dN}{dM} = \frac{dN}{d(k_BT)} \frac{d(k_BT)}{dM} = n(T) \frac{d(k_BT)}{dM},
\]

(23)

so that

\[
n(T) = \frac{3M}{2k_BT} n(M).
\]

(24)

Inserting equation (20) and substituting \(M = M_{\text{vir}} = 1.2 \times 10^{15} h^{-1} M_\odot\), \(k_BT = 7\) keV finally leaves us with

\[
n(k_BT = 7\text{ keV}) = 1.06 \times 10^{-5} h^3 \text{ Mpc}^{-3} \text{ keV}^{-1} \frac{\delta_c}{\sigma_R} \exp \left( -\frac{\delta_c^2}{2\sigma_R^2} \right).
\]

(25)

Equating this to the observed value \(2.0 \times 10^{-7} h^3 \text{ Mpc}^{-3} \text{ keV}^{-1}\) leaves us with the equation

\[
x \exp \left( -\frac{x^2}{2} \right) = 0.019,
\]

(26)

with \(x = \delta_c/\sigma_R\). This equation must be solved numerically. A simple way of doing this is to rewrite it as

\[
x = \sqrt{\frac{2 \ln \left( \frac{x}{0.019} \right)}{2}}
\]

(27)
and iterate on your calculator. Make a first guess for $x$ and plug it in on the right hand side. Use the output as the input for the next iteration. When the result changes little from one iteration to the next, a solution has been found. Starting with the guess $x = 2$, I found the solution $x = 3.20$ after five iterations. So

$$x = \frac{\delta_c}{\sigma_{R=10h^{-1}Mpc}} = 3.2.$$  \hspace{1cm} (28)

Since $ \delta_c = 1.686$ we get

$$\sigma_{R=10h^{-1}Mpc} = \frac{1.686}{3.2} \approx 0.53.$$  \hspace{1cm} (29)

Finally, we can now use equation (19) to find $\sigma_8$:

$$\sigma_{R=10h^{-1}Mpc} = 0.53 = \sigma_8 \left(\frac{10}{8}\right)^{-0.8},$$  \hspace{1cm} (30)

and we receive the fruit of all our labour in the form of the value

$$\sigma_8 = 0.63.$$  \hspace{1cm} (31)

There are several things we have neglected to do here. We have not taken the uncertainties in the values of $n$ and $M$ into account and seen how they propagate to an uncertainty in $\sigma_8$. And we should not, unless we have strong reason to do so, assume an Einstein-de Sitter universe, but rather allow $\Omega_{m0}$ to vary. But still this example illustrates an important application of the mass function and some of the steps that must be taken to relate observations to theory.