UST-KASI, Daejeon, 20004

Field Research Course D: Cosmological likelihoods and parameter estimation

Instructors: Dr. David Parkinson

David Parkinson

2020/08/31

Contents

1 Introduction

The era of precision cosmology started when the cosmological parameters were measured at greater than 10% accuracy. This was about 1998, when type-Ia supernovae were first used accurately to measure the expansion, and balloon and ground-based CMB experiments first unambiguously detected the CMB anisotropies. This was really cemented with the first WMAP results (in 2003), and it has since become a semi-industry, where each new data set is analysed, by itself and combined with others, for the impact it will have on the inferences regarding the cosmological parameters.

In this course we will cover what the cosmological parameters are, how they connect to the physical mechanisms, and how we use cosmological observations to make probabilistic inferences as to their values. We will show how different data from different sources has been used to construct the concordance cosmological model (known as "ΛCDM"). We will demonstrate the methods to forecast the effectiveness of future cosmological surveys, and how the next generation of experiments will further narrow the parameter space. Finally we will discuss how the improved parameter constraints have led to the emergence of cosmic tensions between different experiments.

Some useful text books for this course include:

- *"Cosmological Physics"* by John Peacock, Cambridge University Press, Cambridge (1999)
- *"The Cosmic Microwave Background"* by Ruth Durrer, Cambridge University Press, Cambridge (2008)
- *"Dark Energy: Theory and Observations"* by Amendola and Tsujikawa, Cambridge University Press, Cambridge (2010)

This course is not meant as a replication of that material. If the student wishes to understand the physics of what is going on in the early universe or the late-time acceleration, these textbooks provide a good understanding. The purpose of this course is to learn about how the inferences and measurements of the cosmological parameters are actually made. The physics is important, but the content of this course is very much about how the observations are transformed into constraints, and so will focus on nomenclature and methods, rather than explanations.

During these notes I will make some dimensional analysis of the different quantities considered. The symbols that are used are the same as the SI symbols: $[M]$ for mass, $[L]$ for length, $[t]$ for time, and $[T]$ for temperature.

2 The Cosmological Parameters - ΛCDM and extensions

In this chapter we discuss what the cosmological parameters are, and how they relate to the physical effects that they refer to.

2.1 Learning outcomes

- Understand what the different symbols are
- Understand how they relate to the physical laws of cosmology

2.2 Meaning of symbols

2.2.1 Density, critical density, and physical density

There are three very common symbols used in cosmology to refer to density, which are ρ, Ω and ω. If we are dealing with only the homogeneous universe, on very large scales, then these parameters are only functions of time, or more commonly, redshift.

The first symbol ρ is the ordinary density as we have in physics. It is a dimension-full quantity, meaning that, what ever units you're using, the dimension will be mass per unit length cubed, or $[M][L]^3$. The scaling of the density ρ is given by the continuity equation

$$\frac{d\rho}{dt} = -3\frac{da}{dt}\frac{1}{a}\rho(1+w)\,, \tag{2.1}$$

where t is cosmological time (measured on a constant energy density hyper-surface), a is the cosmological scale factor, and w is the equation of state for that fluid. The continuity equation assumes that the fluid is not being created or destroyed, merely diluted by the expansion of the universe.

The different fluids will have different values of the equation of state w, and so their energy density will scale differently depending on the equation of state.

> Show how the physical density ρ scales with scale-factor a for pressureless matter ($w = 0$), radiation ($w = 1/3$) and the cosmological constant ($w = -1$).

The second symbol, Ω gives the density as a fraction of the critical density,

$$\Omega = \frac{\rho}{\rho_{\rm c}}\,, \tag{2.2}$$

and, as such is dimensionless. Since the critical density is defined as

$$\rho_{\rm c} = \frac{3H^2}{8\pi G}\,, \tag{2.3}$$

then both parts of the fraction in equation 2.2 will scale with expansion, and importantly, scale in different ways.

> Find the condition by which the density as a fraction of the critical density Ω will remain constant for all time. What value will it take?

Finally there is the so-called *physical density*, $\omega = \Omega h^2$, where $h = H/100\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. The reduced Hubble parameter h (also called 'little-h') is dimensionless, and so the physical density ω is also dimensionless. But it scales in the same manner as the ordinary density ρ, and be substituted for ρ in physical equations. Because it is more 'physical' in this manner, many cosmological probes are more sensitive to ω than Ω.

2.2.2 Power spectra

The distribution of fluctuations (of any type, not just density) as a function of spatial position and time is often summarised in terms of a power spectrum. If we can decompose the spatial distribution of fluctuations ($\delta(x)$), in terms of some Fourier basis, then the power spectrum is given by

$$P(k) = \frac{2\pi^2}{k^3}\langle\delta(k)^2\rangle\,, \tag{2.4}$$

where k is the wavenumber. The power spectrum has units of $[L]^3$, in the case the dimension of the fluctuation $\langle\delta^2\rangle$ is dimensionless.

> Show that if you Fourier transform the power spectrum you recover a quantity that varies with scale r, but is dimensionless.

However, in cosmology often the power spectrum will follow some power-law, with a constant spectral index. In this case, the power spectrum can be parameterised by two numbers, the amplitude A, and spectral index n, so that

$$P(k) = A\left(\frac{k}{k_0}\right)^n\,. \tag{2.5}$$

We see that the amplitude A is merely the value of the fluctuation evaluated at some 'pivot-scale' k_0. The amplitude of the fluctuation must therefore also have the same units as the power spectrum, $[L]^3$. The spectral index will be dimensionless. The scale-invariant case, where equal power is distributed on all scales, is set when $n = 0$.

Show that the spectral index n is independent of the choice of pivot scale.

Confusingly though, there are occasions in cosmology where the spectral index n is the not the tilt. In this case, the power spectrum is actually given by

$$P(k) = A \left(\frac{k}{k_0} \right)^{n-1} . \tag{2.6}$$

Now the scale-invariant case, where equal power is distributed on all scales, is now set when $n = 1$.

2.2.3 Dark matter physics

Though we do not know the exact mechanism for the creation of dark matter, it is most commonly assume to be a 'thermal relic'. This does not mean that it couples to the electromagnetic force, and so can be heated and cooled in the same way as baryonic matter. It rather means that it emerges from thermal equilibrium during a much hotter phase of the early universe.[1] During the early universe there were enough high energy particles for the cold dark matter particle to be constantly created and destroyed through interactions. When these interactions are in equilibrium, the dark matter particles are constantly being replenished. However, as the Universe expands, though, it becomes increasingly harder for a dark matter particle to find a partner to annihilate with annihilation reaction shuts off.

At this point, the DM density remains frozen in time. The 'freeze-out' time occurs when the annihilation rate, $\Gamma_{\text{inelastic}}$, is on the order of the Hubble rate, H:

$$\Gamma_{\text{inelastic}} = n_{\text{DM}} \langle \sigma v \rangle \sim H\,, \tag{2.7}$$

where n_{DM} is the dark matter number density and $\langle \sigma v \rangle$ is the velocity-averaged cross section. Cold Dark Matter is non-relativistic at freeze-out, with $n_{\text{CDM}} \sim T^{3/2} \exp^{-m_{\text{CDM}}/T}$, with T the temperature of the DM species; hot dark matter is relativistic at freeze-out, with $n_{\text{CDM}} \sim T^3$. Warm DM falls somewhere in between these two cases (Lisanti, 2017).

After freeze-out, the number of dark matter particles will be fixed, and the number density n will fall only due to the expansion of the Universe. The mass density of dark matter today, $\Omega_{\text{CDM}} h^2$ therefore relates directly to the time since freeze out, mass of the dark matter particle, and the cross-section of interactions.

[1] So it is 'thermal' in the statistical mechanics sense, rather than the infrared heat sense.

2.3 List of symbols

2.3.1 Parameters for cosmological parameter estimation

The parameters commonly used for parameter estimation are listed in table 2.1. Note that I haven't included the total matter density, as $\Omega_M = \Omega_{\mathrm{CDM}} + \Omega_{\mathrm{b}}$. Nor have I included the Cosmological constant density Ω_Λ, as, if you assume a flat universe, then $\Omega_\Lambda = 1 - \Omega_M$.

Table 2.1: List of commonly varied cosmological parameters. These are the minimal set needed to describe the Flat ΛCDM concordance cosmology.

Parameter name	Symbol	Physical meaning
Hubble parameter	H_0	Expansion rate at redshift zero
Cold Dark Matter density	Ω_{CDM}	Density of cold dark matter at redshift zero
Baryon density	Ω_{b}	Density of baryons at redshift zero
Spectra index	n_s	Tilt of power spectrum of density perturbations
Amplitude of fluctuations	A_s	Amplitude of power spectrum of density perturbations
Optical depth	τ	Optical depth to reionisation

There is one important parameter missing from this list, which is the temperature of the cosmic microwave background, T_{CMB}, measured at redshift zero. However, this was measured so well by the COBE-FIRAS experiment (Mather et al., 1994) that it is not commonly varied as part of any analysis, but instead fixed at the mean of the measured value, 2.72548 ± 0.00057 K (Fixsen, 2009).

2.3.2 Derived parameters

The parameters that can be derived from the ones commonly varied for parameter estimation are listed in table 2.2.

2.3.3 Extension parameters

The parameters that can be added to a cosmological analysis, to extend it beyond the standard flat ΛCDM model, are listed in table 2.3.

Table 2.2: List of commonly derived cosmological parameters.

Parameter name	Symbol	Physical meaning
Age of the Universe	t_0	Age of the Universe at redshift zero
Total matter density	Ω_{M}	Density of all pressureless matter at redshift zero
Cosmological constant density	Ω_Λ	Density of the cosmological constant at redshift zero
Galaxy clustering amplitude	σ_8	Amplitude of the power spectrum on the scale of 8 h^{-1}Mpc
Baryon to photon ratio	η	Ratio of the number of baryons to number of photons
Neutrino density	Ω_ν	Density of the massless neutrino sector
Photon density	Ω_γ	Density of photons
Dark matter cross section	σ_{DM}	
Redshift of reionisation	z_{re}	Redshift when the universe reionises
Cosmological constant	Λ	Value of the cosmological constant
Slow-roll parameters	ϵ and η	

Table 2.3: List of common cosmological parameters added as extensions to the concordance Flat ΛCDM model.

Parameter name	Symbol	Physical meaning
Curvature	Ω_k	Homogeneous curvature of the Universe, rescaled by the Hubble parameter
Dark Energy equation of state	w_{DE} (or w)	Equation of state of the dark energy
Running of the spectral index	α_s	Scale dependence of spectral index of power spectrum
Tensor to scale ratio	r	Ratio of the amplitudes of the tensor power spectrum to the scalar power spectrum
Dark matter mass	m_{DM}	Mass of a WIMP cold dark matter particle
Effective neutrino number	N_{eff}	Effective number of neutrino species
Neutrino mass	m_ν	Mass of one (or more) neutrino species

3 Distance measures

In this chapter we discuss what the cosmological distances measurements are, and how they are related to values of the density parameters.

3.1 Learning outcomes

- Learn about the different types of cosmological distances that can be measured
- Learn how these distances relate to the integrated expansion rate and large-scale curvature, and the parameters that define these

3.2 Photon trajectories

3.2.1 Null-geodesics

The most commonly used variant of the standard Freidmann-Lemeitre-Robertson-Walker metric used in cosmology today is hen we are left with

$$ds^2 = dt^2 - a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega \right) . \tag{3.1}$$

This is the comoving distance metric, where $a(t)$ is the scale factor (dimensionless, normalised to $a(t_0) = 1$ today), r is comoving distance (dimension-full, but assumed to be unchanging with time in a homogeneous universe), t is cosmological time, Ω here is the two-dimensional angular volume element $d\Omega = d\theta^2 + \sin^2 \theta d\phi^2$, and k is the curvature. Here k must have inverse units to r^2 (so that the term in the denominator of the fraction attached to the spatial part of the metric is dimensionless), and is sometimes called the Gaussian curvature.

> What values can the curvature parameter k take, and how do these values relate to the combined total density of matter, radiation and dark energy?

All distances in cosmology are measured by the trajectories of photons through spacetime. Therefore, because photons move along null-geodesics, we will set

$$ds^2 = 0 . \tag{3.2}$$

Note that this condition only for photons. If there was a different cosmic observable that involved massive particles (such as galaxies, cosmic rays or massive neutrinos) travelling large distances, then their velocity would be different to the speed of light, and this condition would not hold.

3.2.2 Radial motion

Since photons move along null-geodesics, we can simplify the distance that they travel to be given by the time of flight,

$$c^2dt^2 = a(t)^2\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega\right]. \tag{3.3}$$

We can further simplify this by only considering photons coming directly towards us. (After all, we will never detect photons that do not reach out detectors). In this case $d\Omega = 0$, and so we can simplfy again, to give

$$cdt = \frac{a(t)dr}{\sqrt{1-kr^2}}. \tag{3.4}$$

> Under what condition might a photon have some tangential motion, leading to $d\Omega \neq 0$, and still reach our detector?

Finally, since the start and end point of the photon trajectory are fixed in comoving coordinates, the r will not change with time. This means that equation 3.4 can be rearranged into a time dependent part and a time independent part

$$c\frac{dt}{a(t)} = \frac{dr}{\sqrt{1-kr^2}}. \tag{3.5}$$

We make a substitution, relating the scale factor $a(t)$ to the Hubble rate $H = \frac{da}{dt}\frac{1}{a}$, so that we can change variables on the left-hand side of equation 3.6

$$c\int\frac{da}{H(a)a^2} = \int\frac{dr}{\sqrt{1-kr^2}}. \tag{3.6}$$

Another quick change of variables from scale factor a to redshift z using

$$a = \frac{1}{1+z}, \tag{3.7}$$

and

$$da = \frac{-dz}{(1+z)^2}, \tag{3.8}$$

gives

$$c\int\frac{dz}{H(z)} = \int\frac{dr}{\sqrt{1-kr^2}}. \tag{3.9}$$

There are now three separate cases that need to be considered:

1. A flat universe

In this case the right-hand side of equation 3.9 is just equal to r, and so the solution is

$$r(z) = c\int_0^z \frac{dz'}{H(z')} . \tag{3.10}$$

2. An open universe

In the case of hyperbolic geometry, the gaussian curvature is negative, and so the solution of the right-hands side of equation 3.9 is given by

$$\int \frac{dr}{\sqrt{1-kr^2}} = \frac{1}{\sqrt{|k|}} \sinh\left(\sqrt{|k|}r\right) \tag{3.11}$$

This gives the solution for r to be

$$r(z) = \frac{1}{\sqrt{|k|}} \sinh\left(c\sqrt{|k|}\int_0^z \frac{dz'}{H(z')}\right) . \tag{3.12}$$

We can rewrite this in terms of some dimensionless curvature scale $R_k = \frac{H_0}{c\sqrt{|k|}}$.[1] This gives the solution for r to be

$$r(z) = \frac{cR_k}{H_0} \sinh\left(\frac{1}{R_k}\int_0^z \frac{dz'}{E(z')}\right) , \tag{3.13}$$

where $E(z) = H(z)/H_0$ is the normalised expansion rate.

> What units does this rescaled curvature parameter have, and how does it relate to the ‘curvature density’ parameter introduced in Table 2.3?

3. A closed universe

In the case of elliptical geometry, the gaussian curvature is positive, and so the solution of the right-hands side of equation 3.9 is given by

$$\int \frac{dr}{\sqrt{1-kr^2}} = \frac{1}{\sqrt{k}} \sin\left(\sqrt{k}r\right) \tag{3.14}$$

Again, we can re-arrange and rewrite this in terms of some curvature scale $R_k = \frac{H_0}{c\sqrt{k}}$. This gives the solution for r in a closed universe to be

$$r(z) = \frac{cR_k}{H_0} \sin\left(\frac{1}{R_k}\int_0^z \frac{dz'}{E(z')}\right) . \tag{3.15}$$

[1] Note that $R_k \to \infty$ as $k \to 0$. Physically we can envisage this as the hyperbolic surface becoming ‘flatter’.

What is the maximum distance that can be travelled by a photon in a closed Universe, using equation 3.15?

3.3 Distances

3.3.1 Co-moving radial distance

The comoving distance r is the same as that given in equations 3.10, 3.13, and 3.15 (depending on curvature).

3.3.2 Luminosity distance

Assuming that we know the intrinsic, absolute luminosity L_0 of a certain object, called a standard candle. Let us assume that we observe a flux S from the standard candle with absolute luminosity L at a fixed comoving distance r. The definition of the *luminosity distance* is then

$$S \equiv \frac{L_0}{4\pi d_L^2} \tag{3.16}$$

This is shown in Figure 3.1.

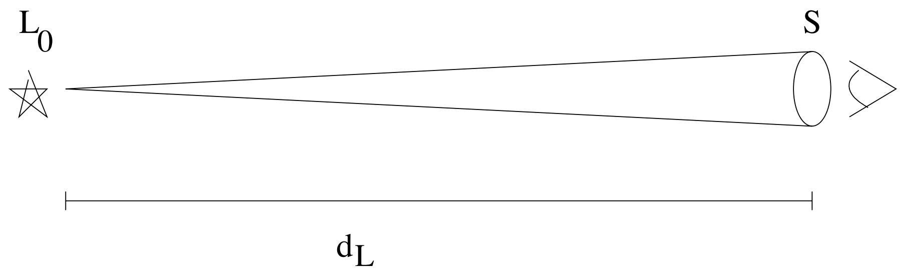

Figure 3.1: Luminosity distance, as defined as the dimming of a distant source by the distance travelled

In terms of the comoving distance, we can find[2]

$$d_L = (1+z)r\,. \tag{3.17}$$

The next question is how does the luminosity distance relate to the cosmological parameters? Let us write the Friedmann equation, in terms of the density parameters

$$H(z) = H_0 \left(\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\gamma(1+z)^4 + \Omega_\Lambda\right)^{1/2} \tag{3.18}$$

[2] See a textbook or cosmology course for a suitable derivation.

In the simple case of a flat, Einstein-de Sitter universe, $\Omega_m = 1$ and $\Omega_k = \Omega_\gamma = \Omega_\Lambda = 0$. In this case:

$$H(z) = H_0(1+z)^{3/2}\,, \tag{3.19}$$

and there is an analytic solution for the radial distance

$$r = \frac{2c}{H_0}\left(1 - \frac{1}{\sqrt{1+z}}\right)\,. \tag{3.20}$$

The luminosity distance is therefore

$$d_L = \frac{2c}{H_0}\left(1 + z - \sqrt{1+z}\right)\,. \tag{3.21}$$

> In this special case, the comoving distance and the luminosity distance exhibit very different behaviour as $z \longrightarrow \infty$. Describe this behaviour, and explain why it happens.

In the case where redshift is small, we can approximate $\sqrt{1+z} \simeq 1 + \frac{z}{2}$. This gives the luminosity distance at low-redshift to be

$$d_L = \frac{cz}{H_0}\,, \tag{3.22}$$

or, equivalently (and assuming $v = cz$),

$$v = H_0 d_L\,. \tag{3.23}$$

This is also known as the Hubble relationship, or Hubble law.

> Is there an analytic form of the luminosity distance in the case of a universe without Λ, but not necessarily flat?

3.3.3 Angular diameter distance

Now we consider the case where we observe an object with known (physical) diameter D under an angle θ. We assume again that the object is at a comoving distance r and the photons which we observe today were emitted at time t_1. From this we define the *angular diameter distance* via the relation $D = \theta d_A$. This is shown in Figure 3.2.

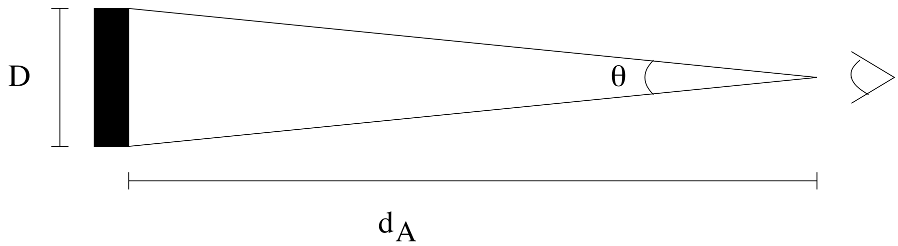

Figure 3.2: Angular diameter distance, as defined as the shrinking of a distant source by the distance of separation

4 Supernova, BAO, GW and other rulers

In this chapter we discuss how the cosmological distances are measured, focussing on type-Ia supernovae, baryon acoustic oscillations, and gravitation waves sirens.

4.1 Learning outcomes

- Learn how the different probes are 'anchored' with respect to the absolute physical measurements
- Understand how to use these different probes in a cosmological analysis

4.2 Luminosity distances

4.2.1 Type-Ia supernova

Type-Ia supernovae are a subtype of type-I, which are identified by a lack of hydrogen lines in their optical spectra. Type-Ia can be identified by the presence of the Si II λ6355 absorption feature at 6150Å.

But it is really the light-curve that is of most interest to cosmologists. As the light from an exploding supernova fades, it does so at a particular rate given by its absolute luminosity. This is know as the Phillips relation (Phillips, 1993), which is sometimes written as

$$M_{\rm max}(B) = -21.726 + 2.698\Delta m_{15}(B)\,, \tag{4.1}$$

where B is the B-band, M is the absolute magnitude and Δm_{15} is the difference in the B-magnitude light curve from maximum light to the magnitude 15 days after B-maximum (where the time here is measured in the galaxy rest frame, not the observer frame). So by continued monitoring the light from the supernova over a few weeks, the observer can determine the **absolute magnitude** of the supernova (M), and compare it to the observed **apparent magnitude** of the supernova (m), through the distance modulus.

$$\mu = m - M = 5\log_{10}\left(\frac{d_L}{10{\rm pc}}\right)\,, \tag{4.2}$$

where d_L is the same luminosity distance as defined in chapter 3 (equation 3.17).

Why is the distance modulus dependent on the base-10 logarithm, and why is it normalised to 10 parsecs?

Though the Phillips relation has since been replaced with something more accurate, it is clear that type-Ia supernova **not** standard candles. After all, if they were truly standard candles, the absolute magnitude would be independent of other measurable quantities, but the same for every supernovae. Instead, they are **standardisable candles**. That is, by measuring the colour and stretch parameters of each supernovae for the MLCS relation (Riess et al., 1996) for example, all type-Ia can be put on the same footing, and so the distance modulus for each can be 'standardised.'

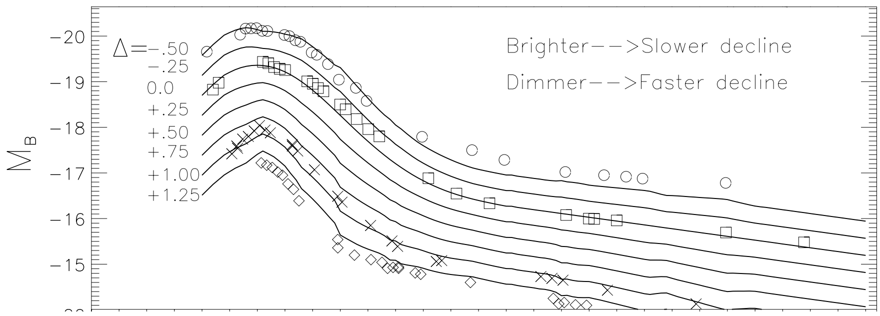

Figure 4.1: Figure showing how the relationship between time and brightness changes for intrinsically brighter and fainter supernova (frrom Riess et al. (1996)).

However, when we use SN-Ia as standard candles, we are making two important assumptions: firstly, that the luminosity of the supernova do not feel the effects of any external environment (such as the galaxy or the nature of the cosmological universe) and so will be the same everywhere in the Universe and at all times, and secondly, that it is well 'anchored' at low redshift.

This question of low-redshift anchor is important, as no type-Ia supernova have been observed in our own galaxy (the Milky Way). As such, the true value of M is not known exactly, and is in fact completely degenerate with the value of the Hubble parameter at redshift zero H_0.

If what we measure is not distance modulus μ but rather apparent magnitude m, use the definition of luminosity distance to demonstrate this degeneracy.

If the distances to the extragalactic supernovae can be fixed through some other method, and so their absolute brightness determined to a very high accuracy, then the Hubble parameter can also be determined. This is the current research of the SH0ES (Supernovae, H_0, for the Equation of State of Dark energy) project (Riess et al., 2016), which uses cepheids and megamasers to determine the distances to galaxies containing supernovae in an independent fashion.

4.2.2 Gravitational waves

Gravitational wave sources can also be used a standard candles. As first suggested by Schutz (1986), the intrinsic luminosity of gravitational wave-generating inspiral events, such as black hole binaries, or binary neutron star systems, can be determined exactly the gravitational wave waveform. As such, and though they are sometimes referred to as 'standard sirens', they are still standard candles, even if the radiation they are producing is gravitational, rather than electromagnetic.

4.3 Angular diameter distances

Angular diameter distances are much harder to measure than luminosity distances, as any object that is large enough to have its apparent size measured will need to be of galaxy size or larger, and therefore likely to be affected by the environment and universe it lives in. There are no reliable objects that have the same size over all cosmic time, except for baryon acoustic oscillations.

4.3.1 Baryon acoustic oscillations

The sound horizon

The baryon acoustic oscillations are the fluctuations in the distribution of galaxies generated by the pressure waves (literally 'sound waves') generated in the early universe. These sound waves propagate through the photon-baryon plasma until the speed of sound in the plasma drops to zero. This finite amount of time that they have available to travel will set a maximum distance that they can travel, which is called the *sound horizon*.

This may be hard to visualise, so imagine a stone dropped into a perfectly still pool of water. The ripple from the stone will propagate out from the point of impact, until it reaches the edge of the pool. Now imagine that before this happens (before the ripple reaches the edge), that someone pours liquid nitrogen over the surface, freezing the water in place, and fixing the position of the ripple relative to the point of impact. The time difference (t) between the stone entering the pool and the pool being frozen, and the speed of sound in the fluid (v_s) will determine how far the ripple can move. This is illustrated in figure 4.2.

For a cosmic sound wave, travelling through the photon-baryon plasma, there is not a single point of impact, but rather the source of sound wave is a gravitational instability, which are present everywhere. So the ripples are often crossing over each other, and this maximum radius can only be determine statistically, by comparing the separation between objects over many different centres. In practice, we use a *correlation function* to measure this scale. This is illustrated in a cartoon created by the Baryon Oscillation Spectroscopic Survey (BOSS) in figure 4.3.

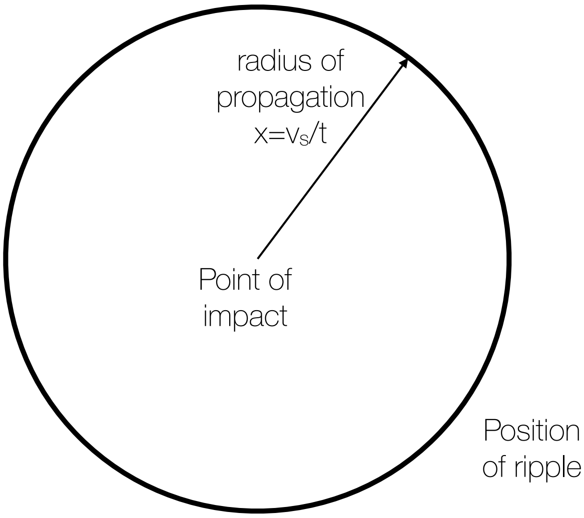

Figure 4.2: Figure showing how the ripple radius depends on the speed of sound.

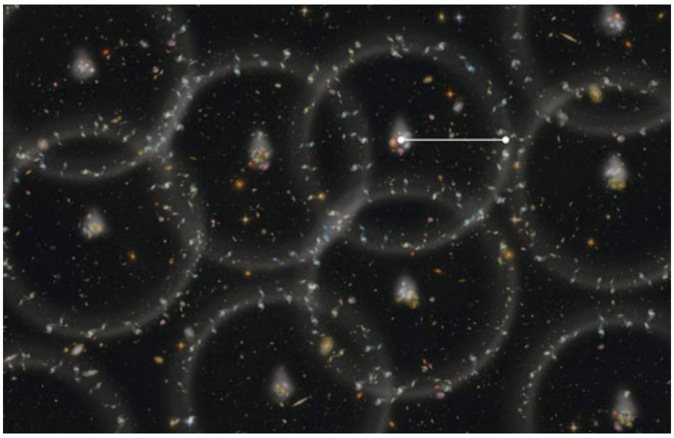

Figure 4.3: A cartoon produced by the BOSS project showing the spheres of baryons around the initial dark matter clumps.

The sound horizon scale is given by s, such that

$$s = \int_{z_{\rm rec}}^{\infty} \frac{c_s dz}{H(z)}\,, \tag{4.3}$$

where c_s is the speed of sound in the plasma, and $H(z)$ is the Hubble parameter at that redshift (as predicted by the Freidmann equation). We can make a prediction for the sound horizon, and assume that the universe is matter dominated over most of the time that the wave is propagating, to give

$$s = \int_{z_{\rm rec}}^{\infty} \frac{c_s dz}{H(z)} = \frac{1}{\sqrt{\Omega_m H_0^2}} \frac{2c}{\sqrt{3 z_{eq} R_{eq}}} \ln \left[\frac{\sqrt{1+R_{rec}} + \sqrt{R_{rec}+R_{eq}}}{1+\sqrt{R_{eq}}} \right]\,, \tag{4.4}$$

where $R \equiv 3\rho_b/4\rho_\gamma \propto \Omega_b h^2/(1+z)$, $z_{eq} = \Omega_m/\Omega_{rad}$ is the redshift of matter-radiation equality and "rec" refers to recombination. The CMB strongly constrains the matter and baryon densities at decoupling and hence the sound horizon, $s = 146.8 \pm 1.8$Mpc.

All BAO measurements quoted in the literature are normalised or 'anchored' in terms of this sound horizon. This symbol s is often written as r_s or r_d. However, the dependence on the scale of the sound horizon means that the standard rulers are not independent on the baryon density. After all, if the baryon density is larger, the speed of sound will be larger, and so the sound horizon will be larger, even if the amount of time remains fixed. This means that, when constraining the cosmological parameters using BAO, the baryon density $\Omega_b h^2$ needs to be included as a *nuisance parameter.*

BAO measurements

5 Sampling likelihoods and exploring degeneracies

In this chapter we discuss how the cosmological parameters are estimated, and their probability distributions are sampled.

5.1 Learning outcomes

- Understand what the likelihood actually is
- Understand what the prior and posterior actually are
- Learn how to sample from the likelihood/posterior, and marginalise

5.2 Likelihood

5.2.1 Definition

The definition of the likelihood distribution is commonly written as a probability density function, given by

$$P(D|\theta, M)\,, \tag{5.1}$$

where D is a random variable that represents the *data*, and θ and M are fixed terms that represent the *free parameters* and *model*.[1] The likelihood is a probability, and so it must be normalised, in this case with respect to the data D, such that

$$\int P(D|\theta, M)\mathrm{d}D \equiv 1\,. \tag{5.2}$$

The likelihood is the probability of observing that data, given that the null-hypothesis is true. So in this case, we can say that the null-hypothesis is that the physical universe obeys those physical laws specified in M, and the unknown parameters take the values specified in θ.

A good example of a likelihood function is the binomial distribution, which models the distribution of outcomes (data) for a coin flip. If we have N flips, k heads, and the probability of a single head is given by π_H, then the probability of that outcome is

$$P(k|N,\pi) = \binom{n}{k}\pi_H^k(1-\pi_H)^{n-k}\,. \tag{5.3}$$

[1]It may sound strange, paradoxical even, to say that the free parameters θ are fixed, but bear with me.

This really is the probability of that particular outcome, not a probability distrtibution for π. It assumes that π is fixed in value, unchanging no matter how many times we flip the coin, and not a random variable. In effect, the value of π_H is a *law of physics* in this universe.

> Show that if we sum over all values of k, the total is unity, confirming equation 5.2.

5.2.2 Gaussian likelihoods

Gaussian likelihoods are really common, and very popular. The reason for this is the *central limit theorem*. Basically the theorem states that, no matter what the probability distribution of a single random variable is, if you have enough of them, and they are independent, the sum will tend towards a Gaussian.

In cosmology, this situation is actually pretty common. Consider:

- a supernova magnitude, meausred from many independent photons, bouncing through the optical system of the telescope, onto the detector,
- a gravitational wave strain, measured over a large number of independent cycles,
- a baryon acoustic oscillation scale, measured from the average separations of a large number of galaxy pairs.

For all of these cases, no matter what the individual probability distribution was for each of these things (photon energy, detector strain, or galaxy separation), the final reported result is well described by a Gaussian.

> What situations exist in cosmology where a Gaussian is not a good description of the data?

The likelihood function, that can be used to compute the probability distribution of Gaussian data, is the χ^2. The χ^2 value for any data set consisting of N independent Gaussian distributions (*'measurements'*) with mean μ and variance σ^2 can be computed using

$$\chi^2 = \sum_{i=1}^{N} \frac{(\mu_i - E_i)^2}{\sigma_i^2}\,, \tag{5.4}$$

where E is the expectation of the theoretical value that μ should take, given some null hypothesis.[2] We often write this as

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{\sigma_i^2}\,, \tag{5.5}$$

[2]So E should be independent of the actual value of μ, but might be related to some control variable. Consider a SN-Ia data set, where the distance modulus depends on redshift z. The prediction for the distance modulus doesn't depend on the measured value of the distance modulus, but does depend on the redshift at which it was measured.

where O refers to observed value. The probability density function of the χ^2 is given by

$$P(\chi^2) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} \left(\chi^2\right)^{\frac{k}{2}-1} \exp\left(-\frac{\chi^2}{2}\right), \quad \text{for} \quad \chi^2 > 0 \tag{5.6}$$

where k is the number of degrees of freedom. The importance of k we will come back to in the next section, but it is basically the number of independent Gaussian distributions, $k \approx N$. So if we're using the χ^2 as a test statistic, we can compute the cumulative probability distribution to estimate the probability of finding a χ^2 equal to, or greater than, some value. This probability is known as the p-value. If that probability is too small, we can then suggest that the probability of that χ^2 value is inconsistent with the null hypothesis, which might be an argument for rejecting it.

> What is the relationship between the commonly used χ^2 per degree of freedom argument, and the pdf in equation 5.6?

However, I am now about to pull the rug out from under you.

You have hopefully realised by now that equation 5.6 is **not** the probability density function for the data. It is the pdf of the χ^2 of the data, and this is not the same thing as the data itself.

In fact, the probability density function of a set of independent Gaussian variables (as we are considering the data, because remember the data is the only thing that is random, at the moment) that we denote as an ensemble $\{D\}$, is just exactly that

$$P(\{D\}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(y-\mu_i)^2}{2\sigma_i}\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\sum_{i=1}^{N} \frac{(y-\mu_i)^2}{2\sigma_i}\right), \tag{5.7}$$

assuming that all values of σ_i are the same in the second step. Here y is the value of the thing being measured, but remember, this takes a fixed value under the null-hypothesis. Since this is the case, we can immediately relate the likelihood to the χ^2.

$$P(\{D\}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\chi^2}{2}\right) \propto e^{-\chi^2/2} \tag{5.8}$$

We have therefore shown why we use the χ^2. It is a shorthand for the probability density function of the data given the null hypothesis, and can also be used as a test statistic for the data itself.

5.2.3 Maximisation

What happens when we want to test different null-hypotheses? Simply put, we can compute different values of the likelihood for the different null-hypotheses, and find the hypothesis that maximises the likelihood (and in the Gaussian case minimises the χ^2).

A good example might be a family of null-hypotheses each with a different value of some parameter. This parameter is not a random variable here, but has a real and fixed value relating to some process that produces the data. A nice example is the probability of heads on a coin, π_H that we discussed in section 5.2.1. This process of finding some best fit likelihood for some parameter value or model is literally called that ("finding the best fit"), but also as **maximisation**.

> Suppose that an unknown coin is flipped 100 times, and produces 13 heads. How would you use likelihood maximisation to estimate the probability of heads, π_H, for that coin?

As an example, let us consider the original Hubble data, that was used to determine that the Universe was expanding (Hubble, 1929). We plot it in figure 5.1, with 10% distance errors, to be generous.

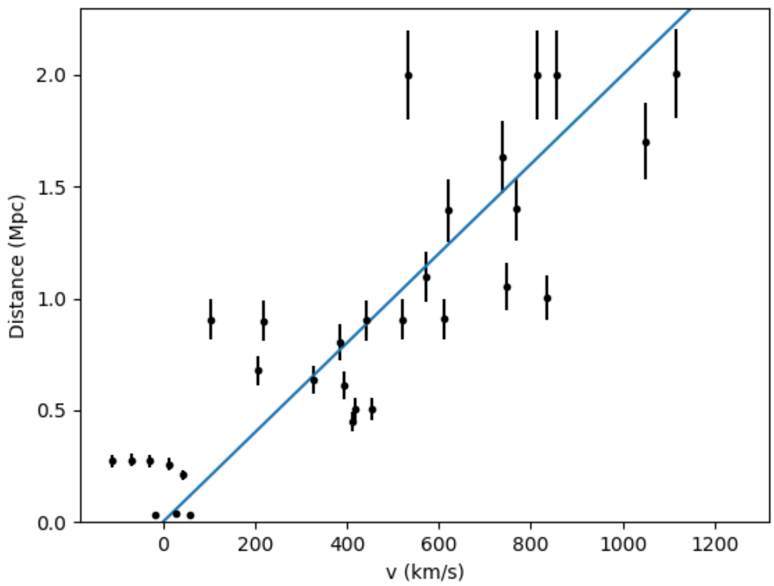

Figure 5.1: Original data for the Hubble expansion, take from Hubble (1929). The axes have been flipped compared to the original figure, and 10% distance errors have been added.

The law that Hubble was demonstrating with this data was the relation that now bears his name

$$v = H_0 D\,. \tag{5.9}$$

We can measure the likelihood of this data, assuming different null-hypotheses. In each hypothesis we test, we assume the Hubble relation is true, and we a test a specific value of H_0. This is called 'fitting' for the parameter. Once we find the parameter that minimises the χ^2 (or maximises the likelihood), we have fitted for the value of H_0. But remember, for each different test, we are still assuming that H_0 is fixed. The Hubble parameter doesn't have a probability associated with it, the data has a probability associated with it, assuming a value of H_0.

> Use the data from the figure above to find the best fit value of H_0. Does it agree with the value we have measured today?

5.2.4 Confidence intervals and coverage

5.3 Posterior probabilities

5.3.1 Bayes theorem

Wait... this entire process of simulating data and computing intervals seems very convoluted! What if we changed our perspective a bit, and decided that probability has a different definition? Would that make things easier, and more intuitive?

Enter two dead white European males: Pierre-Simon, marquis de Laplace, and the Reverend Thomas Bayes.[3]

Laplace and Bayes both had the same idea, that probability should refer to your level of ignorance, and can be applied to anything, not just the outcome of experiments. In his *"A Philosophical Essay on Probabilities"* (Brookes, 1953), Laplace wrote

> "Probability is relative, in part to [our] ignorance, in part to our knowledge."

So if our level of knowledge is incomplete, we can ascribe a probability to this knowledge. This is the basis of Bayes' theorem.

5.3.2 Posterior samples

[3] In another universe, there is a TV show where the old Bayes his young companion Laplace solve crimes using probability and statistics. Alas, no one in this universe has made that show yet.

6 Project 1: what is this universe?

In this project you will use SN-Ia, BAO and GW data to determine the parameters of the universe. This is fake data for a hypothetical universe, and so may not correspond to the cosmology of our Universe as we currently understand it. For this project you will use the techniques you have learnt to answer certain questions, and write a short report.

6.1 Questions

- What are the values of the cosmological parameters?
- How old is the Universe?
- Is it geometrically flat, open or closed?
- What is the eventual fate of this Universe?

6.2 Cosmological parameters

The scientists before you have discovered that dark matter exists, and have determined the baryon to dark matter density ratio has this value:

$$f_{\mathrm{baryon}} = 0.167 \tag{6.1}$$

The have also detected the cosmic microwave background, and measured the temperature to be

$$T_{\mathrm{CMB}} = 1.35K \tag{6.2}$$

Finally, by measuring the ages of cold white dwarf stars, they have found that there exist stars with an estimated age of at least 15 billion years.

6.3 Datasets

There are three different compilations of data that you will use. They are listed below.

6.3.1 Type-Ia supernova

The type-Ia supernovae data set you are using is called 'Parsonage'. It consists of 974 type-Ia supernova. It can be found at https://www.dropbox.com/s/isjt6gq27x6icpj/parsonage.txt?dl=0

6.3.2 Baryon acoustic oscialltions

There are three BAO datasets that you will use: 7dfGS, WDSS LRG sample, and EXMOS.

Table 6.1: GW data

Name	type	Redshift	Measurment	Error
7dfGS	$\frac{r_s}{D_V}$	0.0115	0.08	0.003
WDSS LRG	$D_V\left(\frac{r_{d,\text{fid}}}{r_d}\right)$	0.157	1906.18	66.577
EXMOS	$D_M\left(\frac{r_{d,\text{fid}}}{r_d}\right)$	0.389	3905.35	58.667
	$H\left(\frac{r_d}{r_{d,\text{fid}}}\right)$	0.389	43.56	1.310
	$D_M\left(\frac{r_{d,\text{fid}}}{r_d}\right)$	0.536	4761.80	68.866
	$H\left(\frac{r_d}{r_{d,\text{fid}}}\right)$	0.536	50.83	1.531
	$D_M\left(\frac{r_{d,\text{fid}}}{r_d}\right)$	0.626	5175.36	77.639
	$H\left(\frac{r_d}{r_{d,\text{fid}}}\right)$	0.626	57.19	1.336

These projects all use the same r_d, a value of 101.19Mpc.

6.3.3 Gravitational waves

Only two gravitational wave events with electromagnetic counterparts have been detected so far: GW810929, and GW811137.

Table 6.2: GW data

Name	Redshift	Luminosity distance	Error
GW810929	0.0101	150.9 Mpc	14
GW811137	0.0067	95.4 Mpc	14.5

6.4 Report

You will write a short report, explaining the aim of the project, what you have done, and answering the questions. The report should be around 2500 - 5000 words. The deadline is October 16th.

7 Dark energy parameterisations

In this chapter we discuss what the dark energy equation of state is, how it can be used to choose between different models, and how it is parameterised.

7.1 Learning outcomes

- Learn about the dark energy equation of state, and what values it has to take for the Universe to be accelerating
- Learn about about the different energy conditions in relativity theory
- Discuss the different parameterisations, and how they relate to dark energy models.

8 Non-parametric methods

In this chapter we discuss what non-parametric methods are, and how we can use them when parametric methods fail us, or when alternative approaches are necessary.

8.1 Learning outcomes

- Understand what the difference between parametric and non-parametric is
- Understand how we cannot apply a likelihood analysis approach to non-parametric methods
- Learn how to apply a simple non-parametric approach (cubic spline) to the data

8.2 Parametric vs non-parametric

9 Forecasting and the Fisher matrix

In this chapter we learn about what the Fisher matrix is, and how it can be used to forecast the effectiveness of cosmological observations in measuring the parameters.

9.1 Learning outcomes

- Learn about what the Fisher matrix is, and how it is used
- Apply the Fisher matrix to measurements of cosmological distances
- Compare it to posteriors computed from real data, and learn about the Cramer-Rao bound

9.2 The Fisher Information Matrix

9.2.1 The posterior distribution in the Gaussian case

The Fisher Information Matrix is a way of quantifying the amount of information known about some parameter or set of parameters θ in the context of some random variable X. In Bayesian statistics, we normally interpret this as the size of the posterior probability distribution on some unknown physical parameters, when measured using some data. So θ are the parameters of the model, and X becomes D the observable data.

Here is a reminder of Bayes' theorem.

$$P(\theta|D,M) = \frac{P(D|\theta,M)P(\theta|M)}{P(D|M)}\,, \tag{9.1}$$

where θ are the model parameters, M is the model, and D is the data. Here $P(\theta|D,M)$ is the *posterior* (the probability distribution of the parameters given that the data and the model are true), $P(\theta|M)$ is the *prior* (the probability distribution of parameters, given that the model is true, before any data is taken), and $P(D|\theta,M)$ is the *likelihood* (the probability of the data, given that the model is true, and the parameters are fixed at their specific value). We will ignore the normalising $P(D|M)$, called the *evidence*, for now, as it is not relevant.

In the case where the amount of data is large, the central limit theorem will apply, and so the likelihood can be modelled by a Gaussian. But in this case, the prior can also be assumed to be subdominant (making almost no contribution to the posterior distribution), and so the posterior can also be modelled by a Gaussian. Thus, if we have a set of parameters $\vec{\theta}$, which are all well constrained, then the posterior will be given by

$$P(\vec{\theta}|D,M) = \frac{1}{\sqrt{2\pi|C|}} \exp\left[-\frac{1}{2}(\vec{\theta}-\vec{\theta}^*)^T C^{-1}(\vec{\theta}-\vec{\theta}^*)\right], \tag{9.2}$$

where $\vec{\theta}^*$ are the maximum posterior values of the parameters, and C is the covariance matrix of the parameters. In the case of a single parameter $\vec{\theta} = \theta$, the covariance matrix will be a single value, the variance of the distribution on that parameter $C = \sigma$.

In this case, the Fisher matrix is given by

$$F = C^{-1}. \tag{9.3}$$

So as the power of the data increases, the covariance matrix entries will become smaller (as the dispersion of the distribution of those parameters shrinks). As the covariance matrix becomes smaller, the Fisher matrix grows, and so we learn more about the parameters.

> Remember, more information means larger Fisher matrix, and so smaller covariance matrix.

9.2.2 Computing the Fisher matrix

This may not seem useful. After all, surely you need to compute the covariance matrix first, in order to compute the Fisher matrix. Why the extra step of inverting it?

The answer is that it may be possible to compute the Fisher matrix quickly, without needing to sample the posterior distribution. And we can do this by using the Hessian matrix[1] of the posterior.

First we take twice the negative of the logarithm of the posterior. Using equation 9.2, we find

$$-2\log P(\vec{\theta}|D,M) = (\vec{\theta}-\vec{\theta}^*)^T C^{-1}(\vec{\theta}-\vec{\theta}^*) + \log(2\pi|C|). \tag{9.4}$$

If we are at a posterior maxima, $\vec{\theta} = \vec{\theta}^*$, and the negative log posterior will be a mimima. Since we are at a maxima, the gradients with respect to the parameters be zero, but the Hessian (curvature matrix will be non-zero).

$$\frac{\partial^2(-2\log P(\vec{\theta}|D,M))}{\partial\theta_i\partial\theta_j} = C^{-1} = F \tag{9.5}$$

So we see that the Fisher matrix is equivalent to the Hessian matrix of the (negative log) posterior distribution. Here θ_i and θ_j are individual parameters inside the set of parameters $\vec{\theta}$.

9.3 The Fisher matrix and cosmology

[1] In mathematics, the Hessian matrix or Hessian is a square matrix of second-order partial derivatives of a scalar-valued function, or scalar field https://en.wikipedia.org/wiki/Hessian_matrix

10 Project 2: what is the best experiment for this universe?

In this project, still set in the Universe outlined in part 1, you will *forecast* the effectiveness of one of three possible future cosmological experiments at determining the values of the cosmological parameters. You will be working based on the discoveries of part 1, and using either MCMC or Fisher matrix

For this project you will use the techniques you have learnt to answer certain questions, make a short presentation summarising parts 1 and 2, and write a short report.

10.1 Questions

- Can this experiment be used to determine the parameters of the Universe better than the previous datasets?
- Will this decisively whether this Universe will recollapse?
- How has structure formation been proceeding in this Universe?

10.2 Cosmological parameters

As before, it is part of the cosmological model that dark matter exists, and scientists have determined the baryon to dark matter density ratio has this value:

$$f_{\mathrm{baryon}} = 0.167 \pm 0.03 \tag{10.1}$$

The have also detected the isotropic cosmic microwave background, and measured the temperature to be

$$T_{\mathrm{CMB}} = 1.35K \tag{10.2}$$

Finally, by measuring the ages of cold white dwarf stars, they have found that stars exist with an estimated age of at least 15 billion years.

You will also assume a fiducial cosmology for the forecast, based on your results from part 1.

10.3 Experiments

There are three different proposed experiments, using different types of cosmological data, of which you will forecast one. They are listed below.

10.3.1 Cosmic microwave background

The CMB anisotropy experiment will be called ‘Stave’. It will be an all-sky cosmic variance-limited experiment out to $\ell = 300$.

10.3.2 Baryon Acoustic Oscillations

The BAO survey, called CURE (Collapsing Universe Redshift Explorer) will measure the BAO quantities in the following redshifts

10.3.3 Weak lensing shear survey

The weak lensing survey, called ‘Lobacheski’, will measure the weak lensing shear power spectrum in five redshift bins. It will have a survey area of $15000\mathrm{deg}^2$, with a galaxy surface density of $n_{\mathrm{gal}} = 30\mathrm{arcmin}^{-2}$. The noise generated by intrinsic ellipticity will be $\sigma_\epsilon = 0.21$ (per component). The equation for the number distribution of lensed galaxies is given by

$$n(z) \propto \left(\frac{z}{z_0}\right)^2 \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right], . \tag{10.3}$$

Here $z_m = 0.9/\sqrt{2}$, and the normalisation can be computed such it meets the required surface density. The bins are defined with redshift bounds such that they are equally populated, so the density of each will be $n_{\mathrm{gal}}^{\mathrm{bin}} = 6\mathrm{arcmin}^{-2}$.

10.4 Report

You will write a short report, explaining the aim of the two projects, what you have done, and answering the questions. The report should be around 2500 - 5000 words. The presentation should be a maximum of 15 minutes. The deadline for both is December 18th.

11 Power Spectra

In this chapter we learn about what the power spectrum is, and how it can be measured and predicted.

11.1 Learning outcomes

- Learn about the spatial power spectrum
- Learn about the different basis sets that fluctuations can be decomposed into
- Learn about shot noise, sample variance, and cosmic variance

11.2 Power and fluctuations

Welcome to the inhomogeneous universe!

So far we have assumed that the universe is homogeneous, and though the density can vary in time, it is doing so in a smooth manner such it has the same value everywhere in space. Of course, this is completely unphysical, as we now that gravity is dynamically unstable. Although a smooth universe will remain so, as the balance of forces will leave all particles in the same place, any tiny deviation from perfect smoothness will naturally grow. As the density increases in one place, the gravitational forces will become anisotropic, attracting more particles towards the inhomonegeneity, and increasing the gravitational force further. Thus we have a runaway growth of density.

Today the Universe we see is full of structure, but how do we characterise it? We do that with density fluctuations, given by δ, which is defined as

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}\,, \tag{11.1}$$

where $\bar{\rho}(t)$ is the homogeneous density as a function of time, and $\rho(\vec{x}, t)$ is the inhomogeneous density, which varies in space as well as time.

11.2.1 Decompositions, and power spectra

From a theoretical perspective, is very difficult to make predictions for the quantities as a function of position, and much easier to work in Fourier space. Part of this is because the fluctuations evolved independently to linear order, and part of this is because when you Fourier transform the

12 Cosmic Microwave Background

In this chapter we learn how to make likelihood calculations using the CMB power spectra.

13 Weak lensing

In this chapter we discuss what weak lensing shear data is, how it is used, and how we make predictions for it.

13.1 Learning outcomes

- Understand what the weak lensing shear angular power spectrum is
- Understand how we predict the shear power spectrum
- Learn how we estimate the errors

13.2 Angular power spectra

Similar to the cosmic microwave background, the distribution of shear on the sky can be decomposed by spherical bessel function. This is the weak lensing shear angular power spectra, defined as

To first approximation, the achievable errors on C_ℓ scale simply with the survey area of sky, the sky number density of galaxies available $\bar{n}_{gal}$ and the rms variance of the galaxies' ellipticity distribution σ_ϵ:

$$\sigma_{C_\ell} = \sqrt{\frac{2}{(2\ell+1)f_{sky}}}\left(C_\ell + \frac{\sigma_\epsilon^2}{\bar{n}}\right), \tag{13.1}$$

14 Tensions and Extensions

Bibliography

Amendola, L. and S. Tsujikawa (2010). *Dark Energy: Theory and Observations.* Cambridge University Press.

Brookes, B. C. (1953). A philosophical essay on probabilities. by pierre simon, marquis de laplace. translated from the sixth french edition by f. w. truscott and f. l. emory. pp. viii, 196. $1.25 (paper)$;2.50 (cloth). 1951. (dover publications, new york). *The Mathematical Gazette 37*(320), 148?149.

Durrer, R. (2008). *The Cosmic Microwave Background.* Cambridge: Cambridge University Press.

Fixsen, D. J. (2009, nov). THE TEMPERATURE OF THE COSMIC MICROWAVE BACKGROUND. *The Astrophysical Journal 707*(2), 916–920.

Hubble, E. (1929). A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences 15*(3), 168–173.

Lisanti, M. (2017, January). Lectures on Dark Matter Physics. In J. Polchinski and et al. (Eds.), *New Frontiers in Fields and Strings (TASI 2015*, pp. 399–446.

Mather, J. C., E. S. Cheng, D. A. Cottingham, J. Eplee, R. E., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, J. Moseley, S. H., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson (1994, January). Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument. *ApJ 420*, 439.

Peacock, J. A. (1999, January). *Cosmological Physics.* Cambridge, UK: Cambridge University Press.

Phillips, M. M. (1993, August). The Absolute Magnitudes of Type IA Supernovae. *ApJ Lett. 413*, L105.

Riess, A. G., L. M. Macri, S. L. Hoffmann, D. Scolnic, S. Casertano, A. V. Filippenko, B. E. Tucker, M. J. Reid, D. O. Jones, J. M. Silverman, and et al. (2016, Jul). A 2.4 *The Astrophysical Journal 826*(1), 56.

Riess, A. G., W. H. Press, and R. P. Kirshner (1996, December). A Precise Distance Indicator: Type IA Supernova Multicolor Light-Curve Shapes. *ApJ 473*, 88.

Schutz, B. F. (1986, September). Determining the Hubble constant from gravitational wave observations. *Nature 323*(6086), 310–311.