State University of Londrina
Center of Exact Sciences
Physics Department

Lucas Gomes de Oliveira Corbanez

Statistical inference techniques applied to cosmology

Londrina
2023

State University of Londrina

Center of Exact Sciences
Physics Department

Lucas Gomes de Oliveira Corbanez

Statistical inference techniques applied to cosmology

Monograph supervised by Prof. Dr. Sandro Vitenti, entitled "Statistical inference techniques applied to cosmology" and presented to the State University of Londrina as part of the requirements for obtaining a Bachelor's degree in Physics.

Advisor: Prof. Dr. Sandro Vitenti

Londrina
2023

Cataloging Record

Lucas Gomes de Oliveira Corbanez
Statistical inference techniques applied to cosmology - Londrina, 2023 - 54 p., 30 cm.
Advisor: Prof. Dr. Sandro Vitenti
1. Cosmology. 2. Statistics. 3. Supernovae.
I. State University of Londrina. Physics. II. Statistical inference techniques applied to cosmology.

Lucas Gomes de Oliveira Corbanez

Statistical inference techniques applied to cosmology

Monograph presented to the Physics Course at the State University of Londrina as a partial requirement for obtaining a Bachelor's degree in Physics.

Examination Committee

Prof. Dr. Sandro Vitenti
State University of Londrina
Advisor

Prof. Dr. Paula Fernanda Beinzobaz
State University of Londrina

Prof. Dr. Fabio Luiz Melquiades
State University of Londrina

Londrina, May 20, 2023

To all those who, in some way, contributed
to the realization of this work.

Acknowledgments

I would like to express my heartfelt thanks to those who have been part of this journey of learning and personal growth. Firstly, my eternal gratitude to my family, whose love and support have always been my safe haven. To my girlfriend, I have no words to thank you for all the support, encouragement, and unwavering patience in reviewing my texts over the years.

I would also like to thank the friends I made during college. Our friendship and companionship were essential to my academic and personal development, and I am very grateful to have had you by my side.

I couldn't fail to mention my teachers, who were true mentors on this journey. In particular, I would like to thank Prof. Dr. Thiago Pereira, who inspired me from the beginning to explore the fascinating field of cosmology, and Prof. Dr. Sandro Vitenti, who agreed to guide me in this independent work despite all the complications. I also thank Prof. Dra. Paula Bienzobaz and Prof. Dr. Fabio Melquiades for being the examining board and for all the shared knowledge.

Finally, I thank everyone who, in some way, contributed to this special moment in my life. I know this is just the beginning of a long journey, but I am confident that the lessons learned and the connections established will accompany me throughout my trajectory.

"Asking for help isn't giving up,' said the horse. 'It's refusing to give up."
— Charlie Mackesy The Boy, the Mole, the Fox and the Horse

Lucas Gomes de Oliveira Corbanez. **Técnicas de inferência estatística aplicadas a cosmologia**. 2023. 54 p. Trabalho de Conclusão de Curso em Física - Universidade Estadual de Londrina, Londrina.

Resumo

Este trabalho tem como foco a aplicação de técnicas de inferência estatística em cosmologia, com ênfase particular no uso de dados de supernovas do tipo Ia (SNe Ia) para estimar parâmetros cosmológicos. O estudo começa com uma introdução ao pano de fundo matemático necessário, incluindo geometria diferencial e relatividade geral, que formam a base da cosmologia moderna. Abordamos uma visão geral de diferentes métodos estatísticos, incluindo estimadores estatísticos, um breve estudo de estatística bayesiana e testes de hipóteses. Este trabalho também discute a física subjacente que governa as SNe Ia e por que elas são caracterizadas como "velas padronizáveis". Finalmente, os conceitos estatísticos são aplicados à estimativa de parâmetros cosmológicos usando um Teste de Razão de Verossimilhança, que é implementado em um script `Python`. Dados de SNe Ia obtidos do Supernova Legacy Survey (SNLS) são usados para of fit dos Estimadores de Máxima Verossimilhança e as regiões de confiança.

Palavras-Chave: 1. Cosmologia. 2. Estatística. 3. Supernovae.

Lucas Gomes de Oliveira Corbanez. **Statistical inference techniques applied to cosmology**. 2023. 54 p. Monograph in Physics - Londrina State University, Londrina.

Abstract

This work focuses on the application of statistical inference techniques in cosmology, with a particular emphasis on the use of Type Ia supernovae (SNe Ia) data to estimate cosmological parameters. The study begins with an introduction to the necessary mathematical background, including differential geometry and general relativity, which form the foundation of modern cosmology. We made an overview of different statistical methods, including statistical estimators, a brief study of Bayesian statistics, and hypothesis testing. This work also discusses the underlying physics that governs SNe Ia, and why they are characterized as standardizable candles. Finally, the statistical concepts are applied to the estimation of cosmological parameters using a Likelihood Ratio Test (LRT), which is implemented in a `Python` script. SNe Ia data obtained from the Supernova Legacy Survey (SNLS) are used to fit the Maximum Likelihood Estimators (MLE) and the confidence regions.

Key-words: 1. Cosmology. 2. Statistics. 3. Supernovae.

Contents

1 Introduction

Modern cosmology utilizes a range of tools to examine and validate proposed hypotheses. The combination of theories, simulations, precise observations, and statistical analyses has proven highly effective in addressing intricate inquiries about our universe (DODELSON, 2003). These findings have unveiled profound insights into a cosmos that greatly differs from the one envisioned by philosophers throughout history (NUSSBAUMER, 2014).

All recent discoveries were set in motion at the beginning of the last century with the development of Einstein's General Theory of Relativity, which dethroned space-time of its absolute and impetuous position. Differential geometry stands as the mathematical core of this new theory, where physical entities are promoted to tensors in a differentiable manifold (EINSTEIN; MINKOWSKI, 1920; CARROLL, 2003). Given all these changes, the mere attempt to fit the cosmos into preconceived notions of eternity proved incompatible with its newfound dynamic structure (NUSSBAUMER, 2014). Within a decade of the proposal of the general theory, dynamical cosmological models emerged and were subsequently confirmed by observations. Not only the cosmos was proved to be dynamic, but it was constantly expanding (HUBBLE, 1926; NUSSBAUMER, 2014).

Today's observations are made by huge collaborations, and provide the community with more data than we could ever imagine (ASTIER et al., 2005; CONTRERAS et al., 2010; AGHANIM et al., 2020). With large data sets, we are now capable of applying multiple statistical techniques to fit the parameters of our cosmological models. Consequently, a solid understanding of statistics is crucial for conducting precise analyses of such extensive samples. In particular, knowledge of estimators and hypothesis testing methods becomes essential to rigorously evaluate and validate our models (BARLOW, 1993; RICE, 2003).

One type of observation that proved extremely useful to further understand the expansion of the cosmos was of Type Ia Supernovae (PERLMUTTER et al., 2003; HUSS; MOBERG, 2011). The measurement of the light curve of such phenomenons made possible their classification as standardized candles (WRIGHT; LI, 2018). By evaluating the likelihood of these samples we can estimate a set of parameters of our cosmological model at different confidence levels (ASTIER et al., 2005).

Our objective is to replicate these statistical tests using a sample of Type Ia Supernovae, employing a Likelihood Ratio Test implemented in `Python`. The aim is to determine the confidence levels for the cosmological parameters and assess the effectiveness of our `Python` script as a suitable method for achieving this.

To accomplish this goal, we commence by delving into the understanding of this dynamic universe in Chapter 2. In Section 2.1, we discuss the required mathematical tools,

followed by the construction of Einstein's General Theory in Section 2.2, and its application to the cosmos in Section 2.3. The statistical foundation is elucidated in Chapter 3, where we cover estimators in Section 3.1. Additionally, a brief remark on Bayesian Statistics is provided in Section 3.2. Subsequently, in Section 3.3, we describe the Likelihood Ratio Test. The data set description is found in Chapter 4, incorporating a discussion on the physics underlying Type Ia Supernovae in Section 4.1, and concluding with the application of our script to the data sample in Section 4.2.

2 The dynamical universe

In the second half of the XIX century, physicists were avidly searching for proof of the existence of the *luminiferous ether*. The *ether* was theorized as a required propagation medium and privileged inertial frame for electromagnetic waves. The observation of this entity could bring the unification of magnetism, electricity, and light to a close. One possible method of measurement of the effect of such an absolute entity – which was the approach of the Michelson-Morley experiment – was the velocity of light relative to its flow. This experiment "undermined the whole structure of the old ether theory and thus served to introduce the new theory of relativity" (MINKOWSKI; EINSTEIN; MAHALANOBIS, 2021).

Einstein's acumen lay in the physical interpretation of the result of the Michelson-Morley experiment. Physical laws should be independent of the absolute velocity of an observer. The validity of the Maxwell equations, especially the velocity of light in a vacuum, and the relativity principle are foundations for the Special Theory of Relativity. This first theory is restricted to a specific set of reference frames. The Lorentz Transformations are particular to inertial frames, particles in rectilinear uniform motion (EINSTEIN, 1905; LORENTZ, 1952).

This restrictive characteristic gives the impression that the Galilean coordinates are a fundamental referential, even though there is no reason why to believe it is so. It is easy to see that from a Non-Lorentzian change of coordinates geometrical forces could arise. But this same phenomenon can be used to "transform away" gravitational forces. The Principle of Equivalence states that a gravitational field of force is strictly equivalent to one introduced by a transformation of coordinates and no possible experiment can distinguish between the two (EINSTEIN; MINKOWSKI, 1920).

Einstein's second theory is a generalization to all these possible reference frames, even though we may need to look into complex geometries and non-euclidean coordinate systems. As Einstein says in his 1920 work "The mathematical apparatus useful for the general relativity theory, lay already complete in the Absolute Differential Calculus" (EINSTEIN; MINKOWSKI, 1920).

2.1 Differential Geometry

Our minimal mathematical representation of spacetime is a given set of spacetime points, $\mathcal{M}$, which locally resembles Euclidean space. Furthermore, this space can be parametrized by a set of real numbers, with which we can describe the dynamical evolution of a system using differential equations. The set of spacetime points with these two

characteristics is called a differentiable manifold (ISHAM, 1999).

In a single point of the manifold, we can describe two vector spaces. First, the tangent space of $\mathcal{M}$ at the point p, $\mathcal{T}_p\mathcal{M}$, that is a set of vectors

$$V = V^\mu \partial_\mu, \tag{2.1}$$

where V^μ is the component and ∂_μ is the basis. These objects take a function defined on $\mathcal{M}$ to a real number on $\mathcal{R}$. The vector components transform according to the following law:

$$\widetilde{V}^\nu = \frac{\partial \widetilde{x}^\nu}{\partial x^\mu} V_\mu. \tag{2.2}$$

The other vector space is the cotangent space of $\mathcal{M}$ at the point p, $\mathcal{T}_p^*\mathcal{M}$. This space consists of a set of forms

$$w = w_\mu dx^\mu, \tag{2.3}$$

which in turn take a vector from $\mathcal{T}_p\mathcal{M}$ to a real number in $\mathcal{R}$. Forms components transform accordingly:

$$\widetilde{w}_\nu = \frac{\partial x_\mu}{\partial \widetilde{x}^\nu} w_\mu. \tag{2.4}$$

Knowing a set of vectors $(V^{\mu_1}, V^{\mu_2}, \ldots, V^{\mu_n})$ and a set of form $(w_{\nu_1}, w_{\nu_2}, \ldots, w_{\nu_m})$, their product will transform as:

$$\begin{aligned}\widetilde{V}^{\alpha_1}\widetilde{V}^{\alpha_2}\ldots\widetilde{V}^{\alpha_n}\widetilde{w}_{\beta_1}\widetilde{w}_{\beta_2}\ldots\widetilde{w}_{\beta_m} = \frac{\partial \widetilde{x}^{\alpha_1}}{\partial x^{\mu_1}}\frac{\partial \widetilde{x}^{\alpha_2}}{\partial x^{\mu_2}}\ldots\frac{\partial \widetilde{x}^{\alpha_n}}{\partial x^{\mu_n}}\frac{\partial x_{\nu_1}}{\partial \widetilde{x}^{\beta_1}}\frac{\partial x_{\nu_2}}{\partial \widetilde{x}^{\beta_2}}\ldots \\ \frac{\partial x_{\nu_m}}{\partial \widetilde{x}^{\beta_m}} V^{\mu_1}V^{\mu_2}\ldots V^{\mu_n} w_{\nu_1} w_{\nu_2}\ldots w_{\nu_m}.\end{aligned} \tag{2.5}$$

Objects that respect the transformation law above are called tensors and can be thought of as the product of vectors and forms;

$$A^{\mu_1\mu_2\ldots\mu_n}{}_{\nu_1\nu_2\ldots\nu_m} = V^{\mu_1}V^{\mu_2}\ldots V^{\mu_n} w_{\nu_1} w_{\nu_2}\ldots w_{\nu_m}. \tag{2.6}$$

We say the order of a tensor is (n, m) when it is constructed from n vectors and m forms.

Moreover, following the same line of thought, one can create a different ranked tensor from the product of two other tensors,

$$T^{\mu\nu}{}_{\sigma\lambda} = A^{\mu\nu} B_{\sigma\lambda}. \tag{2.7}$$

Notice how we have combined two lower-ranked tensors, $(2, 0)$ and $(0, 2)$, to create a third with order $(2 + 0, 0 + 2)$.

Another possible operation is the rank reduction of a tensor. This can be done by simply contracting two indexes of a tensor.

$$T^\mu{}_\nu = T^{\mu\sigma}{}_{\sigma\nu}. \tag{2.8}$$

The contraction between two indexes transformed a $(2,2)$ tensor into a $(1,1)$ tensor.

An important feature of tensors is their symmetries. We say a tensor is symmetric if

$$A_{\mu\nu} = A_{\nu\mu} \tag{2.9}$$

and antisymmetric if

$$A_{\mu\nu} = -A_{\nu\mu}. \tag{2.10}$$

It is important to note that the dynamics of an entity are defined at a specific point in our differentiable manifold. For instance, comparing vectors from two distinct points is, so far, a fool's errand. They exist in entirely different mathematical spaces and are, therefore, incomparable. We will soon explore how one might connect such particular spaces.

We have hitherto described arbitrary tensors without any specific usage for the understanding of the general theory. First and foremost, we shall do a brief study of the metric tensor, or as Einstein called it, the covariant fundamental tensor.

The metric tensor is a bilinear map from vectors on the tangent space to a real number. The line element is the distance between two points in a given space and is defined as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{2.11}$$

The metric is a symmetric tensor. Moreover, it is an orthogonal tensor, *i.e.*, $g_{\mu\sigma} g^{\sigma\nu} = \delta_\mu{}^\nu$. For this reason, it is conventional to use the metric to raise and lower indexes;

$$g_{\mu\nu} A^\nu = A_\mu. \tag{2.12}$$

Let us go back to the remark made about the connectivity of points on a manifold. We aim to create an entirely covariant theory; therefore, we expect that our mathematical apparatus should be invariant under the tensor transformation law. It is easy to see that the partial derivative – a primary candidate for our connectivity problem – does not transform accordingly Eq.(2.5). Hence, we shall define a covariant derivative by adding the following correction to the partial derivative:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\sigma} V^\sigma. \tag{2.13}$$

This factor is called connection. Furthermore, we impose that

$$\widetilde{\nabla}_\mu \widetilde{V}^\nu = \frac{\partial x^\alpha}{\partial \widetilde{x}^\mu} \frac{\partial \widetilde{x}^\nu}{\partial x^\beta} \nabla_\alpha V^\beta, \tag{2.14}$$

which makes the connection transform in such a way that it "cancels out" the irregularities in the partial derivative transformation. The covariant derivative of a form is

$$\nabla_\mu w_\nu = \partial_\mu w_\nu - \Gamma^\sigma_{\mu\nu} w_\sigma \tag{2.15}$$

and its effect on a scalar is simply the partial derivative.

Finally, we also expect the connection to satisfy two other characteristics. First, it should be torsion-free, *i.e.*, symmetric in its two lower indexes. Secondly, it should be metric-compatible, meaning the covariant derivative of the metric is null. Following all these rules, we can uniquely define a connection. The Christoffel Symbol,

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left[\frac{\partial g_{\alpha\nu}}{\partial x^{\beta}} + \frac{\partial g_{\beta\nu}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\nu}} \right], \tag{2.16}$$

is the only possible torsion-free, metric-compatible connection, which is the base for Einstein's general theory.

We can define a directional covariant derivative

$$\frac{D}{d\lambda} V^{\nu} = \frac{dx^{\mu}}{d\lambda} \nabla_{\mu} V^{\nu}, \tag{2.17}$$

over which we bring attention to a special case. When the above derivative is null we say that the vector V^{ν} is being parallel transported along the path $dx^{\mu}/d\lambda$. Such requirements give rise to

$$\frac{d}{d\lambda} V^{\nu} + \Gamma^{\nu}_{\alpha\beta} \frac{dx^{\alpha}}{d\lambda} V^{\beta} = 0, \tag{2.18}$$

known as the equation of parallel transport.

With the knowledge we gathered so far, we are now able to derive the geodesic equation of a timelike path. From Eq. (2.11), we shall take ds as the proper time $d\tau$ and parametrize the equation to a real value λ. Furthermore, we can integrate the expression to find τ leading to

$$\tau = \int \sqrt{-g_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda}} \, d\lambda, \tag{2.19}$$

where the minus sign has been added to turn the interval positive. For simplicity, we rename the term inside the square root as f. Additionally, we take an infinitesimal functional variation of the equation, such that,

$$\delta\tau = \int \delta\sqrt{-f} d\lambda = \int \frac{1}{2} (-f)^{-\frac{1}{2}} \delta(-f) d\lambda. \tag{2.20}$$

Using the identity in terms of the four-velocity $U^{\mu} = dx^{\mu}/d\tau$,

$$g_{\mu\nu} \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau} = g_{\mu\nu} U^{\mu} U^{\nu} = -1, \tag{2.21}$$

and parametrizing λ to τ we can rewrite Eq. (2.20) as

$$\delta\tau = -\frac{1}{2} \int \delta f d\tau. \tag{2.22}$$

By adding small variations to the metric and the path,

$$x^\mu \to x^\mu + \delta x^\mu, \tag{2.23}$$

$$g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}, \tag{2.24}$$

we can solve to the first order and analyze the integrand at the boundary using the Leibniz Rule. The resultant integrand should be null by imposing no variations of the proper time.

Finally, the following is true:

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0. \tag{2.25}$$

The equation above represents the path that extremizes the proper time and is called the geodesic equation. Once we are set on a particular metric we can use this equation to find the geodesic paths. It is essential to point out that any parameter linearly related to τ will leave Eq. (2.25) unchanged. These are called affine parameters.

Furthermore, by identifying $dx^\mu/d\tau = V^\mu$, we can see that Eq. (2.25) becomes Eq. (2.18);

$$\frac{D}{d\tau}V^\mu = \frac{dV^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}V^\alpha V^\beta = 0. \tag{2.26}$$

This means that the geodesic is a path in which a vector is parallel transported along itself, *i.e.*, a straight line is a path that parallel transport its tangent vector.

Using the symmetry of $V^\alpha V^\beta$ one can rewrite Eq. (2.18),

$$\frac{dV_\mu}{d\tau} = \frac{1}{2}\partial_\mu g_{\alpha\beta}V^\alpha V^\beta. \tag{2.27}$$

If the metric is independent of some particular coordinate x^μ, then the form V_μ is invariant. In other words, there is an isometry of the metric of which V^μ is a conserved quantity. This construction is highly dependent on coordinates, whereas the following is true in a covariant manner:

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0. \tag{2.28}$$

The Eq. (2.28) is called Killing Equation, and $\vec{K}$ is the Killing Vector. Carrying the metric along the field of vectors $\vec{K}$ leaves it unchanged. Each Killing Vector Field in an n-dimensional manifold represents a unique isometry.

The word curvature is expected to come up when reading about general relativity and differential geometry. In our case, we shall study the intrinsic curvature, a quantity that can be measured within the geometry. This intrinsicality does not require our manifold to be embedded in a higher-dimensional space.

But how does one measure such a thing? The answer lies in the effect curvature has on vectors' cyclic transportation along our geometry. Let us take the commutator of the covariant derivative over a test vector

$$[\nabla_\mu, \nabla_\nu]\, V^\sigma = \nabla_\mu\nabla_\nu V^\sigma - \nabla_\nu\nabla_\mu V^\sigma. \tag{2.29}$$

Using Eq. (2.13) and some algebraic tricks, we end up with:

$$[\nabla_\mu, \nabla_\nu] V^\sigma = \left(\partial_\mu \Gamma^\sigma_{\nu\gamma} - \partial_\nu \Gamma^\sigma_{\mu\gamma} + \Gamma^\sigma_{\mu\lambda}\Gamma^\lambda_{\nu\gamma} + \Gamma^\sigma_{\nu\lambda}\Gamma^\lambda_{\mu\gamma}\right) V^\gamma, \tag{2.30}$$

the expression inside the parenthesis is the transformation acting on the vector after the cycle. The Riemann Tensor is defined as

$$R^\sigma{}_{\gamma\mu\nu} = \partial_\mu \Gamma^\sigma_{\nu\gamma} - \partial_\nu \Gamma^\sigma_{\mu\gamma} + \Gamma^\sigma_{\mu\lambda}\Gamma^\lambda_{\nu\gamma} + \Gamma^\sigma_{\nu\lambda}\Gamma^\lambda_{\mu\gamma}, \tag{2.31}$$

and it is used to measure the intrinsic curvature of a manifold. Reflecting on the covariant nature of the equations above it is easy to see that if the curvature vanishes for a particular coordinate system, it will do so in every other. The curvature tensor is also commonly represented with a lowered first index, $R_{\sigma\gamma\mu\nu}$.

One of the main properties of the curvature tensor is called the Bianchi Identity. We can find it by performing cyclic permutations of the covariant derivative of Eq. (2.31):

$$\nabla_\lambda R^\sigma{}_{\gamma\mu\nu} + \nabla_\sigma R^\gamma{}_{\lambda\mu\nu} + \nabla_\gamma R^\lambda{}_{\sigma\mu\nu} = 0. \tag{2.32}$$

There are two important contractions of Eq. (2.31). The first is called the Ricci tensor. We find it by contracting the first and third indexes of the curvature tensor;

$$R_{\mu\nu} = R^\sigma{}_{\mu\sigma\nu}. \tag{2.33}$$

The second is called the Ricci scalar. It is the contraction of the indexes in the Ricci Tensor:

$$R = R^\mu{}_\mu. \tag{2.34}$$

To create a symmetric $(0, 2)$ tensor from the Ricci Tensor and the Ricci Scalar, Eq. (2.32) can be contracted twice, resulting in

$$\nabla^\mu R_{\mu\nu} - \frac{1}{2}\nabla^\mu R g_{\mu\nu} = 0. \tag{2.35}$$

Applying the linearity of the covariant derivative, we define the Einstein Tensor as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}, \tag{2.36}$$

which sustain the following property:

$$\nabla^\mu G_{\mu\nu} = 0. \tag{2.37}$$

2.2 General Relativity

Now that we paved the way with the necessary mathematical introduction, we can proceed to generalize the gravitational theory. First and foremost, let us go about how

one might take a law of physics usable in flat space to curved spacetime. A sort of straightforward way is called the minimal-coupling principle. The minimal-coupling principle is simply an algorithm for rewriting a given law in a covariant manner. Such transformation can be made by transforming variables into tensors, and the partial derivative into covariant derivatives (CARROLL, 2003).

As previously hinted, gravity experienced by a free-falling particle should be merely an expression of the selected coordinate system; gravity is no different from a geometrical force. Hence, it seems reasonable to replace the gravitational field with a geometrical entity, *e.g.*, the metric tensor or the curvature tensor.

Moreover, another important entity in gravitation is the mass. Since we know that mass is a specific representation of energy, we should promote it to a tensor that encompasses all its possible forms. We want a $(0, 2)$ symmetric tensor that contains information about the distribution of energy in a system. Additionally, this tensor should extend the continuity equation to our curved spacetime.

The described tensor is known; it is the energy-momentum tensor $T_{\mu\nu}$. It can be thought of as a matrix of all combinations of the flux of μ momentums across surfaces of constant x^ν. The tensor $T_{\mu\nu}$ is constructed in a way that

$$\nabla^\mu T_{\mu\nu} = 0, \tag{2.38}$$

i.e., the total energy is conserved. The Eq. (2.38) is the covariant form of the continuity equation.

The gravitational field is described by the Poisson Equation:

$$\nabla^2 \Phi = 4\pi G\rho. \tag{2.39}$$

By following the minimal-coupling principle we can propose a proportionality between the second-order derivative of a $(0, 2)$ symmetric geometry-related tensor and the energy-momentum tensor.

Unfortunately, since the (2.16) is metric-compatible, direct derivations of the metric are not a great candidate. However, the Ricci Tensor does embody this twice-differentiated metric in itself. Sadly, $R_{\mu\nu}$ is not consistent with (2.38).

The required tensor was already derivated in the last section by twice contracting (2.32). The Einstein Tensor respects all requirements for the equivalence between energy and space-time curvature. Finally, we may propose the following field equation:

$$G_{\mu\nu} = \kappa T_{\mu\nu}. \tag{2.40}$$

We can find κ by solving the equation for a massive body at rest in a weak gravitational field. The energy-momentum tensor will be just the energy density across time, *i.e.*, the classical energy density $\rho = T_{00}$. The metric will be the Minkowski flat space plus small

perturbations,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{2.41}$$

known as the weak-field limit.

A few algebra yields

$$\nabla^2 h_{00} = -\kappa\rho, \tag{2.42}$$

where the similarity Eq. (2.39) is evident. Moreover, using the weak-field limit for non-relativistic particles, the geodesic equation shows us

$$h_{00} = -2\Phi. \tag{2.43}$$

Combining the last two equations, we show that

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{2.44}$$

is the generalization of Eq. (2.39). The Eq. (2.44) is known as Einstein's Field Equation. It brings to light the subtle reality that energy is the curvature of spacetime.

2.3 Cosmology

In 1917, Einstein applied his theory to the cosmos by proposing the existence of a cosmological constant, Λ, to balance the universe and keep it static. This idea was based on the prevailing belief at the time that the universe was static and unchanging. However, in the 1920s, Alexander Friedmann disregarded the need for a closed universe, instead, he theorized the possibility of a dynamical cosmos with a variable radius. He used Eq. (2.44) to show that the universe could be either expanding or contracting, which challenged Einstein's idea of a static universe (NUSSBAUMER, 2014).

To find such solutions, we begin with the cosmological principle. The cosmological principle states that the universe is spatially homogeneous and isotropic. These two characteristics give rise to rotational and translational symmetries. Constraining our possible metrics to those with such symmetries we are forced to select from a set of maximally symmetric spaces (DODELSON, 2003).

The set of maximally symmetric spaces contains those spaces that have the same number of Killing vector fields as the Euclidian Space, but with arbitrary constant curvature. This arbitrarity allows three groups: zero, positive, and negative curvature. The only difference within these groups is the scaling, therefore, we will only consider $(-1, 01)$ as possible values (MUKHANOV, 2005).

Since the curvature is constant in such spaces the curvature tensor, (2.31), is a Lorentz-invariant in a locally flat coordinate system $x^{\hat{\mu}}$. Therefore, we can build $R_{\hat{\sigma}\hat{\gamma}\hat{\mu}\hat{\nu}}$ from

another known invariant tensor, *e.g.*, the metric. By trying to match the symmetries of (2.31), we arrive at the following proportionality:

$$R_{\hat{\sigma}\hat{\gamma}\hat{\mu}\hat{\nu}} = A\left(g_{\hat{\sigma}\hat{\mu}}g_{\hat{\gamma}\hat{\nu}} - g_{\hat{\sigma}\hat{\nu}}g_{\hat{\gamma}\hat{\mu}}\right). \tag{2.45}$$

The same is true for every other coordinate system. Moreover, we can contract booth sides twice yielding the general form of the Riemann Curvature Tensor for a maximally symmetric manifold;

$$R_{\sigma\gamma\mu\nu} = \frac{R}{n(n-1)}\left(g_{\sigma\mu}g_{\gamma\nu} - g_{\sigma\nu}g_{\gamma\mu}\right). \tag{2.46}$$

We are interested in $3D$ spaces, which will compose the leaves of constant time. By restricting our curvature to those of three dimensions, we define the Gaussian curvature as

$$K = \frac{R}{6}. \tag{2.47}$$

Let us begin by assuming a general spatially-dynamic metric

$$ds^2 = -dt^2 + a(t)d\sigma^2, \tag{2.48}$$

where $a(t)$ is an arbitrary function of t and $d\sigma^2$ is the metric of a spherically-symmetric 3-space given by

$$d\sigma^2 = e^{2\beta(r)}dr^2 + r^2 d\Omega^2. \tag{2.49}$$

Here $\beta(r)$ is an arbitrary function of r and $d\Omega^2$ is the solid angle. By computing the curvature of this general metric and using the curvature tensor for maximally symmetric spaces we can fix

$$e^{2\beta(r)} = \frac{1}{1 - Kr^2}. \tag{2.50}$$

Rewriting everything and parametrizing the radius, $r = \sqrt{\alpha}r$, to restrain $K = \alpha k$ we arrive at the Friedmann-Lemaître-Robertson-Walker Metric (FLRW Metric)

$$ds^2 = -dt^2 + a(t)\left(\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2\right). \tag{2.51}$$

We have $a(t)$ as the scale factor and $k = -1, 0, +1$ as the curvature.

Before solving Einstein's field equation, it is necessary to describe the energy within the universe. The fundamental assumption is that the cosmos is permeated by a perfect fluid composed of various entities. The following energy-momentum tensor can describe such fluid:

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \tag{2.52}$$

The structure of Eq. (2.52) can be understood by the isotropy of the perfect fluid. Isotropy results in the diagonal form of this tensor. Moreover, it also explains the equal valued pressure in all three directions.

With cosmo's energy content and its related metric, we can solve Eq. (2.44) for the variable radius $a(t)$. Although the process is straightforward, it is long and mechanical. First, we find all Christoffel symbols for the FLRW metric. Secondly, with all possible symbols, we compute the Riemann tensor, contract to find the Ricci tensor, and contract again for the Ricci scalar. Now, we are capable of analyzing the time and spatial portions of the Field Equation separately. By doing so, we arrive at two equations:

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3}\rho(t) - \frac{k}{a(t)^2} \tag{2.53}$$

and

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho(t) + 3p\right). \tag{2.54}$$

These are respectively the First and Second Friedmann Equations. The Eq. (2.53) is derived from the time portion of the Field Equation and the Eq. (2.54) is derived from a substitution of Eq. (2.53) into the spatial portion of the Field Equation.

Not long after Friedmann's remarks on Relativistic Cosmology, the Belgian priest Georges Lemaître arrived at similar equations and theorized about the cosmos at $t = 0$. In 1927 he proposed that the universe began as a "primordial atom", which exploded creating all energy components of the universe. Moreover, his calculations showed a linear relationship between the velocity of stellar bodies and their distances from an observer. In 1929, Hubble published his measurement of the relation between the redshift of galaxies and their distances, solidifying the idea of a cosmological expansion. The Hubble-Lemaître law states

$$v = H_0 d, \tag{2.55}$$

where H_0 is the current Hubble constant, v is the relative velocity, and d is the distance. By replacing the velocity of a stellar body with the velocity of the expansion and the distance by the radius, we can also write the Hubble parameter as

$$H = \frac{\dot{a}}{a}. \tag{2.56}$$

The Hubble parameter, H, represents the rate at which the universe is expanding, and the current rate, H_0, is one of the parameters that can be determined using observational data.

It is useful to define a critical energy density required for the universe to have zero curvature or $k = 0$. By constraining the universe in this way, we can solve Eq. (2.53) for

ρ_{crit};

$$\rho_{crit} = \frac{3H^2}{8\pi G}. \tag{2.57}$$

A general density parameter in terms of ρ_{crit} is given by

$$\Omega = \frac{\rho}{\rho_{crit}}. \tag{2.58}$$

Before we proceed with the analysis of these parameters, it is necessary to understand the distribution of energy density in the universe and how each component evolves with the expansion of the universe. For any perfect fluid (2.52) is conserved, *i.e.*, $\nabla_\mu T^\mu{}_\nu = 0$. The time component of this derivation yields

$$\nabla_\mu T^\mu{}_0 = \partial_0 T^0{}_0 + \Gamma^\mu_{\mu 0} T^0{}_0 - \Gamma^\alpha_{\mu 0} T^\mu{}_0 = 0. \tag{2.59}$$

Computing $T^\mu{}_\nu$ and replacing the symbols obtained from the FLRW metric, we find the fluid equation;

$$\frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}(\rho + p) = 0. \tag{2.60}$$

Each component of the cosmo has a unique equation of state

$$p = w\rho. \tag{2.61}$$

Replacing Eq. (2.61) in Eq. (2.60), and solving for ρ in terms of a, we find

$$\rho = \rho_0 a^{-3(1+w)}. \tag{2.62}$$

The visible components of the universe are matter and radiation. Since the equation of state for each is known, their evolutions are respectively

$$\rho_m = \rho_{m0} a^{-3} \tag{2.63}$$

and

$$\rho_r = \rho_{r0} a^{-4}. \tag{2.64}$$

Another type of energy may also arise as a possible interpretation of the cosmological constant Λ. The most general form of the Field Equation has Λ as an integration constant. One can think of this constant as a part of the energy-momentum tensor that emerges from the vacuum. The central principle regarding the vacuum energy density is how it should evolve with the variation of the scale factor. As the size of the cosmos increases or decreases, the amount of vacuum energy changes accordingly. Therefore, its density should remain constant across the evolution of the universe. For $w = -1$, we can see in Eq. (2.62) a constant vacuum energy density;

$$\rho_\Lambda = \rho_{\Lambda 0}. \tag{2.65}$$

Moreover, one vital corollary of $w = -1$ is

$$p = -\rho_\Lambda. \tag{2.66}$$

Finally, the energy density for the whole universe can be written as

$$\rho(t) = \rho_m(t) + \rho_r(t) + \rho_\Lambda(t) = \rho_{m0}a^{-3} + \rho_{r0}a^{-4} + \rho_\Lambda. \tag{2.67}$$

Dividing by the critical density, we have the density parameter

$$\Omega = \frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} + \Omega_\Lambda, \tag{2.68}$$

where we have omitted the subscript indicating that these parameters represent their current values for simplicity. Just like H_0, the set of parameters $(\Omega_m, \Omega_r, \Omega_\Lambda)$ can be nummerically fixed with observational data.

We are, at last, close to an equation that couples all of our cosmological parameters of interest. We can define a final parameter by rewriting the first Friedmann Equation in terms of the Hubble and density parameters. Dividing Eq. (2.53) by H_0^2, identifying Eq. (2.57), and defining

$$\Omega_k = -\frac{k}{H_0^2}, \tag{2.69}$$

we can solve the result for H and find

$$H^2 = H_0^2\left(\frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} + \frac{\Omega_k}{a^2} + \Omega_\Lambda\right). \tag{2.70}$$

The Eq. (2.70) lacks an observable variable. However, using Eq. (2.25), one can find a relation between the scale factor $a(t)$ and the visible wavelength (DODELSON, 2003). To find this equality, we first define the four-momentum of a particle

$$P^\mu = \frac{dx^\mu}{d\lambda} = (E, p_r, p_\theta, p_\phi). \tag{2.71}$$

From the relativistic energy-momentum relation a massless particle, such as a photon, will have a four-momentum with zero magnitudes. Furthermore, a photon travels radially across the universe, therefore $p_\theta = p_\phi = 0$. Computing $g_{\mu\nu}P^\mu P^\nu$ we find that

$$-E^2 + \frac{a^2}{1 - kr^2}p_r^2 = 0. \tag{2.72}$$

Replacing the Eq. (2.71) into Eq. (2.25) and using the chain rule, the time component of the result is

$$E\frac{dE}{dt} = -\Gamma^0_{rr}p_r^2 = -a\dot{a}\frac{1}{1 - kr^2}p_r^2. \tag{2.73}$$

Replacing p_r^2 from the energy-momentum relation into the equation above we arrive at a solvable relation between energy and the scale factor,

$$\frac{1}{E}\frac{dE}{dt} = -\frac{1}{a}\frac{da}{dt}. \tag{2.74}$$

Taking $a_0 = 1$, the Eq. (2.74) yields

$$E = \frac{E_0}{a}. \tag{2.75}$$

A commonly used observable is the redshift. It is defined as the ratio between the variation in wavelength, $\lambda_{obs} - \lambda_{em}$, and the emitted wavelength, λ_{em}, *i.e.*,

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}. \tag{2.76}$$

By rewriting Eq. (2.75) in terms of λ_{obs} and λ_{em}, an equation of $a(z)$ emerges;

$$a = \frac{1}{1+z}. \tag{2.77}$$

This relation is fundamental. We shall use it to know the scale factor at the moment of the emission of the light we are observing. In close, we can replace the scale factor in Eq. (2.70) by Eq. (2.77), yielding the Parameter Equation as a function of the redshift;

$$H^2 = H_0^2 \left(\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_k(1+z)^2 + \Omega_\Lambda\right). \tag{2.78}$$

The redshift can be used to estimate the distance of a light-emitting object. We can first create a grid over the cosmos and use it to map fixed distances between objects. This way, even though the grid is expanding, the measurement within the grid remains constant. We call this the comoving distance,

$$\chi(z) = \int_0^z \frac{dz}{H(z)}. \tag{2.79}$$

A value of redshift will always be at a fixed comoving distance $\chi(z)$.

We can also estimate the distance using the flux of light measured by an observer. This is the luminosity distance, d_L, and it can be calculated in terms of Eq. (2.79) with

$$d_L = (1+z)\frac{H_0^{-1}}{\sqrt{|\Omega_k|}} S_k\left[H_0\sqrt{|\Omega_k|}\chi(z)\right]. \tag{2.80}$$

$S_k(\chi)$ is one of three possible functions depending on the curvature;

$$S_k(\chi) = \begin{cases} sin(\chi), & k = 1 \\ \chi, & k = 0 \\ sinh\chi, & k = -1 \end{cases}. \tag{2.81}$$

Finally, we can use the magnitude of an object to estimate its distance. The distance modulus, μ, is the difference between the apparent magnitude and the absolute magnitude. It is related to Eq. (2.80) by

$$\mu = 5\log_{10}\left(\frac{d_L}{10}\right). \tag{2.82}$$

3 Statistical Methods

A great example of data analysis in the physical sciences is the work of Tycho Brahe, a Danish astronomer of the 16th century who made detailed observations of the stars and planets over many years and recorded his observations in meticulous notes (later published as the Rudolphine Tables by Johannes Kepler) (EGGEN, 2023; WESTMAN, 2023).

One of Tycho Brahe's most significant contributions was his extensive observations of the planet Mars. Using his observations, he was able to construct accurate tables of the planet's position in the sky. Johannes Kepler, one of Tycho's students, was fascinated by the retrograde motion present in the orbit of Mars. He spent eight years trying to solve this peculiar motion, with the firm belief that by studying the orbit of Mars, we would either arrive at the secrets of astronomy or forever remain in ignorance of them. (SHEEHAN, 1996).

With the knowledge he acquired, he was able to formulate his laws of planetary motion. Using Tycho's observations, he showed that the planets move in elliptical orbits around the sun, rather than in the circular orbits proposed by Ptolemy (RABIN, 1996).

Tycho Brahe's observations were a crucial step in the development of modern astronomy, and his work laid the foundation for the development of more accurate models of the solar system and gravity itself. The use of detailed observations and statistical analysis to understand the natural world was a key breakthrough in the physical sciences and continues to be an important tool in scientific research(EGGEN, 2023).

3.1 Estimators

The primary focus of data analysis is a set of data, observations, $\{x_i\}$, and an underlying relationship between these measurements. These assumptions are largely based on the belief that there is a deeper reality beyond the measurements themselves. Our goal is to extract meaningful information from these data (BARLOW, 1993).

Any operation that can be performed over $\{x_i\}$ to summarize its distribution in a single value is called statistics. These include statistics for central tendencies, such as the arithmetic mean and median, and for dispersion, such as variance and standard deviation.

However, one can also arrive at these statistics through a probability distribution P that theoretically describes the sample. The probability of a measurement x_i given a parameter μ is $P(x_i|\mu)$. For instance, from P one can find the expected value of x_i with:

$$\langle x_i \rangle = \sum_{x_i} x_i P(x_i|\mu). \tag{3.1}$$

The expected value of x and the mean value of the sample $\{x_i\}$ are related by the law of large numbers. Even though the expected value is found by operating on the probability function, and the mean value is found by operating on the real data, they are equal for a sufficiently large sample.

The process of directly finding a parameter from a sample is called estimation. An estimator is any rule, method, or criterion used to estimate the parameter[1]. The average value of a sample is commonly calculated as follows:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{3.2}$$

This expression is an estimator of the expected value, $\langle x_i \rangle = \mu$. To emphasize this relationship, it is common to write $\bar{x}$ as an estimate of the mean, denoted by $\hat{\mu}$.

One might use a different expression or method to find the same parameter. How can one determine if the estimator being used is the best option? As we shall see, an estimator can be consistent, unbiased, and efficient. These three characteristics can help us choose between possible estimators.

Given a certain sample $\{x_i\}$ related to a particular distribution $P(x_i)$ with a parameter μ, $\hat{\mu}$ is considered consistent if as the sample size increases, the estimate $\hat{\mu}(\{x_i\})$ converges to μ, i.e., $\lim_{N\to\infty} \hat{\mu} = \mu$. Furthermore, $\hat{\mu}$ will be unbiased when $\langle \hat{\mu} \rangle = \mu$. Finally, we say that $\hat{\mu}$ is efficient if its variance is sufficiently small.

Selecting a particular estimator can be a challenging task. The analysis of the efficiency of $\hat{\mu}$, for example, is not only a comparison between estimators but also a value that depends on the sample.

Consider a sample $\{x_i\}$, where all values are from the same distribution $P(x_i)$ and $\langle x_i \rangle = \mu$. The Eq. (3.2) is an estimator for the expected value. We can calculate the variance of the estimator, $V(\hat{\mu}) = \langle \hat{\mu}^2 \rangle - \langle \hat{\mu} \rangle^2$. By calculating $\langle \hat{\mu}^2 \rangle$ as

$$\langle \hat{\mu}^2 \rangle = \left\langle \left(\sum_i^N \frac{x_i}{N} \right)^2 \right\rangle = \frac{1}{N^2} \left\langle \sum_i^N x_i^2 \right\rangle + \frac{1}{N^2} \left\langle \sum_{i \neq j}^N x_i x_j \right\rangle, \tag{3.3}$$

we can simplify it further by using $\langle x_i^2 \rangle = V(x) + \mu^2$ and the fact that $\langle x_i x_j \rangle = \mu^2$ for two independent variables x_i and x_j. The sum of a constant over i is N, and over $i \neq j$ is $N^2 - N$. As the Eq. (3.2) is unbiased, the variance of the estimator will be

$$V(\hat{\mu}) = \frac{V(x)}{N}. \tag{3.4}$$

An effective way to determine a suitable estimator for a given parameter is to choose $\hat{\theta}$ that maximizes the probability of the sample $\{x_i\}$. Essentially, the probability of the

[1] Barlow R. (1989), pg. 69 for examples of different estimators.

sample of size N, given $P(x_i|\theta)$, is calculated as the cumulative probability of each measurement. The expression

$$\mathcal{L}(\{x_i\}|\theta) = \prod_i^N P(x_i|\theta) \tag{3.5}$$

is referred to as the likelihood of $\{x_i\}$ given a parameter θ. We define $\hat{\theta}$ as the maximum likelihood estimator (MLE) when

$$\frac{\partial \mathcal{L}}{\partial \theta}\Big|_{\hat{\theta}} = 0. \tag{3.6}$$

MLEs are generally consistent and, for large samples, they tend to be both unbiased and have the smallest variance. A key consideration of MLE is that they are simply a sensible approach to an estimator.

In order to prove that MLEs are the most efficient estimators, we shall show that estimators have a minimum variance bound. Say we have an unbiased estimator $\hat{\theta}(\{x_i\})$, its expected value is

$$\langle\hat{\theta}\rangle = \int \hat{\theta}(\{x_i\})\mathcal{L}d^n x = \theta, \tag{3.7}$$

whose expected values is simply a function of θ. Differentiating the above equation,

$$\int \hat{\theta}\frac{\partial \mathcal{L}}{\partial \theta}d^n x = \int \hat{\theta}\frac{\partial \ln \mathcal{L}}{\partial \theta}\mathcal{L}d^n x = 1 \tag{3.8}$$

and using the fact that the likelihood distribution is normalized, it is easy to show that

$$\int \frac{\partial \mathcal{L}}{\partial \theta}d^n x = \int \frac{\partial \ln \mathcal{L}}{\partial \theta}\mathcal{L}d^n x = 0. \tag{3.9}$$

Finally, by subtracting θ times the right-hand side of the Eq.(3.9) from the right-hand side of the Eq. (3.8), we find

$$\int d^n x \left(\hat{\theta} - \theta\right)\frac{\partial \ln \mathcal{L}}{\partial \theta}\mathcal{L} = 1. \tag{3.10}$$

Before we go any further, let us briefly define the Cauchy-Schwarz inequality. A possible generalization, when we define the norm of a function, is

$$\int f^* f d^n x = \int f^2 d^n x. \tag{3.11}$$

The Cauchy-Schwarz inequality is

$$\left(\int f^2 d^n x\right)\left(\int g^2 d^n x\right) \geq \left(\int fg d^n x\right)^2. \tag{3.12}$$

Identifying $f = \left(\hat{\theta} - \theta\right)\sqrt{\mathcal{L}}$, $g = \frac{\partial \ln \mathcal{L}}{\partial \theta}\sqrt{\mathcal{L}}$, and their product as Eq. (3.10), the inequality becomes:

$$\left(\int \left(\hat{\theta} - \theta\right)^2 \mathcal{L}d^n x\right)\left(\int \left(\frac{\partial \ln \mathcal{L}}{\partial \theta}\right)^2 \mathcal{L}d^n x\right) \geq 1. \tag{3.13}$$

The first integral is simply $\left\langle \left(\hat{\theta}-\theta\right)^2 \right\rangle = V(\hat{\theta})$, therefore,

$$V(\hat{\theta}) \geq \frac{1}{\int \left(\frac{\partial \ln \mathcal{L}}{\partial \theta}\right)^2 \mathcal{L} d^n x} = \frac{1}{\left\langle \left(\frac{\partial \ln \mathcal{L}}{\partial \theta}\right)^2 \right\rangle}. \tag{3.14}$$

The denominator on the right-hand side of the equation is called Information, $I(\theta)$. As we can see, the variance of an unbiased estimator has a minimal value which depends only on the likelihood and its parameters.

With Eq. (3.14), we can show that the variance of an MLE is precisely the minimum possible. First, the Eq. (3.6) is also true to $\ln \mathcal{L}$. With the definition using the log-likelihood, we may express it in a Taylor expansion around the real value of θ;

$$\left.\frac{\partial \ln \mathcal{L}}{\partial \theta}\right|_\theta + \left(\hat{\theta}-\theta\right) \left.\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_\theta = 0. \tag{3.15}$$

Solving for $(\hat{\theta}-\theta)^2$ yields

$$(\hat{\theta}-\theta)^2 = \frac{\left(\left.\frac{\partial \ln \mathcal{L}}{\partial \theta}\right|_\theta\right)^2}{\left(\left.\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_\theta\right)^2}. \tag{3.16}$$

Secondly, we can derivate Eq. (3.9) a second time and find the following equation:

$$\left\langle \left(\frac{\partial \ln \mathcal{L}}{\partial \theta}\right)^2 \right\rangle = -\left\langle \frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2} \right\rangle. \tag{3.17}$$

Finally, taking the expected value of Eq. (3.16), and using Eq. (3.17) we find the variance of an MLE to be equal to the minimum-variance bound;

$$V(\hat{\theta}) = \frac{1}{\left\langle \left(\frac{\partial \ln \mathcal{L}}{\partial \theta}\right)^2 \right\rangle}. \tag{3.18}$$

3.2 Bayesian Statistics

Suppose you are measuring a beam of mesons, composed of 90% pions and 10% kaons, with a Cherenkov detector. The detector only measures pions. The joint probability of a pion hitting the apparatus and it registering a signal in response to the pion is

$$\mathcal{P}(\pi \text{ and signal}) = \mathcal{P}(\pi|\text{signal})\mathcal{P}(\text{signal}). \tag{3.19}$$

$\mathcal{P}(\text{signal}|\pi)$ represents the probability of the apparatus registering a signal, given that a pion caused the signal. However, the same joint probability can be understood "backwards";

$$\mathcal{P}(\pi \text{ and signal}) = \mathcal{P}(\text{signal}|\pi)\mathcal{P}(\pi). \tag{3.20}$$

Comparing these interpretations yields,

$$\mathcal{P}(\pi|\text{signal}) = \frac{\mathcal{P}(\text{signal}|\pi)\mathcal{P}(\pi)}{\mathcal{P}(\text{signal})}. \tag{3.21}$$

Moreover, the overall probability of a signal caused either by a pion or a kaon is

$$\mathcal{P}(\text{signal}) = \mathcal{P}(\text{signal}|\pi)\mathcal{P}(\pi) + \mathcal{P}(\text{signal}|K)\mathcal{P}(K). \tag{3.22}$$

If our measurement device has an effectiveness of 95% and has a 6% chance of giving a signal for a kaon, the probability of a pion causing a signal from our detector is 99.3%. The value found is completely derived from the frequencies of measurements and their possible relationships.

The Eq. (3.21) can be used for any two events (A and B) and

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A)\mathcal{P}(A)}{\mathcal{P}(B)} \tag{3.23}$$

is Bayes' Theorem that we can use it for either Frequentist or Bayesian statistics. The latter is a branch of statistical inference that uses Bayes' Theorem to update our beliefs about the parameters of a model based on data. In other words,

$$\mathcal{P}(\text{parameter}|\text{sample}) = \frac{\mathcal{P}(\text{sample}|\text{parameter})\mathcal{P}(\text{sample})}{\mathcal{P}(\text{parameter})}. \tag{3.24}$$

The main difference between the two approaches is philosophical, rather than mathematical.

In Bayesian Statistics, probabilities are interpreted as degrees of belief, representing the uncertainty in our understanding of a problem. This means that probabilities are assigned to hypotheses or parameters, rather than just to the data. In contrast, Frequentist Statistics probabilities are only assigned to the data, not to the parameters themselves.

3.3 Hypothesis Testing

A hypothesis is a tentative explanation or prediction about a phenomenon or relationship between variables that can be tested through empirical research. A hypothesis is a statement or proposition that is formulated based on prior knowledge, theoretical background, or observations, and can be either supported or refuted by data.

For instance, the probability of a coin toss turning head is half. We assume that, after several tosses, half of the coin flips will result in heads and the other half will result in tails. This is our hypothesis based on prior knowledge. However, in reality, the exact value of half will not be obtained. For this reason, it is good to test our hypothesis to see how far we are from fact. Hypothesis testing is a procedure used to determine whether a particular claim or hypothesis is supported, or refuted, by the data.

We begin by defining our null and alternative hypotheses with their respective acceptance and rejection regions. Suppose we know the Gaussian distribution of a data set and its variance, but we need to determine its mean. One might speculate that $\mu = \mu_0$ (our null hypothesis H_0) or $\mu = \mu_1$ (the alternative H_A).

The null hypothesis is considered true if the measured value falls within a specific region of interest, denoted as $\mathcal{R}$. On the other hand, the alternative hypothesis is considered true if the measured value falls outside of this region; the complement of $\mathcal{R}$ is denoted as $\mathcal{D} - \mathcal{R}$. This complement is called the critical region $\mathcal{C}$. Based on where the measured value falls, we can decide whether to reject or fail to reject the null hypothesis.

To better understand hypothesis testing, we first define the critical value x_c, which is the boundary separating the region of acceptance, denoted as $\mathcal{R}$, from the critical region, which encompasses all values to the right of x_c.

The first question we ask in hypothesis testing is the probability of mistakenly rejecting the null hypothesis (H_0) when it is actually true and measuring a value x_i in the critical region. This is known as a Type-I error. This probability is denoted as α, and it represents the level of significance of the test. In other words, if the probability of making a Type-I error is α, we are willing to accept that in α percent of cases, our conclusion may be incorrect.

The second question is the probability of mistakenly accepting the null hypothesis when it is false and measuring a value x_i outside the critical region. This is known as a Type-II error, and its probability is denoted as β. Furthermore, we define $1 - \beta$ as the power of our test, which is the probability of successfully rejecting the alternative. The table below summarizes both types of errors.

	H_0 True	H_A True
Measurement in Critical Region	Type-I Error	Correct Decision
Measurement outside Critical Region	Correct Decision	Type-II Error

Hypothesis testing involves a trade-off between the probabilities of making Type-I and Type-II errors, which are influenced by several factors such as the chosen critical value, the level of significance, and the power of the test.

To optimize our testing, we aim to minimize both Type-I and Type-II errors, and this is achieved by choosing a critical region that leads to a high power and a low level of significance. Such a test provides the most decisive evidence for failing to reject the null hypothesis.

For simple hypotheses and alternatives the ratio of the power and the significance is

$$\frac{\int_{\mathcal{C}} P_A(x|\mu_A)dx}{\int_{\mathcal{C}} P_0(x|\mu_0)dx} = \frac{1-\beta}{\alpha}. \tag{3.25}$$

One can find the critical region that optimizes values of β for a selected α by maximizing

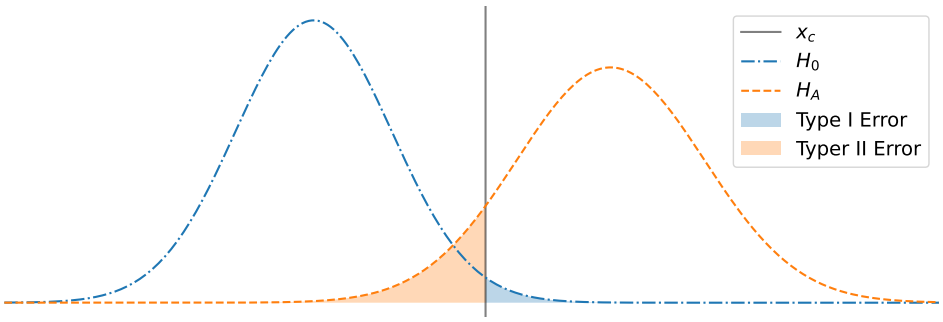

Figure 1 – Illustration of two distinct hypotheses, H_0 and H_A, represented in the figure. The central line represents the critical value x_c. The blue area represents the Type I Error, denoted by α, for measurements falling within the critical region when H_0 is true. On the other hand, the orange region represents β, indicating the probability of Type II Error for measurements falling outside the critical region when H_A is true.

Eq. (3.25) given a critical value c,

$$\frac{P_A(x|\mu_A)}{P_0(x|\mu_0)} > c \,\Big|\, x \in \mathcal{C}. \tag{3.26}$$

The Neyman-Pearson Lemma expresses this optimized ratio regarding the probabilities of the null and alternative hypotheses and states that a critical region defined by the inequality (3.26) generates the most powerful test for a specific significance level.

From the inequality (3.26) one can assess the goodness of a hypothesis from the likelihood of the measured sample. Moreover, the following generalization allows us to assess a composite hypothesis. Say for instance that the null hypothesis admits that our parameters $\vec{\theta}$ belong to a subset Θ_0 of the parameter space Θ, *i.e.*,

$$H_0 : \vec{\theta} \in \Theta_0 \subset \Theta. \tag{3.27}$$

On the contrary, the alternative is indifference, $\mathcal{I}$, to the restriction presupposes conclusion may be incorrect ed by H_0. The estimator of our null hypothesis is

$$\left.\frac{\partial \mathcal{L}}{\partial \theta}\right|_{\hat{\theta}_0 \in \Theta_0} = 0 \tag{3.28}$$

and the estimator of maximum likelihood without our null hypothesis is

$$\left.\frac{\partial \mathcal{L}}{\partial \theta}\right|_{\hat{\theta} \in \Theta} = 0. \tag{3.29}$$

From both estimators, we may rewrite the Neyman-Pearson Lemma as

$$\lambda = \frac{\mathcal{L}\Big(\{x_i\} \,\Big|\, \hat{\vec{\theta}}_0(\{x_i\}|H_0, \mathcal{I})\Big)}{\mathcal{L}\Big(\{x_i\} \,\Big|\, \hat{\vec{\theta}}(\{x_i\}|\mathcal{I})\Big)}, \tag{3.30}$$

where our estimator $\hat{\vec{\theta}}_0$ is a function of the data conditioned to our null hypothesis and every other possible hypothesis, $\mathcal{I}$. Whereas, $\hat{\vec{\theta}}$ is a function of our data solely supposing $\mathcal{I}$.

Since our estimators are of maximum likelihood, it is clear that the ratio above will always be less or equal to 1. The numerator is restricted to a particular subset of our space, while the denominator is free and always greater. This tests how maximal our null hypothesis is compared to an indifferent hypothesis.

For instance, if H_0 is not likely the value of λ will be small; as the likelihood of the sample given the MLE in the restricted subset of the parameter space is smaller than the one from the unrestricted parameter space. Whereas, $\lambda \approx 1$ indicates that H_0 is likely; as the likelihood of the sample given our hypothesis is close to the one with a global MLE.

Just as we did while testing a simple hypothesis, we shall define a critical region for rejection of H_0. This region is defined over the probability of a given $\lambda \leq \lambda_c$ being equal to a significance α, *i.e.*,

$$\mathcal{P}(\lambda \leq \lambda_c | H_0) = \alpha. \tag{3.31}$$

However, to find λ_c, one must know the probability distribution of λ.

Wilk's theorem describes the asymptotic behavior of the log-likelihood ratio. The theorem states that for large samples, $-2\ln\lambda$ approaches a chi-squared distribution with f degrees of freedom, where f equals the total number of parameters minus the number of free parameters in our null hypothesis.

By definition, if $x_i \sim N(0,1)$, the chi-squared distribution with one degree of freedom, χ_1^2, is $u_i = x_i^2 \sim N(0,1)^2$. Furthermore, given f independent u_i, their sum is called the chi-squared distribution with f degrees of freedom, χ_f^2 (RICE, 2003).

A heuristic proof for this theorem can be found with gaussian distribution with one restricted parameter (RICE, 2003). Say we have a sample of n measurements of a random variable $x_i \sim N(\mu, \sigma^2)$. Say $H_0 : \mu = \mu_0$, the ratio with $\mathcal{I}$, where the MLE of μ without restriction in the parameter space yields $\bar{x}$, is

$$\lambda = \exp\left(-\sum_i^n \frac{(x_i - \mu_0)^2}{2\sigma^2} + \sum_i^n \frac{(x_i - \bar{x})^2}{2\sigma^2}\right). \tag{3.32}$$

Being

$$\sum_i^n (x_i - \mu_0)^2 = n(\bar{x} - \mu_0)^2 + \sum_i^n (x_i - \bar{x})^2, \tag{3.33}$$

one can easily show that $-2\ln\lambda$ results in

$$-2\ln\lambda = \frac{n(\bar{x} - \mu_0)^2}{\sigma^2}. \tag{3.34}$$

Assuming that H_0 is true, the MLE $\bar{x}$ is consistent, meaning that as the sample size increases, it approaches the true population mean, μ_0. The variance of the estimator is

given by Eq. (3.4) as $V(\bar{x}) = \sigma^2/n$. The central limit theorem assures us that these are the expected value and variance of a normal distribution for $\bar{x}$. Therefore, we have $\bar{x} \sim N(\mu_0, \sigma^2/n)$. By making a simple variable change, $\bar{z} = \sqrt{n}(\bar{x} - \mu_0)/\sigma \sim N(0, 1)$.

Finally, we can see that the left-hand side of Eq. (3.34) is $\bar{z}^2$, which follows a chi-squared distribution with one degree of freedom, i.e., $\bar{z}^2 \sim \chi^2_1$. This test is known as the Likelihood Ratio Test (LRT). It is a method that belongs to the frequentist school of thought since it is not attributing probabilities to the parameters we are testing.

4 SNe Ia Parameter Estimation

In 1573, Tycho Brahe wrote that

> "Anno praecedente, mense Novembrj, die eiusdem undicesimo, vesperi post Solis occasum, cum meo more sidera coelo sereno contemplarer, novam quandam & inusitatam, praeque aliis admodum conspicuam, iuxta capitis verticem, animadverti fulgere Stellam." (BRAHE, 1573)

A star that was never seen before appeared in the night sky in the constellation Cassiopeia (Figure 2). At its peak, the star shined as bright as Venus. The star, which came to be known as "Nova Stella", was visible to the naked eye for over a year and is considered one of the most significant astronomical events in history. By 1574 the star faded away, leaving many astronomers to speculate about its nature and origin. Thirty years later, another nova was observed by Kepler in the constellation Ophiuchus, again disappearing after a couple of years.

In his 1866 work, On the spectrum of a new star in Corona Borealis, Sir Williams Huggins speculates a possible nature of these "temporary stars". By analyzing the spectroscopy of the object, Huggins said:

> "The character of the spectrum of this star, taken together with its sudden outburst in brilliancy and its rapid decline in brightness, suggests to us the rather bold speculation that, in consequence of some vast convulsion taking place in this object, large quantities of gas have been evolved from it, that the hydrogen present is burning by combination with some other element and furnishes the light represented by the bright lines, also that the flaming gas has heated to vivid incandescence the solid matter of the photosphere. As the hydrogen becomes exhausted, all the phenomena diminish in intensity, and the star rapidly wanes." (HUGGINS, 1899)

In other words, the "new star" could be an outburst of energy from a preexisting star.

Calculations of how much energy was required by a star to reach such levels of luminosity were made by Baade and Zwicky. They figured that the "energy of trapped radiation alone is comparable with the energy of total annihilation of the star [...]"(BAADE; ZWICKY, 1934). Zwicky coined the term "Super Nova" to distinguish this phenomenon from Common Nova, in which a star ignites material on its surface until it is exhausted.

The analysis of the radiation emitted by Supernovae (SNe) is key to the first spectral classification made by Rudolph Minkowski in 1941. Minkowski's taxonomy began with two distinct types: SNe I without hydrogen lines in the spectroscopy and SNe II with hydrogen

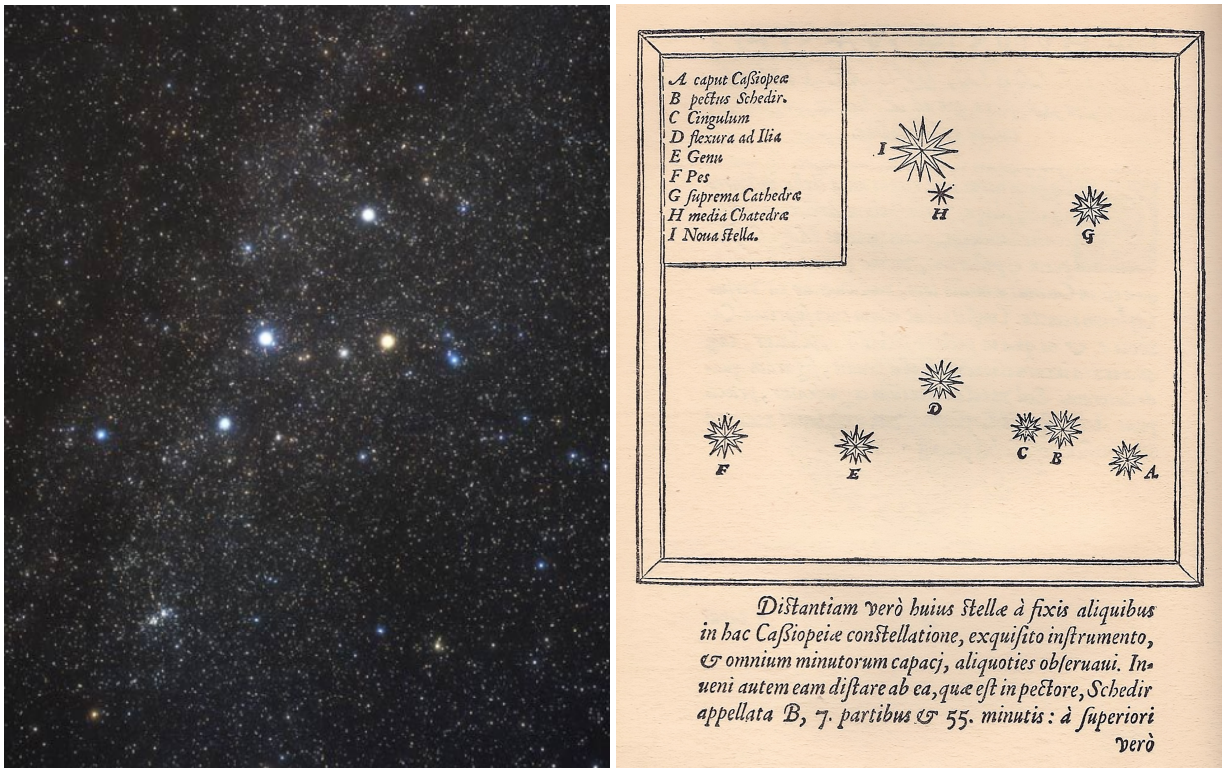

Figure 2 – On the left, a picture of Cassiopeia; from ESA/Hubble of 2006. On the right, a star map of the constellation Cassiopeia shows the position (labeled I) of SN 1572; from Tycho Brahe's De Nova Stella of 1573.

lines (MINKOWSKI, 1941). However, later observations contributed to differentiating various subtypes.

4.1 Type Ia Supernovae

Type Ia Supernovae (SNe Ia), previously known as Minkowski's SNe I, are defined by the absence of hydrogen in the spectral curve. Additionally, the light curve of these observations is characterized by a luminosity peak not long after the initial event. The spectra used to classify these SNe are taken near the peak luminosity.

SNe Ia occurs when a white dwarf restarts nuclear fusion by accreting enough matter from a binary partner. Once fusion starts, the white dwarf's gravitational forces cannot hold the high pressures building up inside, leading to carbon detonation and releasing sufficient energy to destroy the white dwarf. The result is an explosion that ejects all matter and releases the built-up internal energy.

During this process, large quantities of ^{56}Ni are produced. This isotope undergoes the following decay chain:

$$^{56}28\text{Ni} \rightarrow {}^{56}27\text{Co} + {}^{0}_{1}\text{e}^{+} + \gamma \rightarrow {}^{56}_{26}\text{Fe} + 2{}^{0}_{1}\text{e}^{+} + \gamma. \tag{4.1}$$

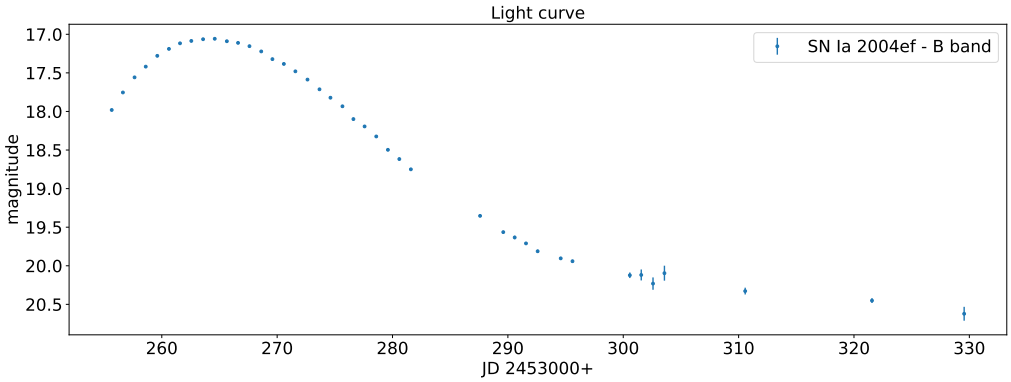

Figure 3 – Measured light curve of a SN Ia. Data extracted from The Carnegie Supernova Project (CONTRERAS et al., 2010)

The gamma radiation released during the decay of ^{56}Ni to ^{56}Fe is initially invisible to us. However, given the cloud of ejected matter (ejecta) surrounding the recently deceased star, this high-energy radiation will broaden its wavelength through various interactions. The ultraviolet, optical, and infrared radiation (UVOIR) generated by these interactions will be measured and used to construct the SNe Ia light curve (Figure 3).

Three simultaneous phenomena can holistically explain the overall shape of the SNe Ia light curve: energy injection from the radioactive decay of ^{56}Ni, the adiabatic conversion of internal energy to kinetic energy of the ejecta, and energy escaping as UVOIR light. We can divide the light curve into four phases (WRIGHT; LI, 2018).

I. At the very beginning of the explosion, the ejecta still is a densely packed opaque cloud surrounding the decaying nickel. The gamma rays created by the decay are losing energy to the ejecta and falling into the UVOIR range. But despite that, this light is not able to escape the current prison of matter; therefore, there is no measurement of the event during this epoch.

II. As the ejecta is dispersed throughout the space, it becomes more translucent allowing the light of all wavelengths to escape. Now, we can observe the supernova and begin the measurement process. The UVOIR luminosity rises significantly, but soon reaches a peak; given that the gamma rays are also leaving with less interaction with any matter whatsoever.

III. At this point, there is little interaction generating rays in the bandwidth of interest. The decreasing light being measured is a remnant from gamma rays only from earlier times of the event, which are now free because the ejecta is fully translucent to UVOIR wavelengths.

IV. Now that the ejecta is completely dispersed, no significant amount of UVOIR is generated in the inner layers of the supernova. This is characterized by the gradual dimming of the supernova across a period until it has completely faded away.

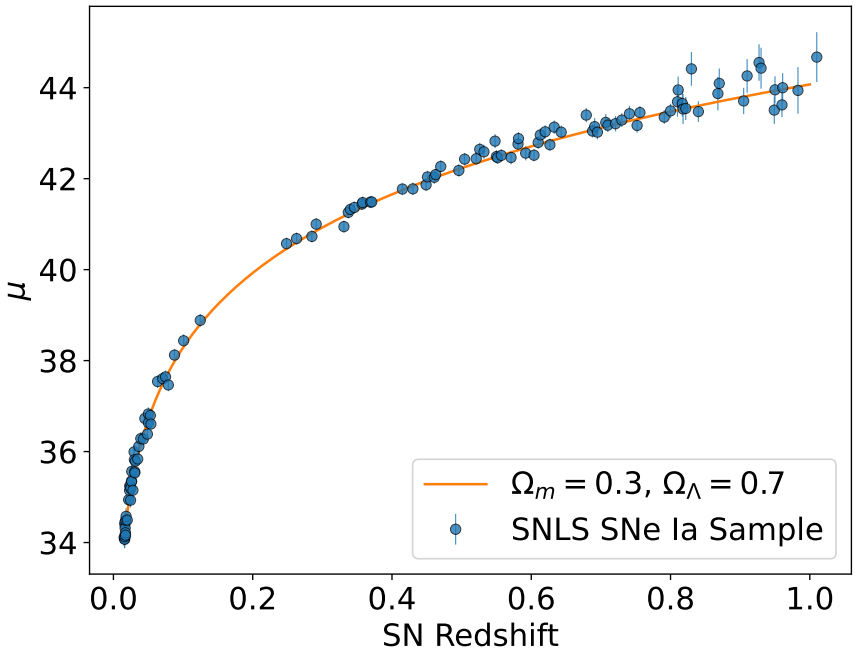

Figure 4 – Hubble diagram for SNLS SNe Ia (ASTIER et al., 2005) and the theoretical curve obtained from equations from Chapter 2 for $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

Precise modeling of the progenitor system and explosion mechanism of SNe Ia reinforce our belief that these objects have a predictable light curve. Moreover, observational data shows us a correlation between the apparent magnitude at a given time and the apparent magnitude at peak brightness. The light curve width-luminosity relation, also known as the Phillips relation, describes a linear relationship between the linear decline parameter (variation of the magnitude) and the peak magnitude (PHILLIPS, 1993).

Assuming that SNe Ia are standard candles at their maximum brightness, we can use the Phillips relation to find the magnitude at peak for a given supernova and compare it with a standard set of supernovae. This means that SNe Ia are standardized candles. By doing so, we can calculate the object's distance modulus, μ; together with the measurement of the redshift, we can plot a Hubble diagram.

As shown in Figure 4, the correlation of distance modulus and redshift of our sample is very similar to what we expected from our cosmological model. This fact reassures us that the usage of SNe Ia as standardized candles is accurate.

SNe Ia surveys (PERLMUTTER et al., 2003) revealed the possibility of an unknown form of energy, dubbed "dark energy". This energy opposes the self-gravitational attraction of matter and causes the Universe's expansion to accelerate. This discovery supports a model with a nonzero cosmological constant, Λ, where the equation of state (2.66) describes the properties of dark energy. In recognition of this breakthrough, the Nobel Prize was awarded in 2011 to the researchers "for the discovery of the accelerating expansion

of the Universe through observations of distant supernovae" (HUSS; MOBERG, 2011).

To investigate dark energy and better describe the possible cosmological model, the Supernova Legacy Survey (SNLS) began to observe high redshift supernovas in 2003. During the first year of observations, they classified 71 SNe Ia – of 400 detected transients – and managed to fit the cosmological parameters Ω_m and Ω_Λ (ASTIER et al., 2005).

With all the data collected in these observations, one might query the difference between the measured distance modulus and the expected value from our cosmological model. In other words, how the difference of μ^{obs} and $\mu(z^{obs}|\theta)$ is distributed? Without adding any complexity beyond necessities we assume that the desired distribution is gaussian with a null mean;

$$P(\mu(z^{obs}|\vec{\theta}) - \mu^{obs}) = \exp\left[-\frac{1}{2}\frac{(\mu(z^{obs}|\vec{\theta}) - \mu^{obs})^2}{\sigma^2}\right]. \tag{4.2}$$

Given that we measure the redshift, we can compute the distance modulus using equations (2.78) to (2.82). Together with the inferred μ and error σ, we compute the probability of this sampling event.

Finally, the likelihood of our SNe Ia sample given our cosmological parameters $\vec{\theta}$ is

$$\mathcal{L}(D|\vec{\theta}) = \exp\left[-\frac{1}{2}\sum_i \frac{(\mu(z_i^{obs}|\vec{\theta}) - \mu_i^{obs})^2}{\sigma_i^2}\right], \tag{4.3}$$

which is used to test different simple and composite hypotheses using the methods described in the last chapter.

4.2 Using statistical methods to estimate the cosmological parameters

In this section, we aim to apply the statistical analysis we previously discussed. First and foremost, we will be using the data collected by SNLS. For a particular set of parameters $\vec{\theta}$, we can compute the likelihood of the sample using Eq. (4.3). However, to simplify the calculations that are performed, we shall use the χ^2 defined as:

$$\chi^2(D|\vec{\theta}) = \left(\frac{\Omega_k - 0.0}{0.01}\right)^2 + \sum_i \frac{(\mu(z_i^{obs}|\vec{\theta}) - \mu_i^{obs})^2}{\sigma_i^2}. \tag{4.4}$$

We have added a factor of Ω_k to constrain its value, given that the data we are using is not taking into consideration a plethora of errors. Furthermore, now rather than finding the parameters that maximize the likelihood, we will find the ones that minimize χ^2. Besides, by switching to $-2\ln\mathcal{L} = \chi^2$, we assure that the ratio test will be distributed as a χ^2_n.

Given the high degeneracy of Ω_r, we opted to leave it fixed in our model. We will rewrite Eq. (2.78) with $\Omega_r = 9.2 \times 10^{-}5$ derived from Planck 2018 Results (AGHANIM

et al., 2020). Moreover, a minor change to Eq. (2.78) will allow us to find the confidence regions for the equation of state parameter for dark energy, w. The resultant parameter equation is

$$H^2 = H_0^2 \left(\Omega_\Lambda (1+z)^{-3(1+w)} + \Omega_m (1+z)^3 + \Omega_k (1+z)^2 + 9.2 \cdot 10^{-5} (1+z)^4 \right). \tag{4.5}$$

Using this expression for H together with other equations from the last section of the first chapter enables the computation of Eq. (4.4) for different redshifts.

To begin our analysis, we will find the set of parameter $\vec{\theta} = (\Omega_\Lambda, \Omega_m, w, H_0)$ which minimizes the χ^2. It is important to note that Ω_k will be calculated within the defined functions since it is constrained as

$$\Omega_k = 1 - (\Omega_\Lambda + \Omega_m + \Omega_r). \tag{4.6}$$

We proceed to perform this minimization using the package `SciPy` of the `Python Programming Language`. The package's object `optimize.minimize` is capable of finding the local minimum of a n-dimensional function $f(\vec{x})$ through an array of methods. We choose to use the Nelder-Mead Optimization algorithm (GAO; HAN, 2012).

This algorithm starts with a simplex, a set of $n+1$ vertices in n dimensions, that are evaluated and sorted. After sorting, the worst-performing point is transformed via a set of linear transformations and reevaluated. This process goes on until we achieve a given convergence criteria. For a deeper understanding of the subject and its implementation see (GAO; HAN, 2012). This optimization yielded the following parameters:

$$\vec{\theta} = (\Omega_\Lambda, \Omega_m, w, H_0) = (0.70808155, 0.29182358, -1.08027404, 70.15387593) \tag{4.7}$$

We will consider the parameters obtained from this optimization as our null hypothesis in the LRT. Therefore, we can calculate χ^2_{null} and save its value for future use.

The next step in our analysis is to find the confidence regions with the LRT. To achieve this, we have developed a script capable of performing the test for a given region of our parameter space. We can test up to two parameters simultaneously, and we can obtain all the necessary plots to create a corner plot of $(\Omega_\Lambda, \Omega_m, w, H_0)$.

Our script runs through a loop for each point in the grid, given a list of parameters, regions, and grid size. It minimizes the remaining parameters and computes the Eq. (4.4). Finally, the log ratio is calculated as

$$-2 \ln \lambda = \chi^2 - \chi^2_{null}; \tag{4.8}$$

and then stored in a tuple with the parameters $(\Omega_\Lambda, \Omega_m, w, H_0, -2 \ln \lambda)$. The obtained 2d-array is then transformed into a dictionary with keys `Omega_l`, `Omega_m`, `w`, `Hubble`, and `Ratio`.

Using the package `Matplotlib` we can create a level plot with `pyplot.tricontour`. Moreover, the desired levels we want to draw can be passed as a list to the object. Knowing

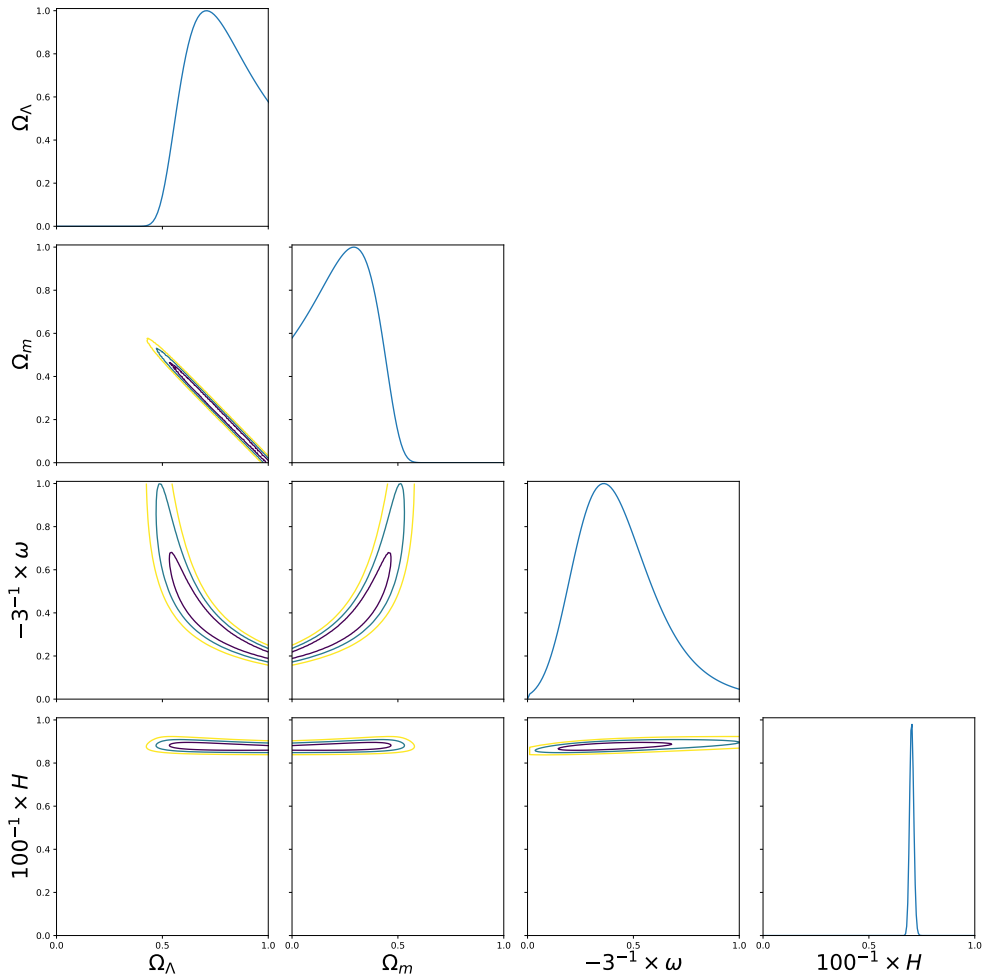

Figure 5 – Corner plot displaying the confidence regions for the cosmological parameters Ω_Λ, Ω_m, w, and H_0. The main diagonal shows the one-dimensional confidence regions for each parameter, while the lower triangular matrix represents the two-dimensional composite confidence regions for pairs of parameters in the corresponding column and row. The displayed confidence regions are for 1σ, 2σ, and 3σ.

that the ratio $-2 \ln \lambda$ is distributed as a chi-square with n degrees of freedom, we can calculate the required percentiles at a given significance level with `scipy.stats.chi2.ppf`. Together with the likelihood in one dimension, we can create the corner plot for the cosmological parameters (Figure 5).

Finally, we compared the performance of our LRT computation method with the one from `NumCosmo`. It was expected that our code would be slower since it is written in `Python` instead of `C`. However, the method of calculation also had a high impact on computation time. On average, it took us 57 minutes to process the two-dimensional ratios for 10,000 points. It is worth noting that our script is performing parallel computations, making use of all available cores. This average value was obtained when running the script on a Quad-Core CPU, therefore, being at least four times greater when measured using a single core. The same result is obtained in less than ten minutes using `NumCosmo` since it computes directly the level curve –and only the level curve– at a given significance level.

5 Conclusion

We began by providing an introduction to the foundations of modern cosmology, encompassing required topics such as differential geometry and general relativity. Moreover, we delved into various statistical methods, covering the essentials of statistical estimators and hypothesis testing. This allowed for a robust understanding of the tools necessary for the subsequent analysis of SNe Ia data.

We have also discussed the underlying physics governing SNe Ia and their characterization as standardizable candles, establishing their crucial role in cosmological studies. Using this knowledge in conjunction with the statistical concepts discussed earlier, we managed to apply the LRT to estimate cosmological parameters.

In addition to the aforementioned content, several notable observations emerged from the analysis. Firstly, when examining the individual two-dimensional confidence regions, their sizes provided interesting insights. It was observed that smaller areas at a given significance level indicated a higher degree of parameter degeneracy. Notably, the plots related to H displayed an exceptionally narrow region. Furthermore, the correlation between Ω_Λ and Ω_m was clearly evident in their plot, highlighting the interdependence of these parameters. This correlation was further accentuated by the utilization of an approximation for Ω_k in Eq. (4.4), which contributed to the slender region observed.

Furthermore, it is worth mentioning that the implemented method involved calculating data for every point in a grid, leading to a larger computational load than strictly necessary for the analysis. Although this approach resulted in sub-optimal computation time, significant reductions were achieved by employing parallel computation techniques. However, it is important to note that further improvements in computational efficiency could be attained by leveraging the power of GPU acceleration.

Overall, our study reaffirms the power of statistical inference techniques applied to SNe Ia datasets to estimate cosmological parameters. Additionally, considerations regarding computational efficiency were addressed, acknowledging both the potential for optimization through parallel computation and the prospects of GPU utilization. These insights contribute to the broader field of cosmology and pave the way for future advancements in parameter estimation and cosmological studies.

Bibliography

AGHANIM, N. et al. Planck 2018 results-vi. cosmological parameters. *Astronomy & Astrophysics*, EDP sciences, v. 641, p. A6, 2020. 17, 48

ASTIER, P. et al. Snls: Measurement of ω_m, ω_λ and ω from the first year data set. In: *American Astronomical Society Meeting Abstracts*. [S.l.: s.n.], 2005. v. 207, p. 15–04. 17, 46, 47

BAADE, W.; ZWICKY, F. On super-novae. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 20, n. 5, p. 254–259, 1934. 43

BARLOW, R. J. *Statistics*. [S.l.: s.n.], 1993. v. 29. (Manchester Physics Ser., v. 29). 17, 33

BRAHE, T. *Tychonis Brahe De nova et nullius aevi memoria prius visa stella: iam pridem anno a nato Christo 1572, mense Novembri primum conspecta, contemplatio mathematica*. Hafniae Impressit Lavrentivs, 1573. Disponível em: <https://books.google.com.br/books?id=kSQLAAAAMAAJ>. 43

CARROLL, S. B. *Spacetime and Geometry*. [S.l.: s.n.], 2003. 17, 25

CONTRERAS, C. et al. The carnegie supernova project: first photometry data release of low-redshift type ia supernovae. *The Astronomical Journal*, IOP Publishing, v. 139, n. 2, p. 519, 2010. 17, 45

DODELSON, S. *Modern Cosmology*. [S.l.]: Academic Press, 2003. 17, 26, 30

EGGEN, O. J. Tycho brahe. Encyclopedia Britannica, apr 14 2023. [Online; accessed 14-03-2023]. Disponível em: <https://www.britannica.com/biography/Tycho-Brahe-Danish-astronomer>. 33

EINSTEIN, A. On the electrodynamics of moving bodies. *Annalen der physik*, v. 17, n. 10, p. 891–921, 1905. 19

EINSTEIN, A.; MINKOWSKI, H. The foundation of the generalised theory of relativity. *On a Heuristic Point of View about the Creation and Conversion of Light*, v. 1, p. 22, 1920. 17, 19

GAO, F.; HAN, L. Implementing the nelder-mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, Springer, v. 51, n. 1, p. 259–277, 2012. 48

HUBBLE, E. P. Extragalactic nebulae. *Astrophysical Journal, 64, 321-369 (1926)*, v. 64, 1926. 17

HUGGINS, W. *The Scientific Papers of Sir William Huggins*. [S.l.]: Ex Libris, 1899. 43

HUSS, E.; MOBERG, A. The nobel prize in physics 2011. Nobel Prize Outreach AB 2023, dec 10 2011. [Online; accessed 20-05-2023]. Disponível em: <https://www.nobelprize.org/prizes/physics/2011/summary/>. 17, 47

ISHAM, C. J. *Modern Differential Geometry for Physicists*. [S.l.: s.n.], 1999. Vol. 61. (World Scientific Lecture Notes in Physics, Vol. 61). 20

LORENTZ, H. A. Michelson's interference experiment. *The Principle of Relativity. Dover Books on Physics. June 1*, p. 1–7, 1952. 19

MINKOWSKI, H.; EINSTEIN, A.; MAHALANOBIS, P. C. *The Principle of Relativity*. Project Gutenberg., 2021. [Online; accessed 2023-01-24]. Disponível em: <https://www.gutenberg.org/cache/epub/66944/pg66944-images.html>. 19

MINKOWSKI, R. Spectra of supernovae. *Publications of the Astronomical Society of the Pacific*, v. 53, p. 224, 1941. 44

MUKHANOV, V. F. *Physical foundations of cosmology*. [S.l.]: Cambridge university press, 2005. 26

NUSSBAUMER, H. Einstein's conversion from his static to an expanding universe. *The European Physical Journal H*, Springer Science and Business Media LLC, v. 39, n. 1, p. 37–62, Feb 2014. Disponível em: <http://dx.doi.org/10.1140/epjh/e2013-40037-6>. 17, 26

PERLMUTTER, S. et al. Supernovae, dark energy, and the accelerating universe. *Physics today*, American Institute of Physics, v. 56, n. 4, p. 53–62, 2003. 17, 46

PHILLIPS, M. M. The absolute magnitudes of type ia supernovae. *The Astrophysical Journal*, v. 413, p. L105–L108, 1993. 46

RABIN, S. J. Johannes kepler. new astronomy. trans. william h. donahue. cambridge and new york: Cambridge university press, 1992. xvi+ 665 pp. 140. *Renaissance Quarterly*, Cambridge University Press, v. 49, n. 1, p. 176–177, 1996. 33

RICE, J. A. *Mathematical statistics and data analysis*. [S.l.]: China machine press, 2003. 17, 40

SHEEHAN, W. *The planet Mars: A history of observation & discovery*. [S.l.]: University of Arizona Press, 1996. 33

WESTMAN, R. S. Johannes kepler. Encyclopedia Britannica, feb 15 2023. [Online; accessed 14-03-2023]. Disponível em: <https://www.britannica.com/biography/Johannes-Kepler>. 33

WRIGHT, B. S.; LI, B. Type ia supernovae, standardizable candles, and gravity. *Physical Review D*, APS, v. 97, n. 8, p. 083505, 2018. 17, 45