



Universidade Federal de Campina Grande
Centro de Educação e Saúde
Unidade Acadêmica de Física e Matemática
Física - Licenciatura

Vanderson Costa Cavalcante

Estatística Bayesiana e Estimativa de Parâmetros
Cosmológicos

Cuité - PB

Maio de 2024

Vanderson Costa Cavalcante

Estatística Bayesiana e Estimativa de Parâmetros Cosmológicos

Monografia apresentada ao curso de Licenciatura em Física, Unidade Acadêmica de Física e Matemática, Centro de Educação e Saúde, Universidade Federal de Campina Grande, como requisito para obtenção do título de Licenciado em Física.

Universidade Federal de Campina Grande
Centro de Educação e Saúde
Unidade Acadêmica de Física e Matemática
Física - Licenciatura

Orientador: Prof. Dr. Joseclécio Dutra Dantas

Cuité - PB
Maio de 2024

C376e Cavalcante, Vanderson Costa.

Estatística Bayesiana e estimativa de parâmetros cosmológicos. / Vanderson Costa Cavalcante. - Cuité, 2024. 93 f.

Trabalho de Conclusão de Curso (Licenciatura em Física) - Universidade Federal de Campina Grande, Centro de Educação e Saúde, 2024. "Orientação: Prof. Dr. Joseclécio Dutra Dantas".

Referências.

1. Cosmologia. 2. Parâmetros cosmológicos. 3. Tratamento de dados cosmológicos. 4. Estatística Bayesiana. 5. Energia escura. 6. Modelos cosmológicos. 7. Centro de Educação e Saúde. I. Dantas, Joseclécio Dutra. II. Título.

CDU 52(043)

Vanderson Costa Cavalcante

Estatística Bayesiana e Estimativa de Parâmetros Cosmológicos

Monografia apresentada ao curso de Licenciatura em Física, Unidade Acadêmica de Física e Matemática, Centro de Educação e Saúde, Universidade Federal de Campina Grande, como requisito para obtenção do título de Licenciado em Física.

Trabalho aprovado. Cuité - PB, 13 de maio de 2024:

Prof. Dr. Joseclécio Dutra Dantas
Orientador

Professor
Prof. Dr. Fábio Ferreira de Medeiros
Examinador

Professor
Prof. Dr. Vladimir Soares Catão
Examinador

Professor
Prof. Dr. Heron Neves de Freitas
Examinador - Suplente

Cuité - PB
Maio de 2024

*Este trabalho é dedicado às crianças adultas que,
quando pequenas, sonharam em se tornar cientistas.*

Agradecimentos

Gostaria de expressar minha profunda gratidão a todos que contribuíram nesta jornada de crescimento pessoal e intelectual, chamada graduação.

Primeiramente, gostaria de agradecer a Deus por ter me guiado nesta trajetória e por ter me dado força para superar todos os empecilhos que surgiram durante o percurso, além de ter colocado pessoas incríveis na minha vida.

Aos meus pais, Ivaneide Silva de Brito Cavalcante e Evandro Costa Cavalcante, por todo o amor, apoio e incentivo que me deram desde o primeiro momento que fui aprovado no curso. Vocês sempre me ensinaram a não ter medo de ir atrás dos meus sonhos. Sem o apoio de vocês, essa conquista não seria possível. Obrigado por tudo, amo vocês.

À minha irmã, Jéssica Cavalcante, e meu sobrinho, Gutiérrez Cavalcante, tios e avós por sempre acreditarem em mim.

Aos primos, em especial Erica Dayana, e Emanuel Pedro, por todo suporte, apoio e ensinamentos recebidos.

A todos os meus amigos que o curso me presenteou, em especial Maria Suêd, Karla, Frank Marcelo, Isac, Simone, Marcília, Layza e Laura, que tornaram esta etapa da minha vida mais leve, com momentos de conversa, risadas e brincadeiras. Vocês tornaram esta etapa da minha vida inesquecível e mais significativa. Agradeço por cada momento compartilhado e por todo apoio e amizade demonstrados ao longo desses anos.

A Dona Roseane, por sua acolhida generosa durante às vezes que fui até sua residência, sempre muito cuidadosa, preparava refeições deliciosas que me traziam boas recordações e faziam com que eu me sentisse em casa. Obrigado por abrir suas portas para mim e saiba que cada momento compartilhado em sua companhia foi verdadeiramente especial.

Ao meu preceptor do Programa Residência Pedagógica (PRP), Pablo José Lima Soares, por todo o apoio, ensinamentos e orientação durante o processo. Com você, obtive diversos aprendizados que impulsionaram meu crescimento pessoal e profissional. Agradeço pelo suporte e pela confiança depositados em mim.

Ao meu orientador, Joseclécio Dutra Dantas, pela orientação, apoio e ensinamentos que recebi ao longo deste processo. Sua orientação foi essencial para o desenvolvimento e conclusão deste trabalho. Além disso, seus ensinamentos foram fundamentais para meu crescimento acadêmico.

Por fim, quero agradecer aos professores do curso de modo geral, especialmente a

João Batista, que sempre esteve disposto a tirar dúvidas em relação à disciplina, sempre muito paciente e atencioso. Se cheguei até aqui, foi graças a sua ajuda.

“a natureza é um enorme jogo de xadrez disputado por deuses, e que temos o privilégio de observar. As regras do jogo são o que chamamos de física fundamental, e compreender essas regras é a nossa meta”.
(Richard Feynman)

Resumo

Este trabalho visa compreender a aplicação da estatística Bayesiana no tratamento de dados cosmológicos. Este é o paradigma estatístico presente na apresentação e atualização de parâmetros cosmológicos, como o parâmetro de Hubble e os parâmetros de densidade dos diversos constituintes do Universo. O tipo de pesquisa desenvolvido é a revisão bibliográfica. Como aplicação, apresentaremos dois modelos cosmológicos: o Λ CDM e um modelo de energia escura descrito por uma parametrização em termos do *redshift*. Com o método de Monte Carlo e Cadeia de Markov (MCMC) em Python, foi realizado o tratamento de dados de Supernovas do Tipo Ia (SNe Ia) e dados de $H(z)$ de Relógios Cósmicos, na tentativa de estimar o parâmetro de densidade de matéria do universo (Ω_{0m}). Também apresentamos uma comparação entre esses dois modelos.

Palavras-chave: Parâmetros cosmológicos, Tratamento de dados, estatística Bayesiana, energia escura, modelos cosmológicos.

Abstract

This work aims to understand the application of Bayesian statistics in the treatment of cosmological data. This is the statistical paradigm present in the presentation and updating of cosmological parameters, such as the Hubble parameter and the density parameters of the various constituents of the Universe. The type of research carried out is bibliographic review. As an application, we present two cosmological models: the Λ CDM and a dark energy model described by a parameterization in terms of redshift. Using the Monte Carlo and Markov Chain (MCMC) method in Python, the treatment of Type Ia Supernova (SNe Ia) data and $H(z)$ data from Cosmic Clocks was performed, in an attempt to estimate the matter density parameter of the universe (Ω_{0m}). We also present a comparison between these two models.

Key-words: Cosmological parameters, Data processing, Bayesian statistics, Dark energy, Cosmological models.

Lista de ilustrações

Figura 1 – Representação gráfica do fator de Bayes.	58
Figura 2 – Importando as bibliotecas necessárias.	63
Figura 3 – Atribuindo valores às constantes conhecidas	64
Figura 4 – Carregando os dados.	65
Figura 5 – Organizando os dados.	65
Figura 6 – Definindo parametrizações.	67
Figura 7 – Quinto Bloco de Código	67
Figura 8 – Definindo o parâmetro de Hubble adimensional.	68
Figura 9 – Distância-luminosidade.	68
Figura 10 – Estimando Ω_{0m} no modelo Λ CDM.	70
Figura 11 – Resultado do Ω_{0m}	71
Figura 12 – Definindo o log da verossimilhança.	71
Figura 13 – Carregando dados dos modelos cosmológicos.	73
Figura 14 – Gerando o gráfico da distância luminosa.	73
Figura 15 – Distância-luminosidade.	74
Figura 16 – Definindo a distribuição a priori.	75
Figura 17 – Definindo a distribuição a posteriori.	76
Figura 18 – Definindo a distribuição a priori.	76
Figura 19 – Definindo a distribuição a posteriori.	77
Figura 20 – Implementando um cálculo de MCMC.	78
Figura 21 – Organizando as amostras.	79
Figura 22 – Gerando o gráfico das distribuições.	79
Figura 23 – Distribuição a posterior para Ω_{0m} durante a amostragem MCMC.	80
Figura 24 – Estimativa do valor de Ω_{0m} do modelo Λ CDM.	81
Figura 25 – Implementando um cálculo de MCMC.	82
Figura 26 – Organizando as amostras.	83
Figura 27 – Estimativa de valores dos parâmetros do modelo de energia escura.	83
Figura 28 – Organizando as amostras.	84
Figura 29 – Distribuição a posteriori unidimensional e contornos bidimensionais para o espaço de parâmetros (w_0, w_1, Ω_{0m}) referentes ao modelo de energia escura 3.3.	85
Figura 30 – Valores dos parâmetros w_0, w_1 e Ω_m	85

Lista de tabelas

Tabela 1 – Parâmetros Cosmológicos	37
Tabela 2 – Distribuições a Priori Conjugadas.	55

Lista de abreviaturas e siglas

Λ CDM	Lambda Cold Dark Matter (Lambda Matéria Escura Fria)
FLRW	Friedmann-Lemaître-Robertson-Walker
RG	Relatividade Geral
CMB	Cosmic Microwave Background (Radiação Cósmica de Fundo em Micro-ondas)
BAO	Baryon Acoustic Oscillations (Oscilações Acústicas de Bárions)
LSS	Large-Scale Structure (Estrutura em Grande Escala)
FDP	Função Distribuição de Probabilidade
FDC	Função de distribuição Cumulativa
FMP	Função Massa de Probabilidade
MCMC	Monte Carlos de Cadeia de Markov

Lista de símbolos

Γ	Letra grega Gama
Λ	Letra grega Lambda
\in	Pertence
ν	Letra grega ni
μ	Letra grega mi
σ	Letra grega sigma
θ	Letra grega teta
ϕ	Letra grega fi
γ	Letra grega gama
η	Letra grega eta
α	Letra grega alfa
δ	Letra grega delta
β	Letra grega beta
ρ	Letra grega rô
π	Letra grega pi
Ω	Letra grega Ômega
Δ	Letra grega Delta
Σ	Somatório
\cup	Representa a união de conjuntos
\cap	Representa a interseção de conjuntos
\propto	Proporcional
χ	Letra grega chi

Sumário

	Introdução	25
1	UMA INTRODUÇÃO À COSMOLOGIA	27
1.1	Equações de Friedmann	27
1.2	Equação de Continuidade	32
1.3	Equação de Estado	33
1.4	Equação de Aceleração	34
1.5	Lei de Hubble	34
1.6	Desvio Para o Vermelho (<i>Redshift</i>)	35
1.7	Modelo Cosmológico ΛCDM	36
1.8	Energia Escura	37
1.8.1	Modelo da Energia Escura	38
1.9	Supernovas do Tipo Ia (SNe Ia)	39
1.10	Análise Utilizando Dados de SNe Ia	40
2	PROBABILIDADE E ESTATÍSTICA	41
2.1	Axiomas da Probabilidade	41
2.2	Função Distribuição de Probabilidade	42
2.3	Função de Distribuição Cumulativa	42
2.4	Distribuição de Probabilidade Discreta	43
2.4.1	Variável Discreta	43
2.4.2	Função Massa de Probabilidade	43
2.5	Principais Distribuições Discretas	43
2.5.1	Uniforme Discreta	43
2.5.2	Distribuição de Bernoulli	44
2.5.3	Distribuição Binomial	44
2.5.4	Distribuição de Poisson	45
2.6	Distribuição de Probabilidade Contínua	45
2.6.1	Variável Contínua.	45
2.6.2	Função Densidade de Probabilidade	45
2.7	Principais Distribuições Contínuas	46
2.7.1	Uniforme Contínua	46
2.7.2	Distribuição Normal	46
2.7.3	Distribuição Exponencial	47
2.8	Probabilidade Condicional	47
2.9	Probabilidade Conjunta	47

2.10	Estatística Bayesiana	48
2.10.1	Frequentista X Bayesiana	48
2.10.2	Teorema de Bayes	49
2.11	A Priori	50
2.11.1	Tipos de Prioris	51
2.11.1.1	Distribuições a Priori Não Informativas	52
2.11.1.2	Distribuições a Priori Fracamente Informativas	52
2.11.1.3	Distribuições a Priori Informativas	52
2.12	Verossimilhança	53
2.12.1	Método de Máxima Verossimilhança	53
2.13	Distribuição a Posteriori	54
2.14	Prioris Conjugadas	55
2.15	Regra de Bayes	55
2.15.1	Probabilidade Preditiva Anterior $P(Y)$	56
2.15.2	Fator de Bayes	56
2.15.3	Verossimilhança Marginal	57
2.16	Métodos de Monte Carlo de Cadeias de Markov (MCMC)	58
2.16.1	O Método de Metropolis	59
2.16.2	Funcionamento Método de Metropolis	60
2.16.3	Metropolis-Hastings	60
3	ESTIMANDO PARÂMETROS COM MCMC	63
3.1	Bibliotecas	63
3.2	Constantes	64
3.3	Conjunto de dados	64
3.4	Definição de funções	66
3.5	Função de Máxima Verossimilhança	69
3.6	Método MCMC	74
	Considerações finais	87
	REFERÊNCIAS	89

Introdução

A Teoria do Big Bang representa um paradigma de grande sucesso. Este sucesso é sustentado por um número considerável de pilares observacionais confiáveis. O diagrama de Hubble, a partir do qual se verifica a expansão do universo; a abundância de elementos leves, que concordam com as previsões da Nucleossíntese do Big Bang (NBB); a temperatura e as anisotropias da radiação cósmica de fundo (RCF), concordantes com a teoria, são exemplos desses pilares (DODELSON; SCHMIDT, 2021).

Em 1929, E. Hubble observou o desvio para o vermelho na luz emitida por galáxias distantes. A lei de Hubble enuncia que a velocidade de recessão das galáxias é proporcional a suas distâncias. Isto representa uma evidência observacional importante da expansão do universo (HASHIMOTO *et al.*, 2018). A ideia de expansão sugeriu a George Gamow (GAMOW, 1946), Ralph Alpher e Robert Herman (ALPHER; BETHE; GAMOW, 1948; ALPHER; HERMAN, 1950) que, em determinada fase, a radiação desacoplou da matéria e se propagou pelo Universo. Esta radiação, a RCF, foi prevista em 1940 e detectada cerca de duas décadas depois, em 1965, pelos radioastrônomos Arno Penzias e Robert Wilson (PENZIAS; WILSON, 1965). A temperatura da RCF, em torno de 2,7 K, é uma das medidas mais precisas na Cosmologia atual (FIXSEN, 2009; EVANS, 2015). A uniformidade observada na temperatura da RCF vinda de diversas direções do céu, concorda com o Princípio Cosmológico que, na Teoria do Big Bang, assume a homogeneidade e isotropia do Universo em grande escala. Essas descobertas/observações fazem da cosmologia hoje uma área muito bem fundamentada.

A cosmologia atual é alimentada por uma grande quantidade de dados observacionais coletados em experimentos e instrumentos que envolvem colaborações internacionais importantes. Dentre os instrumentos, podemos citar os mais conhecidos, em ordem cronológica de início de funcionamento. O Cosmic Background Explorer (COBE) foi lançado pela agência norte americana National Aeronautics and Space Administration (NASA) em 1990. As medidas foram refinadas pelo Wilkinson Microwave Anisotropy Probe (WMAP), lançado pela NASA em 2001. Os dados mais recentes vêm do satélite Planck, uma colaboração da NASA e da European Space Agency (ESA), lançado em 2009 (EVANS, 2015; PETTINARI, 2016).

Os dados coletados nessas observações têm contribuído para que a comunidade científica possa entender melhor o universo, porque os dados confirmam previsões, como a homogeneidade e isotropia em grande escala, mas também alimentam novos desafios, novos problemas a serem solucionados, como é o caso do problema do horizonte, que diz respeito justamente à observação da homogeneidade e isotropia (BAUMANN; MCALIST-

TER, 2015; DIMOPOULOS, 2021). Trabalhos recentes importantes, fundamentados nos dados da missão Planck, apresentam valores de parâmetros cosmológicos extraídos da análise do espectro de potência da RCF (DODELSON; SCHMIDT, 2021; DURRER, 2020; GIOVANNINI, 2008). Observações de supernovas do tipo Ia fornecem as medidas mais precisas de distâncias e da expansão do universo (RIESS et al., 2018; SULLIVAN et al., 2011), apontando para a existência de um conteúdo exótico desconhecido - a energia escura -, ou para a necessidade de uma modificação da teoria da gravitação. Novas descobertas impulsionam novas ideias, novas proposições, novos desafios.

Mas esses novos dados precisam ser interpretados de maneira cuidadosa. E essa interpretação exige um tratamento estatístico que atenda às necessidades de validação de modelos a partir dos dados observacionais e do conhecimento teórico prévio. É nesse contexto que a estatística Bayesiana tem contribuído: no tratamento de dados cosmológicos e refinamento de parâmetros (RIESS et al., 2018; SULLIVAN et al., 2011). A pesquisa desenvolvida é a pesquisa bibliográfica e tem como objetivo compreender os principais conhecimentos relacionados à estatística Bayesiana contribuindo para a compreensão inicial dos principais conhecimentos com o intuito de entender como é feito o tratamento de dados cosmológicos atuais de acordo com esse paradigma estatístico.

No Capítulo 1, realizamos uma breve revisão sobre os principais fundamentos da cosmologia moderna. Iniciamos com as equações de Friedmann, que são soluções das equações de Einstein, onde é possível entender como ocorre a expansão do universo. Descrevemos a lei de Hubble e concluímos o capítulo explicando como é feita a análise de dados de uma supernova. O Capítulo 2 é dedicado a uma introdução abrangente da estatística Bayesiana. Inicialmente, foi realizada uma revisão de alguns axiomas e conceitos de probabilidades que consideramos necessários para o entendimento dos paradigmas Bayesianos. Em seguida, explicamos o teorema de Bayes e como é feita a análise e interpretação de dados na perspectiva Bayesiana. Para finalizar, no Capítulo 3, serão apresentados e comentados o passo a passo de um código em Python que utiliza o MCMC para estimar parâmetros de modelos cosmológicos.

1 Uma Introdução à Cosmologia

A teoria do Big Bang se fundamenta na ideia cosmológica de que o universo é homogêneo e isotrópico em grande escala, um conceito apoiado por várias observações que confirmam a expansão cósmica. A partir das soluções das equações de campo de Einstein nesse contexto, é possível reconstruir o passado da expansão do cosmos (DIMOPOULOS, 2021). Esse conceito é o cerne do modelo padrão da cosmologia, denominado Λ CDM, que explica com êxito a evolução e a configuração do Universo visível, incorporando a matéria escura, energia escura e a aceleração da expansão cósmica. Neste capítulo, apresentaremos as equações de Friedmann, que são soluções das equações de Einstein da Relatividade Geral, assim como a Lei de Hubble, o Modelo Cosmológico Λ CDM, uma discussão sobre Energia Escura e sobre a Análise de dados de Supernovas do Tipo Ia (SNe Ia), dentre outras coisas.

1.1 Equações de Friedmann

O modelo cosmológico do espaço-tempo Friedmann-Lemaître-Robertson-Walker (FLRW) é representado por um elemento de linha, solução das equações de Einstein, que assume que o espaço-tempo tem as seguintes propriedades geométricas: é homogêneo, isotrópico e possui 4 dimensões (AMENDOLA; TSUJIKAWA, 2010).

O elemento de linha FLRW envolve coordenadas espaço-temporais (como tempo t e coordenadas espaciais (x, y, z)) e é dado pela expressão

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a^2(t) d\sigma^2, \quad (1.1)$$

onde:

- μ e ν variam de 0 a 3.
- $g_{\mu\nu}$ é um tensor métrico (AMENDOLA; TSUJIKAWA, 2010), que é responsável por determinar como as distâncias entre eventos são medidas em diferentes sistemas de coordenadas e em diferentes pontos do espaço-tempo. Além disso, ele captura as informações sobre a curvatura e a distância de um espaço-tempo curvo.
- $a(t)$ é um fator de escala com o tempo cósmico t (AMENDOLA; TSUJIKAWA, 2010). É a relação entre o tamanho do universo em um determinado tempo t com o tamanho do universo em algum momento de referência. De modo que a taxa de mudança de $a(t)$ com relação ao tempo cósmico representa a velocidade com que o universo está se expandindo.

- $d\sigma^2$ é a métrica tridimensional com curvatura constante K (AMENDOLA; TSUJIKAWA, 2010). Fornece a medida da distância entre pontos do espaço tridimensional. A métrica é independente o tempo. Dessa forma, a geometria do espaço tridimensional permanece a mesma ao longo do tempo. Assim propriedades do espaço com a curvatura permanecem constantes. $d\sigma^2$ pode ser calculado por meio da expressão abaixo:

$$d\sigma^2 = \gamma_{ij} dx^i dx^j = \frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2), \quad (1.2)$$

onde K representa a curvatura constante e pode assumir três valores, $K = +1, 0, -1$ que representam geometrias diferentes (DIMOPOULOS, 2021; AMENDOLA; TSUJIKAWA, 2010), descritas a seguir:

- $K = +1$: universo fechado, correspondente à geometria de uma esfera tridimensional (MUKHANOV, 2005).
- $K = 0$: representa uma geometria do universo plana (MUKHANOV, 2005).
- $K = -1$: universo aberto com geometria hiperbólica (MUKHANOV, 2005).

Os índices i e j variam de 1 a 3. Na equação (1.2) estamos usando as coordenadas polares, em que $(x^1, x^2, x^3) = (r, \theta, \phi)$ e $y_{11} = (1 - Kr^2)^{-1}$, $y_{22} = r^2$ e $y_{33} = r^2 \sin^2 \theta$.

Para buscar soluções específicas das equações de Friedmann, é frequentemente útil trocar o tempo físico t pelo tempo conforme η (MUKHANOV, 2005), o qual é definido como:

$$\eta \equiv \int \frac{1}{a} dt. \quad (1.3)$$

O Tempo conforme facilita a descrição matemática da expansão do espaço-tempo. Segundo Rodrigues et al. (2011) podemos entendê-lo como sendo “um relógio cuja a medida diminuiu com o aumento do universo, de forma que a distância comóvel por unidade de tempo conforme é constante”.

A partir da equações de campo de Einstein, podemos chegar às equações dinâmicas de movimento. A primeira etapa para encontrar as equações dinâmicas é calcular os símbolos de Christoffel, que podem ser obtidos por meio da métrica $g_{\mu\nu}$ e suas derivadas (DIMOPOULOS, 2021; AMENDOLA; TSUJIKAWA, 2010).

$$\Gamma_{\lambda\mu}^{\nu} = \frac{1}{2} g^{\mu\alpha} (g_{\alpha\nu,\lambda} + g_{\alpha\lambda,\nu} - g_{\nu\lambda,\alpha}), \quad (1.4)$$

onde:

- $g_{\alpha\nu,\lambda} \equiv \frac{\partial g_{\alpha\nu}}{\partial x^\lambda}$.

- $g_{\mu\nu}$ satisfaz a relação $g_{\mu\alpha}g^{\alpha\nu} = \delta_{\mu}^{\nu}$, que conecta o coeficiente métrico com a sua inversa.
- δ_{μ}^{ν} é o delta de Kronecker, que é igual a 1 quando os índices forem iguais ($\mu = \nu$) e 0, quando forem diferentes ($\mu \neq \nu$) (AMENDOLA; TSUJIKAWA, 2010).

Os símbolos de Christoffel $\Gamma_{\lambda\mu}^{\nu}$ descrevem como as coordenadas do espaço-tempo mudam ao longo de trajetórias geodésicas. De acordo com Fleisch (2017), cada símbolo de Christoffel é um coeficiente que indica a projeção da derivada de um vetor de base em determinada direção:

$$\Gamma_{\lambda\mu}^{\nu} \vec{e}_{\nu} = \frac{\partial \vec{e}_{\lambda}}{\partial x_{\mu}}. \quad (1.5)$$

Em (1.5), $\Gamma_{\lambda\mu}^{\nu}$ representa a projeção do vetor $\partial \vec{e}_{\lambda} / \partial x_{\mu}$ na direção \vec{e}_{ν} .

Podemos definir o tensor de Ricci em termos dos símbolos de Christoffel, pela seguinte expressão:

$$R_{\mu\nu} = \Gamma_{\mu,\alpha\nu}^{\alpha} - \Gamma_{\mu\nu,\alpha}^{\alpha} + \Gamma_{\mu\nu}^{\alpha} \Gamma_{\alpha\beta}^{\beta} - \Gamma_{\mu\beta}^{\alpha} \Gamma_{\alpha\nu}^{\beta}. \quad (1.6)$$

Essa equação mostra como a taxa de mudança dos símbolos de Christoffel influenciam a curvatura de um ponto específico do espaço-tempo.

Somando as componentes dos tensor de Ricci, chegamos a sua contração, que é dada pelo escalar de Ricci (AMENDOLA; TSUJIKAWA, 2010):

$$R = g^{\mu\nu} R_{\mu\nu}. \quad (1.7)$$

O escalar de Ricci é o cálculo da curvatura média do espaço-tempo em um determinado ponto.

Agora podemos definir o tensor de Einstein, que é dado por:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R. \quad (1.8)$$

Como podemos ver, o tensor de Einstein é formado pelos tensor e escalar de Ricci (AMENDOLA; TSUJIKAWA, 2010). Resolvendo as equações de Einstein, podemos obter a dinâmica cosmológica (AMENDOLA; TSUJIKAWA, 2010). As equações de Einstein são escritas da forma.

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \quad (1.9)$$

em que:

- $G_{\mu\nu}$ é o tensor Einstein.
- $T_{\mu\nu}$ é o tensor de energia-momento dos componentes da matéria.

- G é a constante gravitacional.

A equação (1.9) relaciona a curvatura do espaço-tempo, representada pelo tensor de Einstein, com a distribuição de energia e momento dos componentes de matéria, representada pelo tensor de energia-momento.

Para a métrica FLRW, Equação (1.1), as componentes não nulas dos símbolos de Christoffel são:

$$\begin{aligned}\Gamma_{ij}^0 &= a^2 H \gamma_{ij}, \\ \Gamma_{0j}^i &= \Gamma_{j0}^i = H \delta_j^i,\end{aligned}$$

$$\begin{aligned}\Gamma_{11}^1 &= \frac{Kr}{1 - kr^2}, \\ \Gamma_{22}^1 &= -r(1 - Kr^2), \\ \Gamma_{33}^1 &= -r(1 - Kr^2) \sin^2 \theta,\end{aligned}$$

$$\begin{aligned}\Gamma_{33}^2 &= -\sin \theta \cos \theta, \\ \Gamma_{12}^2 &= \Gamma_{21}^2 = \Gamma_{13}^2 = \Gamma_{31}^2 = \frac{1}{r},\end{aligned}$$

$$\Gamma_{23}^3 = \Gamma_{32}^3 = -\cot \theta,$$

onde H é o parâmetro de Hubble, dado pela razão entre a derivada temporal do fator escala, $\dot{a}(t)$, e fator escala, $a(t)$:

$$H \equiv \frac{\dot{a}}{a}. \quad (1.10)$$

O parâmetro H descreve a taxa de expansão do universo.

Com base nas equações (1.6) e (1.7), obtemos os componentes do tensor de Ricci e o escalar de curvatura, como segue:

$$\begin{aligned}R_{00} &= -3 \left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} \right), \\ R_{0i} &= R_{i0} = 0, \\ R_{ij} &= a^2 \left(2 \frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + 2K \right) \gamma_{ij},\end{aligned}$$

$$R = 6 \left(2H^2 + H + \frac{K}{a^2} \right).$$

Utilizando a relação $G^\mu_\nu = g^{\mu\alpha}G_{\alpha\nu}$, determinamos os componentes do tensor de Einstein:

$$\begin{aligned} G^0_0 &= -3 \left(H^2 + \frac{K}{a^2} \right), \\ G^0_i &= G^i_0 = 0, \\ G^i_j &= - \left(3H^2 + 2\dot{H} + \frac{K}{a^2} \right) \delta^i_j. \end{aligned}$$

No espaço-tempo FLRW, o tensor de energia-momento assume a forma de fluido perfeito (AMENDOLA; TSUJIKAWA, 2010), que é um fluido com as seguintes características: homogêneo, isotrópico, sem viscosidade ou dissipação de calor. Neste caso,

$$T^\mu_\nu = (\rho + P)u^\mu u_\nu + P\delta^\mu_\nu, \quad (1.11)$$

onde:

- u^μ é a quadri-velocidade - um vetor que descreve a velocidade de uma partícula de matéria em movimento no espaço-tempo, em que $u_\mu = (1,0,0,0)$, ou seja, não possui componentes espaciais (DIMOPOULOS, 2021).
- ρ e P são as densidades de energia e pressão dos fluidos (DIMOPOULOS, 2021).

As componentes $(0,0)$ e (i,j) do tensor de energia-momento são:

$$\begin{aligned} T^0_0 &= -\rho, \\ T^i_j &= P\delta^i_j. \end{aligned}$$

Utilizando as componentes $(0,0)$ e (i,i) da equação (1.9), obtemos:

$$H^2 \equiv \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{Kc^2}{a^2}, \quad (1.12)$$

$$3H^2 + 2\dot{H} = -8\pi GP - \frac{Kc^2}{a^2}. \quad (1.13)$$

A equação (1.12) é conhecida como equação de Friedmann. Ela nos mostra que o universo não é estático quando $\rho \neq 0$. A equação de Friedmann mostra que a evolução do Universo é resultado da combinação entre sua estrutura geométrica e sua composição (DIMOPOULOS, 2021).

A equação (1.13) também é umas das equações de Friedmann. Diferentemente da equação (1.12), essa equação inclui um termo adicional \dot{H} que representa a variação temporal da taxa de expansão do Universo. Ou seja, essa equação considera a aceleração cósmica.

Para simplificar os cálculos, consideramos $c = 1$. Logo as equações acima assumem as formas:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2}, \quad (1.14)$$

$$3H^2 + 2\dot{H} = -8\pi GP - \frac{K}{a^2}. \quad (1.15)$$

1.2 Equação de Continuidade

A equação (1.12) apresenta dois termos que precisam ser conhecidos para determinar a evolução do Universo. São eles o fator de escala $a(t)$ e a densidade de matéria-energia $\rho(t)$. Para encontrar as funções que relacionam essas grandezas, procedemos como segue. Sabemos que a densidade é definida como a razão entre massa e o volume espacial (DIMOPOULOS, 2021):

$$\rho = \frac{M}{V} \quad (1.16)$$

Em um espaço tridimensional, o volume cresce como $V \propto a^3$, onde $a \propto r$. Logo, derivando V em relação ao tempo, temos:

$$\dot{V} = \frac{dV}{dt} = \frac{dV}{da} \frac{da}{dt} = 3a^2\dot{a}. \quad (1.17)$$

Assim:

$$\frac{\dot{V}}{V} = \frac{3\dot{r}}{r} = \frac{3\dot{a}}{a}. \quad (1.18)$$

Para a densidade, encontramos:

$$\frac{\dot{\rho}}{\rho} = \frac{\dot{M}}{M} - \frac{\dot{V}}{V}. \quad (1.19)$$

Substituindo a equação (1.18) em (1.19), obtemos:

$$\frac{\dot{\rho}}{\rho} = \frac{\dot{M}}{M} - \frac{3\dot{a}}{a}. \quad (1.20)$$

Pela relação (1.10), temos que $\dot{a} = Ha$. Logo:

$$\frac{\dot{\rho}}{\rho} = \frac{\dot{M}}{M} - 3H. \quad (1.21)$$

Dividindo ambos os lados dessa equação por ρ e isolando o termo $\frac{\dot{M}}{M}$, chegamos à relação:

$$\frac{\dot{M}}{M} = \dot{\rho} + 3H\rho. \quad (1.22)$$

Se a massa que está se expandindo com o universo permanecer constante ($M = cte$), então $\dot{M}/M = 0$, e a equação (1.22) se torna:

$$\dot{\rho} + 3H\rho = 0. \quad (1.23)$$

No caso em que a massa do universo permanece constante, à medida que o universo se expande, a densidade de massa diminui ao longo do tempo (DIMOPOULOS, 2021). Isso porque a quantidade de matéria permanece constante, mas o volume em que está distribuída está aumentando.

Em outros processos, pode também acontecer a variação da densidade de massa. Dessa forma, utilizamos a equação abaixo, que é mais geral:

$$\dot{\rho} + 3H\left(\rho + \frac{P}{c^2}\right) = 0. \quad (1.24)$$

A equação acima é denominada de equação da continuidade, e expressa a conservação de energia local. $P(t)$ é a pressão isotrópica, que representa o trabalho feito pela expansão do universo (DIMOPOULOS, 2021), e é dada por:

$$P = -\frac{\rho c^2 \dot{M}}{3H M}. \quad (1.25)$$

1.3 Equação de Estado

A equação (1.24) introduz uma nova variável a ser conhecida, $P(t)$. Para fechar o sistema de equações que descrevem a dinâmica do Universo, utilizaremos a equação de estado na forma $P_i = P_i(\rho_i)$, o que significa que, para cada tipo de fluido barotrópico (“fluido no qual a pressão depende somente da densidade” (PARDO et al., 2013)), existe uma relação entre a pressão e a densidade, ou seja, a cada tipo de densidade está associado um tipo de pressão (DIMOPOULOS, 2021).

A equação de estado dos fluidos perfeitos cosmológicos é dada por:

$$w \equiv \frac{\sum_i P_i}{(\sum_i \rho_i)c^2}, \quad (1.26)$$

em que cada i corresponde a um tipo específico de fluido.

Para um fluido barotrópico, a velocidade do som é dada por:

$$c_s^2 = \frac{\partial \rho}{\partial P} = wc^2. \quad (1.27)$$

Para garantir que a velocidade do som não seja maior que a da luz, devemos ter $w \leq 1$. Dessa forma, valores negativos de w correspondentes a $P < 0$ podem ser considerados.

A equação de continuidade para cada fluido independente pode ser escrita da seguinte forma:

$$\dot{\rho}_i + 3(1 + w_i)H\rho_i = 0. \quad (1.28)$$

Para encontrar a relação da densidade de energia ρ_i com a o fator de escala $a(t)$, basta integrar a equação acima, mantendo w constante em relação ao tempo. Dessa forma, encontramos:

$$\rho_i \propto a^{-3(1+w_i)}. \quad (1.29)$$

1.4 Equação de Aceleração

Através das equações de Friedmann (1.12) e da continuidade (1.24), podemos chegar a uma forma mais conveniente da equação da aceleração diferente da (1.13), mas contendo as mesmas informações (DIMOPOULOS, 2021). Vejamos:

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3} \left(\rho + \frac{3P}{c^2} \right). \quad (1.30)$$

A equação de aceleração cósmica nos mostra como a taxa de expansão do universo está mudando ao longo do tempo. A equação (1.30) também pode ser escrita da seguinte forma:

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi}{3}(1 + 3w)G\rho. \quad (1.31)$$

Para que a taxa de expansão do universo aumente com o tempo, $\ddot{a}(t) > 0$, o que implica que a segunda derivada é positiva. Mas isso só é possível se a pressão for negativa o suficiente para superar a influência gravitacional, de modo que $w < -\frac{1}{3}$, onde consideramos que $a(t) > 0$ (DIMOPOULOS, 2021). Tais condições estão estritamente relacionadas à presença de energia escura que é responsável por impulsionar a expansão do universo.

1.5 Lei de Hubble

Em 1929, com base em observações, o astrônomo Edwin Hubble concluiu que as galáxias distantes estavam se afastando de nós, e a velocidade com que ocorria esse afastamento era proporcional à sua distância ($v \propto r$) (DIMOPOULOS, 2021), ou seja, quanto mais distante a galáxia, maior sua velocidade de afastamento. A partir dessas observações, chegou-se à conhecida lei de Hubble:

$$v = H_0 r, \quad (1.32)$$

onde:

- v é a velocidade de recessão.

- H_0 é a constante/parâmetro de Hubble.

O parâmetro de Hubble (H_0) é expresso da seguinte forma:

$$H_0 = 100hkm, s^{-1}Mpc^{-1} = 2,1332h \times 10^{-42}Gev. \quad (1.33)$$

O termo h representa a incerteza presente no valor de H_0 . Segundo [Dimopoulos \(2021\)](#), seu valor hoje se restringe a:

$$h = 0,674 \pm 0,005. \quad (1.34)$$

Com base na lei de Hubble, o universo não é estático; ele está se expandindo, o que nos leva a concluir que o universo atual teve origem em uma “grande explosão”, o Big Bang ([DIMOPOULOS, 2021](#)). Podemos fazer uma estimativa da idade do universo a partir do Big Bang, o que chamamos de tempo de Hubble (t_H) ([AMENDOLA; TSUJIKAWA, 2010](#)):

$$t_H = \frac{1}{H_0} = 13,7 \cdot 10^9 h^{-1}anos. \quad (1.35)$$

O fato de o universo ter uma idade finita significa que existe uma distância finita. Assim, considerando que nada se move mais rápido que a velocidade da luz (c) e introduzindo a lei de Hubble ($v \leq c$), chegamos à distância máxima observável, expressa abaixo ([DIMOPOULOS, 2021](#)):

$$D_H = \frac{c}{H_0} = 2998h^{-1}Mpc. \quad (1.36)$$

1.6 Desvio Para o Vermelho (*Redshift*)

De acordo com o que vimos até agora, temos indícios de que o universo está em expansão, o que resulta em galáxias se afastando de nós ([DIMOPOULOS, 2021](#)). Através da observação de um desvio nas frequências das linhas espectrais de galáxias distantes em comparação com seus valores observados em laboratórios terrestres ([WEINBERG, 2021](#)), obtemos mais uma evidência de que o universo está realmente se expandindo.

Esse desvio nas linhas espectrais pode ser explicado por meio do efeito Doppler nas ondas de luz. Resumidamente, o efeito Doppler explica que, se uma fonte de luz está se afastando de nós, as ondas são esticadas, fazendo com que seu comprimento de onda aumente e sua frequência diminua. Assim, as linhas espectrais são deslocadas para o vermelho, o que chamamos de *redshift*. Por outro lado, se a fonte de luz está se aproximando, o comprimento de onda diminui e a frequência aumenta. Nesse caso, as linhas espectrais sofrem um desvio para o azul, o *blueshift* ([DIMOPOULOS, 2021](#)).

O que E. Hubble observou em 1929 foi justamente o desvio para o vermelho na luz emitida por galáxias distantes, representando uma forte evidência da expansão do universo

(HASHIMOTO et al., 2018). Dessa forma, à medida que as galáxias estão se afastando de nós, ocorre o alongamento do comprimento de onda, resultando no *redshift*.

O *redshift* pode ser definido como a quantidade pela qual as linhas espectrais são deslocadas do seu comprimento de onda inicial no repouso para comprimentos de ondas maiores, devido à expansão do universo (JONES, 2017). Isso ocorre porque, à medida que a luz emitida por corpos luminosos distantes viaja pelo Universo em expansão, ela é “esticada”. O *redshift* é representado pela letra z e calculado pela equação abaixo.

$$z \equiv \frac{\lambda_0}{\lambda} - 1, \quad (1.37)$$

onde λ_0 é comprimento de onda observado e λ é o comprimento de onda emitido.

O fator de escala parametriza o crescimento do universo, logo $a \propto \lambda$, o que resulta em:

$$\frac{\lambda_0}{\lambda} = \frac{a_0}{a} \implies \frac{a_0}{a} = z + 1. \quad (1.38)$$

1.7 Modelo Cosmológico Λ CDM

O modelo *Lambda Cold Dark Matter* (Λ CDM) é o modelo padrão da cosmologia. Ele explica diversos resultados observados, desde a Radiação Cósmica de Fundo em Micro-ondas até a formação de estrutura em grande escala (LSS), (LI; KOYAMA, 2019; HELM, 2024) utilizando apenas sete parâmetros cosmológicos, sendo eles: o parâmetro de densidade de bárions físicos ($\Omega_b h^2$); o parâmetro de densidade de matéria física (Ω_m); a idade do universo (t_0); o índice espectral escalar (n_s); a amplitude de flutuação de curvatura (Δ_R^2); a profundidade óptica de reionização T e a densidade de energia escura (Ω_Λ) (HELM, 2024).

Segundo o modelo Λ CDM, a densidade de energia do Universo é composta por 5% de matéria conhecida, 25% de matéria escura, e 70% é composta pela constante cosmológica. Por meio da constante cosmológica, é possível explicar a expansão acelerada (LI; KOYAMA, 2019). E a base do modelo Λ CDM é construída sobre três hipóteses principais:

1. Em larga escala, o Universo é uniforme e isotrópico (LI; KOYAMA, 2019). Isto é também conhecido como o Princípio Cosmológico.
2. A gravidade é descrita pela Relatividade Geral (RG) em todas as escalas (LI; KOYAMA, 2019).
3. O Universo contém matéria na forma de Matéria Escura Fria (CDM), bárions e radiação (LI; KOYAMA, 2019).

No modelo cosmológico Λ CDM, o parâmetro de estado w é igual a -1 , correspondendo à constante cosmológica Λ , que está associada à energia escura do modelo (HELM, 2024). Isso significa que a energia escura se comporta como uma constante, que é responsável pela aceleração da expansão do universo.

Além do parâmetro w , o modelo Λ CDM possui outros parâmetros que estão presentes na Tabela 1. Esses parâmetros foram obtidos pela Colaboração Planck (HELM, 2024).

Tabela 1 – Parâmetros Cosmológicos

Descrição	Símbolo	Valor
Parâmetros Independentes		
Densidade bariônica física	$\Omega_b h^2$	0.02230 ± 0.00014
Densidade de matéria escura física	$\Omega_c h^2$	0.1188 ± 0.0010
Idade do universo	t_0	$13.799 \pm 0.021 \times 10^9$ anos
Índice espectral escalar	n_s	0.9667 ± 0.0040
Amplitude da flutuação de curvatura	Δ_R^2	$(2.441 + 0.088 - 0.092) \times 10^{-9}$
Profundidade óptica de reionização	T	0.066 ± 0.012
Parâmetros Fixos		
Parâmetro de densidade total	Ω_{tot}	1
Equação de estado da energia escura	w	-1
Soma das massas de três neutrinos	$\sum m_\nu$	$0.06 \text{ eV}/c^2$
Número efetivo de graus de liberdade relativísticos	N_{eff}	3.046
Amplitude escalar	A_s	2.215 ± 0.13
Razão tensor/escalar	r	0
Corrida do índice espectral	$\frac{dn_s}{d \ln k}$	0
Parâmetros Calculados		
Constante de Hubble	H_0	$67.74 \pm 0.46 \text{ km} \cdot \text{s}^{-1} \cdot \text{Mpc}^{-1}$
Parâmetro de densidade bariônica	Ω_b	0.0486 ± 0.0010
Parâmetro de densidade de matéria escura	Ω_c	0.2589 ± 0.0057
Parâmetro de densidade de matéria	Ω_m	0.3089 ± 0.0062
Parâmetro de densidade de energia escura	Ω_Λ	0.6911 ± 0.0062
Densidade crítica	ρ_{crit}	$(8.62 \pm 0.12) \times 10^{-27} \text{ kg}/\text{m}^3$
Amplitude de flutuação em 8 Mpc/h	σ_8	0.8159 ± 0.0086
Desvio para o vermelho no desacoplamento	z^*	1089.90 ± 0.23
Idade no desacoplamento	t^*	$377,700 \pm 3200$ anos
Desvio para o vermelho da reionização	z_{re}	$8.5_{-1.1}^{+1.0}$

Fonte: (HELM, 2024)

1.8 Energia Escura

Dados atuais apontam que cerca de 95% do universo é composto por um substrato invisível que só é percebível através de sua influência gravitacional. Esses 95% constituem um “setor escuro”, composto por matéria não bariônica que se divide em matéria escura e energia escura. A energia escura é uma componente exótica que possui uma pressão negativa e desempenha um papel dominante na dinâmica do Universo. Segundo a Teoria

da Relatividade Geral de Einstein, o fato de a energia escura apresentar uma pressão negativa explicaria a expansão acelerada do Universo. Por outro lado, a matéria escura é uma forma de matéria sem pressão (ZIMDAHL, 2021).

Através de observações, obtivemos várias evidências da existência de energia escura. Entre elas temos: a idade do Universo comparada às estrelas mais antigas; observações de supernovas; Radiação Cósmica de Fundo em Micro-ondas (CMB); oscilações acústicas de bárions (BAO); e estrutura em grande escala (LSS) (AMENDOLA; TSUJIKAWA, 2010).

A evidência mais convincente da existência de energia escura veio por meio de observações de supernovas do tipo Ia (SN Ia) (AMENDOLA; TSUJIKAWA, 2010). Elas serviram como “velas padrão” para medir a distância em escala do universo e, assim, permitiram descobrir que algumas supernovas estavam mais distantes do que deveriam, indicando que o universo está se expandindo a uma taxa acelerada.

1.8.1 Modelo da Energia Escura

Para começar a compreender o modelo de energia escura, temos que partir da suposição de que a densidade de energia possui uma pressão significativamente negativa, o que implica em $w = \frac{p}{\rho} < 0$. Como já foi mencionado anteriormente, para que a taxa de expansão do universo esteja em aceleração, a condição $w < -\frac{1}{3}$ precisa ser respeitada. Dessa forma, vamos começar com as equações de Friedmann e Raychaudhuri para descrever um universo que é plano do ponto de vista espacial (ESCAMILLA-RIVERA, 2016):

$$E(z)^2 = \left(\frac{H(z)}{H_0}\right)^2 = \frac{8\pi G}{3}(\rho_m + \rho_{DE}) = \Omega_{0m}(1+z)^3 + \Omega_{0(DE)}f(z), \quad (1.39)$$

e

$$\frac{\ddot{a}}{a} = -\frac{H^2}{2}[\Omega_m + \Omega_{DE}(1+3w)], \quad (1.40)$$

em que:

- O índice 0 indica o valor atual do parâmetro de Hubble H_0 e das densidades de matéria e energia Ω_0 .
- ρ_m é a densidade de matéria não relativística, expressa pela equação

$$\rho_m(z) = \rho_{0m}(1+z)^3. \quad (1.41)$$

- ρ_{DE} é a densidade de energia escura, que evolui da forma

$$\rho_{DE}(z) = \rho_{0DE}f(z). \quad (1.42)$$

- A função $f(z)$ é dada por:

$$f(z) = \exp\left[3\int_0^z \frac{1+w(z')}{1+z'}dz'\right]. \quad (1.43)$$

O modelo de energia escura, neste caso, é parametrizado pela equação de estado $w(z)$.

1.9 Supernovas do Tipo Ia (SNe Ia)

Essa classe de supernovas ocorre em sistemas estelares binários, nos quais uma das estrelas é uma anã branca que adquire matéria da estrela que a acompanha, até atingir aproximadamente a massa de 1,4 massas solares. Nesse momento, seu núcleo atinge a temperatura necessária para iniciar a fusão do carbono e oxigênio em níquel, o que rapidamente libera energia suficiente para desencadear a supernova (PADILLA et al., 2021; MOURA, 2016).

As SNe Ia se comportam como velas padrão devido à sua luminosidade e massa equivalentes. Sua luminosidade é bastante consistente, chegando a rivalizar com a galáxia na qual está localizada; por isso, são utilizadas como indicadoras de distância. (PADILLA et al., 2021; MOURA, 2016).

Podemos medir a distância de luminosidade dessas estrelas por meio da equação abaixo:

$$D_L = \sqrt{\frac{L}{4\pi S}}, \quad (1.44)$$

onde:

- L representa a luminosidade, que é a quantidade de energia emitida por unidade de ângulo sólido por segundo (PADILLA et al., 2021).
- S denota a densidade de fluxo de radiação, que é a quantidade de energia recebida por unidade de área por segundo (PADILLA et al., 2021).

Outra medida crucial na observação de supernovas é o módulo de distância padronizado (PADILLA et al., 2021), dada por

$$\mu = m_B^* - M_B + \alpha X_1 - \beta C. \quad (1.45)$$

Na equação (1.45), cada símbolo tem um significado que será descrito abaixo:

- m_B^* é a magnitude de pico observada na banda azul (B) da luz emitida pela supernova (PADILLA et al., 2021).
- M_B representa a magnitude absoluta da supernova na mesma banda azul (B), que é uma medida de sua luminosidade intrínseca (PADILLA et al., 2021).
- α, β e X_1 são parâmetros que dependem das propriedades da galáxia hospedeira da supernova (PADILLA et al., 2021).

- C é a cor da supernova no momento de seu brilho máximo, que é utilizado para corrigir medições de brilho para variações causadas por diferenças na cor das supernovas (PADILLA et al., 2021).

A relação entre as equações (1.44) e (1.45) é representada pela expressão a seguir (PADILLA et al., 2021):

$$\mu = 5 \log_{10} \left(\frac{D_L}{10 \text{ pc}} \right). \quad (1.46)$$

1.10 Análise Utilizando Dados de SNe Ia

A suposição da dominação de energia escura surgiu recentemente, com base em observações e percepção de que a taxa de expansão do universo está acelerando. Para entender esse fenômeno, os astrônomos necessitam de dados observacionais, como, por exemplo, os que estão relacionados ao desvio para o vermelho. Os dados preferíveis são aqueles que apresentam um baixo desvio para o vermelho (z), pois fornecem informações do universo relativamente recente, quando a energia escura começou a dominar. Para estudar esse fenômeno, são utilizadas duas ferramentas importantíssimas: as “velas padrão” e as “régua padrão” (ESCAMILLA-RIVERA, 2016).

As velas padrão são objetos com luminosidade intrínseca bem determinada, como as supernovas do tipo (SNe Ia). Por meio das observações, é possível determinar sua distância. As régua padrão são objetos com um tamanho intrínseco conhecido, como as Oscilações Acústicas Bariônicas (BAO). Essas ferramentas fornecem a medição da distância em função do desvio para o vermelho, permitindo que possamos adquirir informações cruciais sobre a energia escura e como ela afeta a expansão do universo (ESCAMILLA-RIVERA, 2016).

Para interpretar esses dados, é necessário cautela, o que requer uma abordagem estatística adequada para validar modelos com base nos dados observacionais. Nesse cenário, a estatística bayesiana e os algoritmos de Monte Carlo de Cadeia de Markov (MCMC) encontraram seu lugar, por serem ferramentas matemáticas e numéricas poderosas que auxiliam na estimação de parâmetros e na comparação de modelos (PADILLA et al., 2021).

2 Probabilidade e Estatística

No universo científico, existem diversos conceitos de probabilidade, mas podemos defini-la como uma medida numérica que quantifica a incerteza associada a eventos futuros. Através da teoria da probabilidade, temos algumas ferramentas matemáticas para calcular e manipular probabilidades, como a função de distribuição de probabilidade que descreve a probabilidade de cada valor possível de uma variável aleatória, e o teorema de Bayes que permite atualizar a probabilidade de um evento com base em novas informações (FASTERCAPITAL, 2024; SILVA; FERNANDES; ALMEIDA, 2015). Tais ferramentas são essenciais para a resolução de problemas de decisão e modelagem em diversos campos do conhecimento. Neste capítulo exploramos alguns princípios essenciais para compreender a estatística Bayesiana e introduzimos os algoritmos e métodos de amostragem MCMC.

2.1 Axiomas da Probabilidade

O fundamento da teoria da probabilidade pode ser enunciado nos três Axiomas de Kolmogorov (ANDREON S.; WEAVER, 2015). Considerando um evento E no espaço amostral W , definimos $P(E)$ como sendo a probabilidade do evento E ocorrer nesse espaço amostral. Dessa forma, tal valor deve satisfazer os seguintes axiomas:

1. O primeiro axioma diz respeito ao intervalo que limita as probabilidades, indicando que a probabilidade $P(E)$ de um evento E ocorrer está entre zero e um:

$$0 \leq P(E) \leq 1. \quad (2.1)$$

2. O segundo axioma nos diz que, dado o espaço amostral W , a probabilidade de se observar qualquer um dos eventos E desse conjunto é de 100%, ou seja:

$$P(W) = 1. \quad (2.2)$$

Como consequência, a probabilidade de não se observar qualquer um dos eventos desse conjunto é nula:

$$P(\emptyset) = 0. \quad (2.3)$$

3. Regra da soma: A medida de probabilidades da união finita de conjuntos é dada por:

$$P(E_1 \cup E_2 \cup \dots) = \sum_i^N P(E_i), \quad (2.4)$$

em que E_i são eventos mutuamente exclusivos, ou seja, $E_i \cap E_j = \emptyset$, para $i \neq j$. Para uma união infinita de conjunto, temos:

$$P(E_1 \cup E_2 \cup \dots) = \sum_i^\infty P(E_i). \quad (2.5)$$

Na maioria das vezes, simplificamos as coisas especificando um conjunto de hipóteses que são:

- **Mutuamente exclusivas**, o que significa que apenas uma delas pode ser verdadeira.
- **Coletivamente exaustivas**, o que significa que uma delas deve ser verdadeira (DOWNEY, 2021).

2.2 Função Distribuição de Probabilidade

Em estatística, uma distribuição é um conjunto de resultados possíveis e suas probabilidades correspondentes (DOWNEY, 2021). A função de distribuição de probabilidade (FDP) é a uma função que atribui uma probabilidade a cada possível resultado de um experimento aleatório. É uma descrição matemática de um fenômeno aleatório em termos de seu espaço amostral e das probabilidades dos eventos (STOROPOLI, 2021), ou seja, para qualquer valor de x pertencente ao conjunto de números reais (\mathbb{R}), a FDP retorna um número real entre 0 e 1, como expresso abaixo:

$$P(x) : x \rightarrow \mathbb{R} \in [0, 1]. \quad (2.6)$$

2.3 Função de Distribuição Cumulativa

A função de distribuição cumulativa (FDC) é uma função que indica a probabilidade de uma variável aleatória X ser menor ou igual a um determinado valor x (STOROPOLI, 2021). Em outras palavras, ela descreve a probabilidade acumulada de uma variável aleatória até um certo ponto. Matematicamente pode representar a FDC da seguinte forma:

$$F(x) = P(X \leq x), \quad (2.7)$$

onde:

- $F(x)$ é a função cumulativa de probabilidade de X .
- $P(X \leq x)$ é a probabilidade de que a variável aleatória X seja menor ou igual a x .

São características da FDC:

1. **Não decrescente:** À medida que x aumenta, $F(x)$ aumenta ou permanece constante.
2. **Limite à direita:** $F(x)$ tende a 1 quando x tende ao infinito.
3. **Limite à esquerda:** $F(x)$ tende a 0 quando x tende a menos infinito.

2.4 Distribuição de Probabilidade Discreta

A distribuição de probabilidade discreta descreve as probabilidades associadas a uma variável aleatória discreta (STOROPOLI, 2021).

2.4.1 Variável Discreta

Definição: Seja X uma variável aleatória. Se o número de valores possíveis de X assumir um conjunto finito ou enumerável de valores, definimos X como **variável discreta** (MEYER, 1965).

2.4.2 Função Massa de Probabilidade

Nas distribuições discretas, a atribuição de probabilidade a determinados valores é referida como “massa de probabilidade”. A função massa de probabilidade (FMP) é uma expressão que descreve a probabilidade de variável aleatória X assumir o valor x (FERREIRA, 2021), ou seja,

$$p(x) = P(X = x). \quad (2.8)$$

2.5 Principais Distribuições Discretas

A distribuição de probabilidade de uma variável mostra como as probabilidades são distribuídas entre os diferentes valores que a variável aleatória pode assumir. A seguir veremos algumas distribuições discretas.

2.5.1 Uniforme Discreta

A distribuição uniforme discreta é uma distribuição de probabilidade simétrica em que um conjunto finito de valores possíveis de uma variável aleatória são igualmente prováveis de ocorrer (FARIAS; KUBRUSLY; SOUZA, 2016). Essa distribuição é geralmente utilizada quando não há necessidade de favorecer um determinado valor, e todos os valores são igualmente prováveis. Também é empregada quando não há informação prévia sobre um problema. Pois, na ausência de informações, supõe-se que todos os valores possíveis de uma variável aleatória têm a mesma probabilidade de ocorrência.

A distribuição uniforme discreta possui dois parâmetros a e b , podendo ser expressa da seguinte forma:

$$\text{Unif}(a, b) = f(X, a, b) = \frac{1}{b - a + 1}, \quad (2.9)$$

para $a \leq X \leq b$.

2.5.2 Distribuição de Bernoulli

A distribuição de Bernoulli é utilizada para descrever um evento binário, onde há apenas dois resultados possíveis, representados por 0 e 1. Geralmente associamos 1 ao sucesso e 0 ao fracasso (STOROPOLI, 2021), em que os termos “sucesso” e “fracasso” são designações arbitrárias, podendo representar “sim” ou “não”, “ganhar” ou “perder”, “certo” ou “errado”, ou outros pares de resultados.

Essa distribuição possui apenas um parâmetro p , utilizado para modelar resultados binários discretos (FARIAS; KUBRUSLY; SOUZA, 2016), como segue:

$$\text{Bernoulli}(p) = f(x, p) = p^x(1 - p)^{1-x}. \quad (2.10)$$

O resultado “sucesso” corresponde a $x = 1$, e a probabilidade de sucesso é $f(1, p) = p$. O resultado “fracasso” corresponde a $x = 0$, e a probabilidade de fracasso é $f(0, p) = 1 - p$.

2.5.3 Distribuição Binomial

A distribuição binomial descreve a probabilidade do número de sucessos em uma sequência de n experimentos independentes. O número n é mantido fixo e cada experimento tem a mesma probabilidade de sucesso. O resultado de cada experimento segue a distribuição de Bernoulli, com dois resultados possíveis: sucesso ou fracasso (FERREIRA, 2021).

Essa distribuição é amplamente aplicada em diversas situações, fornecendo uma estrutura matemática simples e poderosa para analisar e prever eventos binários, possuindo apenas dois parâmetros n e p , a probabilidade da distribuição binomial é dada por:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad (2.11)$$

em que:

- X é o número de sucessos.
- k é o número desejado de sucessos.
- n é o número total de tentativas.
- p é a probabilidade de sucesso em uma tentativa.

E

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}. \quad (2.12)$$

2.5.4 Distribuição de Poisson

A distribuição de Poisson descreve a probabilidade de um número específico de eventos ocorrerem em um intervalo de tempo ou espaço fixo, dado um número médio de ocorrências durante esse intervalo. Ela é descrita por (FARIAS; KUBRUSLY; SOUZA, 2016):

$$P(X = k) = \text{Pois}(\lambda) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad (2.13)$$

onde:

- λ é o número médio de ocorrências no intervalo de interesse.
- k é o número específico de ocorrências que se deseja calcular a probabilidade.

O modelo de Poisson é comumente empregado em análises de filas, eventos raros de doenças, falhas de equipamentos, entre outras aplicações (FERREIRA, 2021).

2.6 Distribuição de Probabilidade Contínua

Um distribuição de probabilidade contínua descreve as probabilidades associadas a variáveis aleatórias contínuas.

2.6.1 Variável Contínua.

Variável aleatória contínua é aquela que pode assumir um número infinito de valores dentro de um intervalo específico. Por exemplo: temperatura, pressão e qualquer elemento medido numa escala contínua.

Definição: Segundo Meyer (1965), dizemos que X é uma variável aleatória contínua se existir uma função de X que satisfaça às seguintes condições:

- $f(x) \geq 0$ para todo x .
- $\int_{-\infty}^{+\infty} f(x) dx = 1$.
- Para quaisquer a, b , como $-\infty < a < b < +\infty$, teremos $P(a \leq X \leq b) = \int_a^b f(x) dx$.

Essa função é denominada **função densidade de probabilidade**.

2.6.2 Função Densidade de Probabilidade

A Função Densidade de Probabilidade é um conceito fundamental na teoria da probabilidade e estatística. Ela descreve a probabilidade de uma variável aleatória contínua estar dentro de um intervalo de valores. A Função Densidade de Probabilidade fornece

ferramentas poderosas para calcular e interpretar distribuições de variáveis contínuas, possibilitando a modelagem e análise estatística (FERREIRA, 2021).

2.7 Principais Distribuições Contínuas

A função densidade de probabilidade é fundamental para a compreensão de variáveis aleatórias contínuas em contextos estatísticos e probabilísticos. A seguir, serão apresentadas algumas das principais funções densidade de probabilidade.

2.7.1 Uniforme Contínua

A distribuição uniforme contínua é um tipo de distribuição de probabilidade simétrica, em que existe um número de valores com a mesma probabilidade de serem observados dentro de um intervalo. São características da distribuição uniforme:

- **Intervalo fixo:** A distribuição uniforme contínua é definida em um intervalo fixo de valores reais.
- **Probabilidade constante:** A probabilidade de um evento ocorrer em qualquer subintervalo de igual comprimento dentro do intervalo fixo é a mesma.
- **Função densidade de probabilidade:** A função densidade de probabilidade associada à distribuição uniforme contínua é dada por (FARIAS; KUBRUSLY; SOUZA, 2016):

$$\text{Unif}(a, b) = f(X, a, b) = \begin{cases} \frac{1}{b-a}, & \text{se } x \in [a, b]; \\ 0, & \text{caso contrário.} \end{cases} \quad (2.14)$$

2.7.2 Distribuição Normal

A função densidade de probabilidade normal é amplamente utilizada para representar variáveis contínuas cujas distribuições não são conhecidas. Além disso, ela é uma das distribuições contínuas mais utilizadas devido às suas diversas aplicações em modelagem e inferência estatística (FERREIRA, 2021).

A distribuição normal possui dois parâmetros: a média da distribuição (μ) e o desvio padrão (σ^2). É escrita da seguinte forma:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (2.15)$$

2.7.3 Distribuição Exponencial

A função densidade de probabilidade exponencial é utilizada para representar a distribuição de intervalos de tempo entre eventos que ocorrem de forma contínua e independente a uma taxa média constante (FERREIRA, 2021).

Possuindo apenas um parâmetro λ , a função de distribuição exponencial é dada por:

$$\text{EXP}(\lambda) = f(X, \lambda) = \lambda e^{-\lambda X}, \quad (2.16)$$

para $X > 0$.

2.8 Probabilidade Condicional

A probabilidade condicional é uma probabilidade que está sujeita a uma condição (DOWNEY, 2021). Em outras palavras, se temos dois eventos, A e B , e queremos a probabilidade de A ocorrer dado que B já ocorreu, estamos calculando uma probabilidade condicional. A condição, portanto, é o evento B já ter ocorrido.

Para calcular a probabilidade condicional, utilizamos a seguinte equação:

$$P(A|B) = \frac{P(A, B)}{P(B)}, \quad (2.17)$$

onde:

- $P(A|B)$ é a probabilidade condicional de A dado que B ocorreu. O sinal $|$ significa “dado que”.
- $P(A, B)$ é a probabilidade de ambos os eventos A e B ocorrerem.
- $P(B)$ é a probabilidade de B ocorrer.

A probabilidade condicional é não comutativa, ou seja:

$$P(A|B) \neq P(B|A). \quad (2.18)$$

2.9 Probabilidade Conjunta

A probabilidade conjunta é a probabilidade de que dois eventos ocorram simultaneamente (DOWNEY, 2021). Ou seja, ela representa a chance de que ambos os eventos A e B ocorram ao mesmo tempo.

A fórmula para a probabilidade conjunta é a seguinte:

$$P(A, B) = P(B)P(A|B), \quad (2.19)$$

onde:

- $P(A, B)$ é a probabilidade conjunta de A e B .
- $P(B)$ é a probabilidade do evento B ocorrer.
- $P(A|B)$ é a probabilidade de A ocorrer dado que B já ocorreu.

A probabilidade conjunta é comutativa. Significa que:

$$P(A, B) = P(B, A). \quad (2.20)$$

2.10 Estatística Bayesiana

A introdução da abordagem Bayesiana a problemas de inferência estatística teve início em 1763, quando Richard Price publicou a obra póstuma do Rev. Thomas Bayes, intitulada *An Essay Towards Solving a Problem in the Doctrine of Chances*, (PAULINO, 2018). Posteriormente, Pierre Laplace redescobriu o teorema no final do século XVIII na França, mesmo sem conhecimento da existência do teorema de Bayes, ao utilizar a probabilidade para inferência estatística com dados do censo napoleônico. Assim, o teorema ficou mais conhecido como teorema de Bayes, por ter sido formulado inicialmente por Thomas Bayes.

Atualmente, a estatística Bayesiana vem se tornando cada vez mais popular no tratamento de dados estatísticos, pois é uma ferramenta importante para modelagem e previsão estatística. Segundo Kruschke (2011), o século XXI está se tornando Bayesiano. Isso se deve à popularização de métodos computacionais muito poderosos, tornando o processo de modelagem mais fácil e preciso, o que proporcionou mudanças significativas na maneira como as estatísticas Bayesianas eram utilizadas anos atrás. Atualmente, a estatística Bayesiana é aplicada através da utilização de códigos computacionais, promovendo a construção de modelos potentes e complexos (MARTIN; KUMAR; LAO, 2021).

2.10.1 Frequentista X Bayesiana

A natureza exata da probabilidade é um tema que divide a estatística em duas escolas de pensamento: a frequentista e a Bayesiana. A escola frequentista interpreta a probabilidade como a frequência com que um evento ocorre em uma série finita de repetições do mesmo experimento, definindo a probabilidade como a razão entre o número de vezes que um evento ocorre e o número total de tentativas. Por outro lado, a escola Bayesiana considera que a probabilidade reflete o grau de confiança em uma hipótese, podendo incorporar tanto informações prévias quanto os novos dados observados (MARCH, 2013).

A principal diferença entre a inferência estatística Bayesiana e os métodos estatísticos frequentistas está na natureza dos parâmetros desconhecidos. Por exemplo, na

abordagem frequentista, os parâmetros são considerados desconhecidos, mas fixos, ou seja, acredita-se que existe apenas um parâmetro verdadeiro para os dados. Na perspectiva da estatística Bayesiana, os parâmetros também são desconhecidos, mas são tratados como incertos e, portanto, podem ser expressos como uma função de probabilidade, refletindo a incerteza sobre seus possíveis valores (SCHOOT et al., 2014). Isso significa que, na abordagem Bayesiana, o parâmetro pode estar dentro de um intervalo de valores com diferentes probabilidades, o que chamamos de intervalo de credibilidade, enquanto na frequentista apenas um único valor é considerado como verdadeiro ou provável.

A capacidade da estatística Bayesiana incorporar conhecimentos prévios em suas análises permite que as informações já conhecidas sejam integradas ao processo de inferência, promovendo assim um aprimoramento da estimativa. Isso elimina a necessidade de testar repetidamente a hipótese nula, uma prática comum nos métodos frequentistas, pois a incerteza dos parâmetros é representada através de distribuições de probabilidade. Em contraste, a hipótese nula nos métodos frequentistas assume que não há diferença significativa ocorrendo (SCHOOT et al., 2014).

2.10.2 Teorema de Bayes

Partindo da equação (2.19) podemos chegar ao Teorema de Bayes. Estabelecemos que a probabilidade conjunta é comutativa, conforme representado pela relação (2.20). Portanto, podemos afirmar que:

$$P(B)P(A|B) = P(A)P(B|A). \quad (2.21)$$

Dividindo ambos os lados por $P(B)$, obtemos o Teorema de Bayes:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \quad (2.22)$$

A estatística Bayesiana é uma abordagem que utiliza o teorema de Bayes para atualizar o conhecimento prévio sobre os parâmetros de um modelo estatístico com base nos dados observados (AL, 2021). Além disso, pode-se afirmar que o que fundamenta a Estatística Bayesiana é o Teorema de Bayes e que o fundamento do Teorema de Bayes é a probabilidade condicional (DOWNEY, 2021).

Podemos reescrever a equação (2.22), substituindo o A por um parâmetro de interesse θ , que é desconhecido, e o B pelo dado observacional Y . Temos, portanto,

$$P(\theta|Y) = \frac{P(\theta)P(Y|\theta)}{P(Y)}. \quad (2.23)$$

A fórmula de Bayes, Eq. (2.23), é composta por quatro partes distintas. Cada um de seus termos recebe uma denominação de acordo com o que representa. Vejamos:

- O primeiro termo, $P(\theta)$, é denominado **a priori**. Esta função representa nossa crença prévia.
- O segundo termo, $P(Y|\theta)$, é denominado **verossimilhança**; é a probabilidade Y condicionada a θ . Descreve a probabilidade de os dados serem descritos pelo modelo teórico construído sobre o parâmetro (ou conjunto de parâmetros) θ .
- O terceiro termo, $P(\theta|Y)$, é chamado de probabilidade **posterior**. É a quantidade na qual estamos mais interessados. Representa a atualização do conhecimento sobre o parâmetro (ou conjunto de parâmetros) θ depois de conhecidos os dados Y .
- O quarto termo, $P(Y)$, é denominado **evidência**, e pode ser obtido pela integral:

$$P(Y) = \int P(\theta)P(Y|\theta)d\theta, \quad (2.24)$$

tomada sobre todo o espaço de parâmetros.

A evidencia é considera uma constante de normalização e calcula-lá não é uma tarefa fácil. Por isso, em alguns casos, ela pode ser omitida ou cancelada do cálculo, de forma que o teorema de Bayes pode ser representado como proporcionalidade (MARTIN; KUMAR; LAO, 2021):

$$P(\theta|Y) \propto P(\theta)P(Y|\theta). \quad (2.25)$$

Uma característica da probabilidade posterior é que sempre assume uma distribuição, o que possibilita inferir afirmações sobre os parâmetros. Além disso, é válido considerar que a distribuição posterior resulta da integração de um modelo com os dados, garantindo que as declarações probabilísticas derivadas deles sejam matematicamente consistentes. No entanto, para o mundo real, devemos levar em consideração que nossos resultados estão condicionados tanto aos dados quanto ao modelo (MARTIN; KUMAR; LAO, 2021). Dessa forma, podemos calcular a probabilidade de um conjunto desconhecido de parâmetros $\theta = \theta_1, \dots, \theta_i, \dots, \theta_n$, conhecidos os dados Y . A equação (2.25) pode ainda ser escrita de maneira condicional a um modelo particular, M_j , da forma:

$$P(\theta|Y, M_j) \propto P(\theta, M_j)P(Y|\theta, M_j). \quad (2.26)$$

Na estatística bayesiana, um modelo específico M_j é definido como uma função particular, com um certo número de parâmetros livres, e com as probabilidades a priori desses parâmetros.

2.11 A Priori

A priori refere-se à distribuição de probabilidade inicial atribuída a um parâmetro ou hipótese antes dos dados experimentais serem observados (SCHOOT et al., 2014). Ou

seja, ela representa a informação prévia que temos sobre o parâmetro antes de qualquer evidência ser considerada. Para melhor entendimento, vamos tratar de um problema comum no contexto da Estatística Bayesiana, conhecido como o problema dos biscoitos.

O problema dos biscoitos

Suponha que temos duas tigelas de biscoito com as seguintes características:

- A tigela 1 contém 30 biscoitos de baunilha e 10 de chocolate.
- A tigela 2 contém 20 biscoitos de baunilha e 20 de chocolate.

Agora suponha que você escolha uma das tigelas aleatoriamente e, sem olhar, escolha um biscoito aleatoriamente. Se o biscoito for de baunilha, qual é a probabilidade de que ele tenha vindo da tigela 1 ([DOWNEY, 2021](#))?

Neste caso, nossa distribuição a priori seria a probabilidade de escolher a tigela 1 e, como há apenas duas tigelas e só podemos escolher uma delas, nossa probabilidade seria $1/2$. Percebe-se que é a probabilidade sem considerar os dados, pois não levamos em consideração a quantidade de biscoitos em ambas as tigelas, algo que é considerado pela verossimilhança, conforme veremos na próxima seção. A priori representa a probabilidade primária antes de serem observados os dados. Daí vem o nome “a priori”.

Existem diferentes distribuições a priori que podem ser baseadas em conhecimentos prévios, experiências passadas, opiniões individuais ou até mesmo uma escolha subjetiva. Isso dependerá do contexto do problema e da quantidade de informação disponível. Com base nisso, podemos escolher a distribuição a priori que melhor se adéqua ao problema em questão.

A distribuição a priori é uma parte essencial do Teorema de Bayes, pois fornece uma base inicial para a inferência estatística antes de considerar os dados observados. Posteriormente, essa distribuição é atualizada com base nos dados experimentais ([MAIOLI, 2014](#)). O uso de conhecimento prévio relevante permite que essas informações influenciem positivamente o resultado da inferência estatística, o que pode ser muito útil em situações com dados limitados. No entanto, é importante notar que a escolha da distribuição a priori pode afetar os resultados da inferência e, portanto, deve ser feita com cuidado.

2.11.1 Tipos de Prioris

As distribuições a priori são extremamente importantes para inferência estatística Bayesiana. Elas refletem as crenças iniciais sobre os parâmetros antes de analisar os dados observados. Segundo [Gelman et al. \(1995\)](#), de modo geral, podemos ter três tipos de distribuições a priori em uma abordagem Bayesiana. São elas:

- Distribuições a priori não informativas;
- Distribuições a priori fracamente informativas;
- Distribuições a priori informativas.

2.11.1.1 Distribuições a Priori Não Informativas

As priori não informativas são utilizadas quando não se têm conhecimentos prévios sobre os parâmetros e geralmente são expressas por distribuições uniformes, as quais atribuem a mesma probabilidade a todos os valores possíveis do parâmetro dentro de um intervalo. A priori não informativa parte da premissa de que “tudo é possível”; não há limites ou restrições para a crença prévia. Assim, a priori não informativa tem pouca influência no resultado da inferência, permitindo que os dados predominem (ALVES, 2019).

2.11.1.2 Distribuições a Priori Fracamente Informativas

Nas priori fracamente informativas, não partimos da premissa de que “tudo é possível”. São consideradas algumas informações disponíveis sobre o parâmetro, porém essas informações ainda são limitadas. Para essa priori, podemos ter distribuições que atribuem uma maior probabilidade a certos valores do parâmetro com base em estimativas vagas, podendo ser utilizadas distribuições normal e exponencial (BATISTA, 2021). Porém, mesmo com um pouco de conhecimento prévio, ainda assim, os dados têm mais influência na inferência.

2.11.1.3 Distribuições a Priori Informativas

Diferentemente das distribuições a priori não informativas ou fracamente informativas, que são mais neutras e flexíveis, as distribuições a priori informativas fornecem crenças ou conhecimentos prévios sólidos e detalhados sobre os parâmetros, podendo influenciar no resultado final da inferência Bayesiana. Elas são frequentemente utilizadas quando existe abundância de informações relevantes sobre o parâmetro, e é comum utilizar a distribuição normal para expressá-las. Nessa distribuição, com base no conhecimento disponível, atribuímos probabilidades diferentes aos parâmetros.

Em resumo, as distribuições a priori não informativas são mais neutras e possibilitam que os dados controlem a inferência, as distribuições a priori fracamente informativas fornecem pouca informação inicial, enquanto as distribuições a priori informativas fornecem informações detalhadas e específicas sobre o parâmetro antes de analisar os dados. Vale salientar que a escolha entre esses tipos de distribuições depende do contexto do problema, da quantidade de informação disponível e do objetivo da inferência estatística (SCHOOT et al., 2014).

2.12 Verossimilhança

A verossimilhança é um conceito fundamental presente na maioria dos métodos estatísticos. Ela é a base para o método de máxima verossimilhança na estatística frequentista e fundamental para a análise da estatística Bayesiana.

A verossimilhança é um conceito um tanto peculiar, pois não é uma probabilidade, mas é proporcional a uma. Isso porque ela não segue alguns axiomas da probabilidade, como, por exemplo, a soma das verossimilhanças não necessariamente precisa ser igual a 1. Para estimar a verossimilhança de uma hipótese θ descrever os dados Y , calculamos a probabilidade de obter Y dado θ , multiplicados por uma constante positiva arbitrária K (ETZ, 2018):

$$L(\theta) = KP(Y|\theta). \quad (2.27)$$

A constante K garante a normalização de $L(\theta)$.

Na maioria dos casos, uma hipótese representa um valor de um parâmetro em um modelo estatístico, como a média de uma distribuição normal.

Na perspectiva Bayesiana, a função de verossimilhança é usada para atualizar a distribuição a priori para uma distribuição a posteriori. Ela representa a informação obtida a partir de novos dados, sendo proporcional à probabilidade de os dados serem observados, assumindo que uma hipótese proposta é verdadeira, ou seja, $P(Y|\theta)$ (ETZ, 2018). Refletindo sobre o problema dos biscoitos definido na seção anterior, como poderíamos representar a verossimilhança?

De acordo com o nosso problema, estamos considerando a hipótese de escolher a tigela 1, na qual há 30 biscoitos de baunilha e 10 de chocolate, totalizando 40 biscoitos. A probabilidade de tirar um biscoito de baunilha na tigela 1 é a quantidade de biscoitos de baunilha dividida pelo total. Logo a verossimilhança será proporcional a $P(Y|\theta) = 3/4$, que, juntamente com a priori, resultará em nossa distribuição a posteriori.

A verossimilhança também forma a base do fator de Bayes, uma ferramenta para conduzir testes de hipóteses Bayesianos (ETZ, 2018), que será definida a seguir.

2.12.1 Método de Máxima Verossimilhança

A estimação de máxima verossimilhança é um método frequentista utilizado para encontrar o valor do parâmetro que maximiza a função de verossimilhança, com o objetivo de determinar a distribuição de probabilidade que torna os dados observados mais prováveis (MYUNG, 2003). Resumidamente, esse método procura os parâmetros que tornam os dados observados mais prováveis de acordo com o modelo estatístico.

Na análise Bayesiana, é possível interpretar estimativas pontuais clássicas obtidas através do método de máxima verossimilhança. Para amostras grandes, esse método fornece

praticamente todas as informações disponíveis sobre o parâmetro de interesse (GELMAN et al., 1995). Dessa forma, a máxima verossimilhança pode ser utilizada para estimar os melhores parâmetros para o modelo.

2.13 Distribuição a Posteriori

De uma perspectiva Bayesiana, parece natural incorporar o que foi aprendido até agora em nossa análise. Dessa forma, com a priori e a verossimilhança em mãos, combinamo-las através do teorema de Bayes para gerar nossa distribuição a posteriori. Assim, os conhecimentos prévios são atualizados através dos dados (verossimilhança) para produzir uma priori atualizada que chamamos de posteriori (SCHOOT et al., 2014).

No processo de inferência Bayesiana, há uma transição da distribuição a priori $P(\theta)$ para a distribuição a posteriori $P(\theta|Y)$. É razoável antecipar que algumas relações gerais possam existir entre essas duas distribuições. Por exemplo, é esperado que, como a distribuição posteriori incorpora as informações dos dados, ela seja menos variável do que a distribuição a priori (GELMAN et al., 1995).

Retornando ao nosso problema dos biscoitos, a partir do teorema de Bayes, combinando a priori $P(\theta)$ e a verossimilhança $P(Y|\theta)$, obtemos a distribuição a posteriori $P(\theta|Y)$. Temos que: $P(\theta) = 1/2$ e $P(Y|\theta) = 3/4$. Dessa maneira, nossa a posteriori não normalizada será:

$$P(\theta|Y) = 1/2 \cdot 3/4 = 0,375. \quad (2.28)$$

Se somássemos as posteriores não normalizadas, o resultado não seria igual a 1, o que não satisfaria a propriedade da probabilidade de que a soma das probabilidades de todos os eventos deve resultar em 1. Dessa forma, utilizamos o processo de normalização para ajustar os valores dessas probabilidades entre 0 e 1, facilitando a comparação e análise dos resultados. Normalizando este resultado, obtemos $P(\theta|Y) = 0,6$, ou seja, a chance de tirar uma biscoito de baunilha da tigela 1 é de 60%.

Este nosso problema dos biscoitos é relativamente simples; no entanto, geralmente nos deparamos com problemas mais complexos. Na maioria dos casos, obtemos a distribuição posterior por meio de simulação computacional, utilizando os chamados métodos de Monte Carlo de Cadeias de Markov (MCMC).

Em resumo, os três componentes mencionados constituem o teorema de Bayes. Este teorema afirma que nossa compreensão atualizada dos parâmetros de interesse, conhecidos nossos dados atuais, depende do equilíbrio entre nosso conhecimento prévio sobre os parâmetros (expresso pela distribuição a priori) e a evidência atual sobre esses parâmetros de interesse (expressa pela verossimilhança). Em outras palavras, o teorema de Bayes nos permite atualizar nossas crenças ou estimativas sobre os parâmetros de um modelo estatístico com base em novas informações observadas nos dados (SCHOOT et al., 2014).

2.14 Prioris Conjugadas

Como foi expresso em seções anteriores, existem diversas distribuições e diferentes tipos de priori, mas como saber quais distribuições a priori e de verossimilhança escolher para um determinado parâmetro, de modo que sua forma matemática facilite os cálculos, como, por exemplo, os cálculos de integrais?

Na verdade, existem alguns pares de distribuições cujo produto pode ser integrado diretamente de forma analítica. Essas distribuições podem servir como distribuições de probabilidades a priori e de verossimilhança (JONES, 2017), conforme mostra a Tabela 2.

Tabela 2 – Distribuições a Priori Conjugadas.

Distribuições	Parâmetros	Prior Conjugada
Binomial	p [sucesso]	Beta
Multinomial	p [sucesso]	Dirichlet
Poisson	Média	Gamma
Exponencial	1/Média	Gamma
Normal	Média μ (dada a variância)	Normal
Normal	Variância σ^2 (dada a média)	Gamma Inversa

Fonte: (JONES, 2017).

Essas distribuições são denominadas de prioris conjuntas. Podem ser integradas analiticamente para obter expressões de distribuições posteriores.

2.15 Regra de Bayes

Vamos supor a existência de dois eventos: μ representa a verdade de uma hipótese que um pesquisador mantém; e $-\mu$ é a negação dessa hipótese. Estes eventos juntos formam um conjunto disjunto $\{\mu, -\mu\}$, porque as hipóteses não podem ser verdadeiras simultaneamente, (VANDEKERCKHOVE, 2018). Aplicando a equação (2.23) a μ , de posse dos dados experimentais Y , obtemos:

$$P(\mu|Y) = \frac{P(\mu)P(Y|\mu)}{P(Y)}. \quad (2.29)$$

De maneira análoga, a fórmula de Bayes para $-\mu$ fornece

$$P(-\mu|Y) = \frac{P(-\mu)P(Y|-\mu)}{P(Y)}. \quad (2.30)$$

A probabilidade de uma hipótese depois de conhecer os dados é igual à probabilidade da hipótese antes de ver os dados, multiplicada pela probabilidade de os dados serem descritos se essa hipótese for verdadeira, dividida pela nossa evidência. Esta última é expressa pela probabilidade preditiva anterior dos dados observados $P(Y)$, que será definida a seguir (VANDEKERCKHOVE, 2018).

2.15.1 Probabilidade Preditiva Anterior $P(Y)$

A probabilidade preditiva anterior $P(Y)$ é a probabilidade de observar um resultado específico no experimento. Isso pode ser considerado como a média das probabilidades dos resultados previstos pelas hipóteses, ponderada pelas probabilidades anteriores de cada hipótese. Podemos obter $P(Y)$ a partir da regra da soma, adicionando as probabilidades conjuntas $P(\mu, Y)$ e $P(-\mu, Y)$. Logo:

$$P(Y) = P(\mu, Y) + P(-\mu, Y) = P(\mu)P(Y|\mu) + P(-\mu)P(Y|-\mu). \quad (2.31)$$

Agora podemos substituir a equação (2.31) em (2.29) e escrever

$$P(\mu|Y) = \frac{P(\mu)P(Y|\mu)}{P(\mu)P(Y|\mu) + P(-\mu)P(Y|-\mu)}. \quad (2.32)$$

$P(Y)$ é equivalente à soma do numerador da Regra de Bayes para todas as hipóteses concorrentes, fornecendo uma probabilidade média ponderada de observar o resultado Y , (VANDEKERCKHOVE, 2018). A equação 2.32 considera apenas uma hipótese; para k hipóteses, obtemos:

$$P(\mu_i|Y) = \frac{P(\mu_i)P(Y|\mu_i)}{\sum_{k=1}^k P(\mu_k)P(Y|\mu_k)}, \quad (2.33)$$

onde estamos considerando K hipóteses concorrentes e mutuamente exclusivas, sendo uma delas μ_i .

2.15.2 Fator de Bayes

Nas expressões (2.29) e (2.30), obtemos as probabilidades posteriores de duas hipóteses, $P(\mu|Y)$ e $P(-\mu|Y)$. A regra de Bayes nos apresenta uma aplicação interessante: ao tomar a razão dessas probabilidades a posteriori, $P(\mu|Y)/P(-\mu|Y)$, teremos o que chamamos de “odds posterior”. Ela compara e quantifica a nossa crença relativa dessas duas hipóteses.

As odds são uma forma de representar a chance ou probabilidade de um evento acontecer. Elas indicam a relação entre a probabilidade de um evento ocorrer e a probabilidade de não ocorrer (DOWNEY, 2021). Até agora, trabalhamos com o teorema de Bayes na forma de probabilidade, equação (2.23), mas podemos expressar o teorema de Bayes na forma de odds. A partir da razão das probabilidades, obtemos:

$$\frac{P(\mu|Y)}{P(-\mu|Y)} = \frac{\frac{P(\mu)P(Y|\mu)}{P(\mu)P(Y|\mu) + P(-\mu)P(Y|-\mu)}}{\frac{P(-\mu)P(Y|-\mu)}{P(\mu)P(Y|\mu) + P(-\mu)P(Y|-\mu)}}. \quad (2.34)$$

Como os denominadores se cancelam, a expressão se reduz a:

$$\frac{P(\mu|Y)}{P(-\mu|Y)} = \frac{P(\mu)}{P(-\mu)} \frac{P(Y|\mu)}{P(Y|-\mu)}. \quad (2.35)$$

Esta é a **regra de Bayes**, que diz que as odds posteriores são as odds anteriores multiplicadas pela razão de verossimilhança (DOWNEY, 2021), onde:

$$\frac{P(\mu|Y)}{P(-\mu|Y)} \text{ é nossa odds posterior;}$$

$$\frac{P(\mu)}{P(-\mu)} \text{ é nossa odds prior e}$$

$$\frac{P(Y|\mu)}{P(Y|-\mu)} \text{ é o nosso fator de Bayes.}$$

O fator de Bayes faz uma média ponderada da razão de verossimilhanças em todos os valores possíveis de μ . Uma vantagem importante do fator de Bayes é que ele pode ser usado para comparar quaisquer dois modelos, independentemente de sua forma. Assim, são feitas as seguintes análises:

- Se o fator de Bayes for maior que 1, isso significa que os dados são mais prováveis sob μ do que sob $-\mu$.
- Se o fator de Bayes for menor que 1, isso significa que os dados são menos prováveis sob μ do que sob $-\mu$.
- Se o fator de Bayes for exatamente 1, os dados são igualmente prováveis sob qualquer uma das hipóteses.

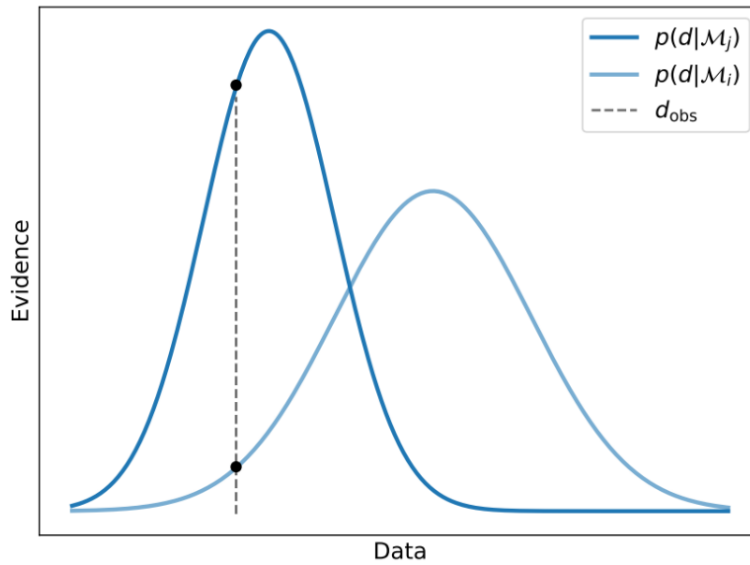
É importante diferenciar o fator de Bayes e a odds a posteriori. Ambos desempenham um papel crucial. As probabilidades a posteriori nos ajudam a determinar nossa crença geral após considerar os dados e tirar conclusões, enquanto os fatores de Bayes atuam como um indicador de aprendizado que nos informa quanto de evidência os dados forneceram.(VANDEKERCKHOVE, 2018)

A Figura 1 ilustra uma comparação entre modelos utilizando o fator de Bayes. $P(d|\mathcal{M}_i)$ e $P(d|\mathcal{M}_j)$ são as probabilidades de o dado d ser descrito pelos modelos \mathcal{M}_i e \mathcal{M}_j respectivamente. Ou seja, são as verossimilhanças para cada modelo. A linha tracejada, que intersecta ambas as distribuições, corresponde aos dados observados reais. O fator de Bayes é simplesmente a razão entre os pontos de interseção, de forma que podemos observar que \mathcal{M}_j é mais favorável sobre \mathcal{M}_i (KARAMANIS, 2023).

2.15.3 Verossimilhança Marginal

O fator de Bayes é, de fato, uma ferramenta importante para a análise de dados. Mas como calculamos as verossimilhanças? Sabemos que $P(Y|\mu)$ é a probabilidade de

Figura 1 – Representação gráfica do fator de Bayes.



Fonte: (KARAMANIS, 2023)

descrever o dado Y com base na hipótese μ . Se μ depender de um conjunto de parâmetros a serem determinados $\theta = (\theta_1, \dots, \theta_n)$ e incorporarmos esse fato à nossa notação, a hipótese parametrizada será um modelo de $\{\theta; \mu\}$ e podemos escrever $P(Y|\mu)$ da seguinte forma: (JONES, 2017)

$$P(Y|\theta) = \int P(Y|\theta, \mu)P(\theta|\mu)d\theta. \quad (2.36)$$

A equação (2.36) é chamada de verossimilhança marginal para o modelo $\{\theta; \mu\}$. Essa expressão calcula a média ponderada das verossimilhanças dos dados para todos os possíveis conjuntos de parâmetros. Uma vantagem dessa abordagem Bayesiana é que podemos incorporar diversos parâmetros ao modelo (JONES, 2017).

Quando temos apenas um ou dois parâmetros em nosso modelo, podemos visualizar a posteriori $P(\theta|Y, M)$ em um gráfico simples. Talvez não consigamos calcular analiticamente medidas importantes, como médias e variâncias, dessa distribuição, mas podemos fazer isso numericamente. No entanto, quando temos mais parâmetros, expressar a posteriori em um gráfico se torna uma tarefa difícil, sendo necessário o uso de algoritmos especiais, como os métodos de Monte Carlo de Cadeias de Markov (MCMC), para lidar com essas situações de alta dimensionalidade (JONES, 2017).

2.16 Métodos de Monte Carlo de Cadeias de Markov (MCMC)

A Cadeia de Markov Monte Carlo (MCMC) é um método estatístico e computacional de análise amplamente utilizado para extrair informações de uma distribuição, sendo frequentemente empregado na inferência Bayesiana para estimar a distribuição posterior.

Uma característica essencial do MCMC é sua capacidade de caracterizar uma distribuição sem a necessidade de conhecimento completo de todas as suas propriedades, gerando de forma aleatória valores dessa distribuição. Além disso, o MCMC é bastante útil mesmo quando temos conhecimento limitado sobre a distribuição, sendo capaz de calcular apenas a densidade para diferentes amostras dela (BROWN, 2018). Essa capacidade torna o MCMC uma ferramenta valiosa em situações onde a distribuição é complexa ou desconhecida.

O MCMC combina duas propriedades: Monte Carlo e cadeia de Markov. O método de Monte Carlo consiste na geração de um grande número de amostras aleatórias de uma distribuição. Através da média dessas amostras, são estimadas características dessa distribuição. Por exemplo, em vez de calcular diretamente a média de uma distribuição normal a partir das equações da distribuição, uma abordagem de Monte Carlo envolveria a extração de um grande número de amostras aleatórias dessa distribuição e, em seguida, o cálculo da média dessas amostras (BROWN, 2018). Isso ocorre porque calcular a média de uma amostra grande é mais fácil do que calcular diretamente a média da distribuição normal.

A propriedade da cadeia de Markov que utilizamos no MCMC é conhecida como propriedade de Markov e afirma o seguinte: a probabilidade de transição para o próximo estado depende apenas do estado atual e não dos estados anteriores. Em outras palavras, em uma sequência amostral, cada amostra depende da anterior, mas uma nova amostra não depende da anterior à anterior (BROWN, 2018) e (ANGELOTTI et al., 2008). Isso significa que o futuro é independente do passado e depende apenas do presente.

O MCMC é bastante útil na inferência Bayesiana, utilizado para aproximar distribuições posteriores, pois em muitos casos a distribuição posterior é complexa e calculá-la de maneira direta por métodos analíticos não é tão fácil; requer um processo trabalhoso e, em alguns casos, nem é possível. Dessa forma, o MCMC permite aproximar aspectos importantes da distribuição posterior, como a média, a mediana, a variância e os intervalos de credibilidade, mesmo em situações onde a forma exata da distribuição posterior é desconhecida ou intratável analiticamente (ROCCA, 2019).

2.16.1 O Método de Metropolis

O método de metropolis é uma variação específica do MCMC que é frequentemente utilizada para amostragem em inferência Bayesiana. Ele fornece simulações que baseiam-se na técnica de aceitar ou rejeitar alterações na distribuição de pontos no espaço (ANGELOTTI et al., 2008).

Na prática, o método de Metropolis é utilizado para fornecer amostras de uma distribuição de interesse quando essa distribuição é desconhecida ou quando enfrentamos dificuldade em obter amostras diretamente. O Metropolis consiste em construir uma

sequência de pontos em um espaço de parâmetros, em que cada ponto representa uma possível configuração do sistema que está sendo modelado. Para gerar um novo ponto na sequência, o algoritmo sugere uma alteração no ponto atual e depois decide se aceita ou rejeita a modificação, baseando-se na regra de aceitação.

A regra de aceitação é usada com o objetivo de garantir que a sequência de pontos gerados se aproxime ao máximo da distribuição de probabilidade desejada, de modo que os pontos que estão sendo gerados ao longo do tempo se concentrem em regiões do espaço onde a densidade de probabilidade é maior, com base na distribuição alvo.

2.16.2 Funcionamento Método de Metropolis

O funcionamento básico do método de Metropolis pode ser expresso pelos seguintes passos:

1. Criar uma amostra inicial de valores dos parâmetros (pode ser escolhida aleatoriamente).
2. Propor um novo conjunto de valores dos parâmetros com base na amostra atual.
3. Calcular a razão entre a densidade de probabilidade dos novos valores propostos e a densidade de probabilidade dos valores atuais.
4. Aceitar os novos valores propostos com uma probabilidade igual à razão calculada. Se os novos valores aumentam a densidade de probabilidade, eles são sempre aceitos. Caso contrário, eles são aceitos com uma probabilidade igual à razão de probabilidade.
5. Repetir esse processo para gerar uma sequência de amostras que, após um número suficiente de iterações, converge para a distribuição de probabilidade desejada.

2.16.3 Metropolis-Hastings

O algoritmo de Metropolis-Hastings é uma forma generalizada do algoritmo de Metrópolis ([KRUSCHKE, 2014](#)). O algoritmo Metropolis-Hastings foi proposto por Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, August M. Teller e Edward Teller em 1953, e posteriormente estendido por W.K. Hastings em 1970 ([CHIB; GREENBERG, 1995](#)).

Análogo ao método de Metropolis, o Metropolis-Hastings produz uma sequência de pontos em um espaço de parâmetros. Cada ponto representa uma configuração do sistema em questão. Porém, o Metropolis-Hastings é mais flexível, pois os pontos são propostos usando uma distribuição de probabilidade de proposta arbitrária, que não necessariamente é simétrica. Além disso, ele permite que a distribuição de proposta seja diferente da distribuição de destino ([GUNDERSEN, 2019](#)).

O algoritmo Metropolis-Hastings é operado da seguinte maneira:

1. Inicialização: Comece com um ponto inicial no espaço de parâmetros.
2. Proposta de movimento: Propõe-se um novo ponto no espaço de parâmetros com base em alguma distribuição de probabilidade de proposta.
3. Aceitação ou rejeição: Decida se o novo ponto proposto será aceito ou rejeitado com base em uma regra de aceitação.
4. Atualização: Se o novo ponto for aceito, atualize a sequência de pontos com o novo ponto. Caso contrário, mantenha o ponto atual na sequência.
5. Repetição: Repita os passos 2 a 4 até obter uma sequência suficientemente longa de pontos.

3 Estimando Parâmetros com MCMC

Neste capítulo será utilizado o método MCMC em Python para estimar parâmetros do modelo Λ CDM e de um modelo cosmológico específico de energia escura. Para isso usaremos os dados de supernovas do tipo Ia (SNe Ia) da União 2.1 e dados de $H(z)$ dos Relógios Cósmicos (CC). A seguir será comentado o passo a passo do código python utilizado, onde foi realizada uma implementação da versão do código desenvolvida por [Castillo \(2021\)](#).

3.1 Bibliotecas

A primeira coisa a fazer é importar as bibliotecas necessárias para a modelagem estatística. Isso é feito a partir das linhas apresentadas na [Figura 2](#). Cada biblioteca tem um objetivo específico que será detalhando em seguida.

Figura 2 – Importando as bibliotecas necessárias.

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3 import random as rand
4 from scipy.optimize import minimize
5 from scipy import integrate, interpolate
6 import emcee
7 import corner
8 from IPython.display import display, Math
```

Fonte: ([CASTILLO, 2021](#)).

Comentando as linhas de código

- Linha 1: A biblioteca **NumPy** fornece suporte para arrays e matrizes multidimensionais, além de funções matemáticas de alto nível para operar nesses arrays ([OLIPHANT, 2006](#)).
- Linha 2: A biblioteca **Matplotlib** serve para a visualização de dados ([SILVA, 2023](#)).
- Linha 3: Para geração de números aleatórios, a biblioteca **random** foi importada ([AWARI, 2023](#)).

- Linhas 4 e 5: Nessas linhas incorporamos algumas funções específicas da biblioteca **SciPy**. A **minimize** é usada para otimização, enquanto **integrate** e **interpolate** são usadas para integração numérica e interpolação, respectivamente (JONES et al., 2001–).
- Linha 6: A biblioteca **emcee** é uma ferramenta para realização de amostragem de Monte Carlo usando o método de Metropolis-Hastings (SALVATIER; WIECKI; FONNESBECK, 2016).
- Linha 7: A biblioteca **corner** é comumente usada para visualização de resultados de análise estatística, especialmente em análise de cadeias de Markov de Monte Carlo (MCMC).
- Linha 8: Nessa linha utilizamos algumas funções específicas de exibição do **IPython**. A **display** é usada para exibir conteúdo na célula do notebook IPython, enquanto **Math** é usada para exibir expressões matemáticas formatadas em LaTeX.

3.2 Constantes

Em seguida, definimos os valores das constantes conhecidas. O H_0 é o valor atual da constante de Hubble em km/s.Mpc. A constante c é a velocidade luz em km/s. Isso é feito conforme ilustrado na Figura 3.

Figura 3 – Atribuindo valores às constantes conhecidas

```
[ ] 1 H0 = 69 #km/s*Mpc
     2 c = 299792.458 #km/s
```

Fonte: (CASTILLO, 2021).

3.3 Conjunto de dados

Na célula da Figura 4 carregamos dois conjuntos de dados de observações astronômicas: aqueles advindos das Supernovas do tipo Ia (SCOLNIC et al., 2018; CHANG; ZHAO; ZHOU, 2019); e aqueles denominados *Cosmic Chronometer* (relógios cósmicos), que fornecem valores estimados do parâmetro de Hubble em diferentes redshifts (YU; RATRA; WANG, 2018; MORESCO, 2015).

Comentando as linhas de código

- Linhas 1 e 2: Para carregar os dados utilizamos a função **np.loadtxt**.

Figura 4 – Carregando os dados.

```

1 union2 = np.loadtxt("/content/drive/MyDrive/Union2-Supernovae_sample.txt", usecols = (1,2,3))
2 relojes = np.loadtxt("/content/drive/MyDrive/Hz_sample.txt")
3
4 z_union2 = union2[:,0]
5 mu_z_union2 = union2[:,1]
6 sigma_mu_z_union2 = union2[:,2]
7

```

Fonte: (CASTILLO, 2021).

- Linhas 1 e 2: `/content/drive/MyDrive/Union2-Supernovae_sample.txt` é o caminho para o arquivo de dados das supernovas e `/content/drive/MyDrive/Hz_sample.txt` é o caminho para o arquivo de dados dos *Cosmic Chronometer*.
- Linha 1: A função `usecols=(1,2,3)` é usada para especificar que só serão carregadas do arquivo as segunda, terceira e quarta colunas.
- Linhas 4, 5 e 6: Aqui são criadas matrizes como os valores de redshift, módulo da distância-luminosidade erros associados (desvio padrão), respectivamente, presentes na matriz `union2`.
- Linha 4: A informação `[:,0]` indica que estamos selecionando todas as linhas (`:`) da coluna 0 da matriz `union2`. O mesmo é feito nas linhas 5 e 6; a diferença é que selecionamos as segunda e terceira colunas.

No trecho do código apresentado na Figura 5, organizamos os dados de supernova de acordo com o redshift e extraímos os dados dos *cosmic chronometer*, com o intuito de facilitar a futura análise dos dados.

Figura 5 – Organizando os dados.

```

8 #Sort the data:
9 inds = z_union2.argsort()
10 mu_z_union2s = mu_z_union2[inds]
11 sigma_mu_z_union2s = sigma_mu_z_union2[inds]
12 z_union2s = np.sort(z_union2)
13
14 #These are already ordered
15 z_relojes = relojes[:,0]
16 H_relojes = relojes[:,1]
17 sigma_relojes = relojes[:,2]

```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 10: O comando `mu_z_union2[inds]` é usada para reorganizar os valores de μ de acordo com a ordem crescente do redshift z .

- Linha 11: O mesmo processo da linha 10 é realizado, desta vez para a matriz `sigma_mu_z_union2`, que contém os erros associados, de acordo com a ordem dos redshifts correspondentes.
- Linha 12: Utilizamos a função `np.sort()` para fazer uma cópia classificada dos valores de redshift das supernovas.
- Linhas 15, 16 e 17: separamos os dados da matriz de `relojes` que contém os dados dos *cosmic chronometer*, onde `[:,0]`, `[:,1]` e `[:,2]` extraem os valores das colunas 0, 1 e 2, respectivamente.

3.4 Definição de funções

É conveniente neste ponto reescrever a equação (1.39), que descreve a evolução do parâmetro de Hubble adimensional com o redshift, desta vez fazendo $\Omega_{0(DE)} = \Omega_X$:

$$E(z)^2 = \frac{H(z)^2}{H_0^2} = \Omega_{0m}(1+z)^3 + \Omega_X \cdot f(z), \quad (3.1)$$

em que Ω_{0m} é o valor atual do parâmetro de densidade de matéria, e o parâmetro de densidade de energia escura deve respeitar a seguinte restrição:

$$\Omega_X = 1 - \Omega_{0m}. \quad (3.2)$$

A função $f(z)$, cuja forma geral é dada pela equação (1.43), permite pensar em diferentes parametrizações para a evolução do parâmetro de densidade de energia escura Ω_X . A parametrização pode ser feita diretamente na função $f(z)$, como segue (CASTILLO, 2021):

$$f(z) = \left(\frac{\log(w_1)}{\log\left(1 + \frac{z}{w_1}\right)} \right)^{3(1+w_0)\log(w_1)}, \quad (3.3)$$

ou pode ser feita na equação de estado w_X , como, por exemplo, da forma:

$$w_X(z) = \left(\frac{\log(w_1)}{\log\left(1 + \frac{z}{w_1}\right)} \right) (1 - w_0) - 1. \quad (3.4)$$

Nestas equações, z é o redshift; w_0 e w_1 são parâmetros que descrevem a evolução do parâmetro de densidade de energia escura ao longo do tempo. O modelo Λ CDM considera que a equação de estado é constante ao longo do tempo. Neste modelo, essas funções devem retornar valores fixos, respectivamente, $f = 1$ e $w_X = -1$, refletindo essa propriedade do modelo.

A seguir, como mostra a Figura 6, escrevemos essas funções em linhas de código.

Figura 6 – Definindo parametrizações.

```

1 # Our DE models
2 def f(z, w0, w1):
3     return ( np.log(w1)/np.log(w1/(1+z)) )**(3*(1+w0)*np.log(w1))
4
5 def w(z, w0, w1):
6     return (1 + w0)*np.log(w1)/np.log(w1/(1+z)) - 1

```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linhas 2 e 3: Definimos uma função $f(z)$ dada pela equação (3.3).
- Linhas 5 e 6: Definimos a equação de estado $w_X(z)$ dada pela equação (3.4).

Posteriormente, definimos duas funções que modelam o comportamento do parâmetro de densidade de energia escura em um modelo cosmológico conhecido como Λ CDM (Lambda Cold Dark Matter).

Figura 7 – Quinto Bloco de Código

```

1 # LCDM
2 def f_L(z, w0, w1):
3     return 1
4
5 def w_L(z, w0, w1):
6     return -1

```

Fonte: (CASTILLO, 2021).

- Linha 2: Definimos a função f_L que recebe três parâmetros z , w_0 e w_1 , onde w_0 e w_1 são parâmetros do modelo Λ CDM. Segundo esse modelo, o parâmetro de densidade de energia escura é considerado constante e igual a -1 . Logo neste contexto w_0 e w_1 não têm efeito.
- Linha 3: A função retorna o valor de 1, para quaisquer valores de z , w_0 e w_1 .
- Linha 5: Definimos a função w_L , que também recebe três parâmetros z , w_0 e w_1 .
- Linha 6: A função retorna o valor de -1 para quaisquer valores de z , w_0 e w_1 .

Nas linhas de código da Figura 8, escrevemos o parâmetro de Hubble adimensional $E(z)$, expresso da seguinte forma:

$$E(z) = \sqrt{\Omega_m(1+z)^3 + (1 - \Omega_{0m}) \cdot f(z, w_0, w_1)}. \quad (3.5)$$

Figura 8 – Definindo o parâmetro de Hubble adimensional.

```
1 def E_inv(z, Om, w0, w1, fun):
2     return 1/np.sqrt( Om*(1+z)**3 + (1-Om)*fun(z, w0, w1) )
```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

A função E_{inv} recebe cinco parâmetros, sendo eles:

- z é o redshift.
- Om é o parâmetro de densidade de matéria Ω_{0m} que representa a fração da densidade de matéria.
- w_0, w_1 são parâmetros do modelo relacionados à evolução temporal do parâmetro de densidade de energia escura.
- fun é uma função que modela o comportamento do parâmetro de densidade de energia escura em função do redshift z .

Na sequência definimos o módulo de distância-luminosidade. Essa função é utilizada para relacionar o redshift de uma fonte com sua distância, sempre considerando a expansão do universo, ou seja, usaremos para calcular a distância de luminosidade de supernovas.

Figura 9 – Distância-luminosidade.

```
1 def mu(z, Om, w0, w1, fun):
2     return 5*np.log10(((1+z)*c*integrate.quad(E_inv, 0, z, args=(Om,
3         w0, w1, fun)))[0]/(H0*1e-5) )
```

Fonte: (CASTILLO, 2021).

O código da Figura 9 define uma função $\mu(z, \Omega_m, w_0, W_1, fun)$ que calcula o módulo da distância para cada supernova. A função é expressa da seguinte forma:

$$\mu(z) = 5 \log_{10} \left(\frac{(1+z) \cdot c}{H_0 \cdot 10^{-5}} \int_0^z \frac{1}{E(z')} dz' \right). \quad (3.6)$$

Assim, a função $\mu(z, \Omega_m, w_0, w_1, fun)$ calcula o módulo da distância usando a expansão cósmica e os parâmetros cosmológicos fornecidos.

3.5 Função de Máxima Verossimilhança

Uma maneira de estimar parâmetros é utilizando a função de máxima verossimilhança, que é uma função que busca maximizar a verossimilhança dos dados observados. Dessa forma, com os dados em mãos, podemos obter o conjunto mais provável de parâmetros do modelo. Escolhemos uma normalização gaussiana, e dessa forma, maximizar a verossimilhança é equivalente a minimizar o qui-quadrado χ^2 (CASTILLO, 2021). A verossimilhança L e qui-quadrado χ^2 estão relacionados por meio da expressão:

$$L \propto e^{-\frac{\chi^2}{2}}. \quad (3.7)$$

Estamos utilizando dois conjuntos de dados, dados de supernovas do tipo Ia (SNe Ia) da União 2.1 e dados de $H(z)$ de Relógios Cósmicos (CC). Para o primeiro conjunto de dado o qui-quadrado é dado pela expressão (CASTILLO, 2021):

$$\chi_{\text{SN}}^2 \equiv \sum_{i=1}^{N_{\text{Union2.1}}} \left[\frac{\mu(z_i; \Omega_m, w_0, w_1) - \mu_{\text{obs}}(z_i)}{\sigma_{\mu,i}} \right]^2, \quad (3.8)$$

onde:

- $N_{\text{Union2.1}}$ é o número total de pontos no conjunto de dados.
- $\mu_{\text{obs}}(z_i)$ é a módulo de distância observado para cada supernova i e $\sigma_{\mu,i}$ é o erro associado.

Para o conjunto de dados “Relógios Cósmicos” (CC), utilizaremos:

$$\chi_{\text{CC}}^2 \equiv \sum_{i=1}^{N_{\text{CC}}} \left[\frac{H(z_i; \Omega_m, w_0, w_1) - H_{\text{obs}}(z_i)}{\sigma_{H,i}} \right]^2, \quad (3.9)$$

onde:

- N_{CC} é o número total de pontos no conjunto de dados.
- $H_{\text{obs}}(z_i)$ é o parâmetro de Hubble observado em cada desvio para o vermelho z_i e $\sigma_{H,i}$ é seu erro.

O código na Figura 12 implementa uma função de verossimilhança logarítmica para o modelo cosmológico Λ CDM usando dados de supernovas (SNe).

Comentando as linhas de código

- Linha 1: Definimos a função `log_likelihood_LCDM_sne(theta, z, mu_obs, sig_mu)`.
- z é o redshift das SNe.

Figura 10 – Estimando Ω_{0m} no modelo Λ CDM.

```

1 def log_likelihood_LCDM_sne(theta, z, mu_obs, sig_mu):
2     Om = theta
3     model = []
4     for i in z:
5         model.append(mu(i, Om, 1, 1, f_L))
6
7     sigma2 = sig_mu**2
8     return -0.5 * np.sum((model - mu_obs)**2 / sigma2 + np.log(2*np.pi*sigma2))
9
10 nll_LCDM_sne = lambda *args: -log_likelihood_LCDM_sne(*args)
11 initial_LCDM_sne = np.array([0.3]) # We take as initial value Om= 0.3.
12 sol_LCDM_sne = minimize(nll_LCDM_sne, initial_LCDM_sne, args=(z_union2s,
13     mu_z_union2s, sigma_mu_z_union2s))

```

Fonte: (CASTILLO, 2021).

- mu_obs é o módulo de distância observado das SNe.
- sig_mu é o desvio padrão associado ao módulo de distância observado.
- $theta$ é o parâmetro que queremos encontrar.
- Linha 2: Definimos $theta = Om$, ou seja, $\theta = \Omega_m$.
- Linha 3: Cria uma lista vazia a qual serão adicionados os valores dos módulos de distância para cada redshift fornecido.
- Linha 4: Itera sobre todos os valores de redshift z .
- linha 5: Adiciona valores à lista “model”. Estes valores são calculados pela função “mu” definida nas linhas da Figura 9. Esta função utiliza como parâmetros Om (Ω_m) e outro parâmetro fixo $(1, 1, f_L)$, onde f_L é a função definida nas linhas da Figura 7.
- Linha 7: Calcula o quadrado do desvio padrão, que será utilizado na função abaixo do código.
- Linha 8: Calcula a negativa da log-verossimilhança, definida pela seguinte expressão:

$$\log_likelihood = -\frac{1}{2} \sum \left(\frac{[\mu(z_i, \Omega_m, \omega_0, \omega_1) - \mu_{obs}(z_i)]^2}{\sigma_{\mu,i}^2} + \log(2\pi\sigma_{\mu,i}^2) \right), \quad (3.10)$$

onde:

- μ_{obs} são valores observados do módulo de distância das supernovas.
- σ^2 é a variância associada aos valores observados μ_{obs} .
- $\log(2\pi\sigma^2)$ é uma constante de normalização.

- Linha 10: Define a função `nll_LCDM_sne` como uma função anônima (lambda) que retorna a negativa da log-verossimilhança.
- Linha 11: Define o valor a priori de Ω_{0m} como sendo 0,3.
- Linha 12: Minimiza a função de verossimilhança negativa usando o método **minimize**, considerando o valor inicial de Ω_{0m} e os dados observados fornecidos como argumentos. Após o cálculo, o resultado é armazenado na função `sol_LCDM_sne`.

O parâmetro de densidade de matéria Ω_{0m} para o modelo Λ CDM, calculado com o método da máxima verossimilhança, e usando dados de supernovas tem valor a posteriori dado pela saída apresentada na Figura 11:

Figura 11 – Resultado do Ω_{0m}

```
1 sol_LCDM_sne.x[0]
```

```
0.3152160602252687
```

Fonte: (CASTILLO, 2021).

No código da Figura 12 escrevemos a função de verossimilhança logarítmica para nosso modelo de energia escura em termos dos parâmetros ω_0 , ω_1 , Ω_m , no intuito de realizar uma análise de ajuste de parâmetros para o modelo cosmológico que representa os dados extraídos de supernovas e, assim, encontrar os melhores valores para esses parâmetros e ter uma melhor compreensão da natureza da energia escura e da expansão do universo.

Figura 12 – Definindo o log da verossimilhança.

```
1 def log_likelihood_f_sne(theta, z, mu_obs, sig_mu):
2     w0, w1, Om = theta
3     model = []
4     for i in z:
5         model.append(mu(i, Om, w0, w1, f))
6
7     sigma2 = sig_mu**2
8     return -0.5 * np.sum((model - mu_obs)**2 / sigma2 + np.log(2*np.pi*sigma2))
9
10 nll_f_sne = lambda *args: -log_likelihood_f_sne(*args)
11 initial_f_sne = np.array([-1.1, 0.5, 0.3])
12 sol_f_sne = minimize(nll_f_sne, initial_f_sne, args=(z_union2s,
13             mu_z_union2s, sigma_mu_z_union2s))
14 sol_f_sne.x
```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Define a função `log_likelihood_f_sne` que calcula a log-verossimilhança, e que tem três parâmetros: $\theta, z, \mu_{obs}, \text{sig_}\mu$. O θ , nesse caso, representa o conjunto de parâmetros do modelo: $\omega_0, \omega_1, \Omega_m$.
- Linha 2: Definição do conjunto de parâmetros `theta = ($\omega_0, \omega_1, \Omega_m$)`.
- Linha 3: Cria uma lista vazia denominada `model` que será preenchida com os valores de μ para cada redshift.
- Linha 4: Itera sobre todos os valores de redshift z , onde cada iteração i atribuirá um valor de z .
- Linha 5: Utilizamos a função `mu` definida na Figura 9, para calcular o módulo de distância referente a cada valor de z .
- Linha 7: Define a variância `sigma2`, ou seja, o quadrado do desvio padrão.
- Linha 8: Calcula a negativa da log-verossimilhança.
- Linha 10: Define a função `nll_f_sne` como função antônima que, para qualquer argumento `args` ela retorna a negativa da log-verossimilhança.
- Linha 11: Define os valores iniciais para os parâmetros $\omega_0, \omega_1, \Omega_m$.
- Linha 12 e 13: utiliza o método `minimize` para minimizar função de verossimilhança. Com os valores iniciais dos parâmetros já definidos e os dados observados fornecidos como argumento, o resultado será armazenado em `sol_f_sne`.
- Linha 14: Exibe os valores dos parâmetros que minimizam a função de verossimilhança.

Como resultado do código acima, obtemos uma matriz com os valores de w_0, w_1 e Ω_m , respectivamente: `array([-0.92878995, 0.50112008, 0.31807275])`.

Dando continuidade, no código da Figura 13, utilizamos dois modelos para calcular o módulo da distância para cada valor de redshift presente em `z_union2s`. Os modelos utilizados são o Λ CDM e modelo f que está definido nas linhas de código da Figura 8.

- Linha 1: Cria uma lista vazia `model_LCDM_sne`, em que serão adicionados os valores de módulo da distância para cada redshift.
- Linhas 2 e 3: Calcula o valor da distância módulo para i através da função `mu`, levando em consideração o parâmetro obtido pela função `sol_LCDM_sne.x[0]`, mantendo os valores fixos (1, 1, `f_L`). E acrescenta esses valores à lista `model_LCDM_sne`.

Figura 13 – Carregando dados dos modelos cosmológicos.

```

1 model_LCDM_sne = []
2 for i in z_union2s:
3     model_LCDM_sne.append(mu(i, sol_LCDM_sne.x[0], 1, 1, f_L))
4
5 model_f_sne = []
6 for i in z_union2s:
7     model_f_sne.append(mu(i, sol_f_sne.x[2], sol_f_sne.x[0],
8                          sol_f_sne.x[1], f))

```

Fonte: (CASTILLO, 2021).

- Linhas 5, 6 e 7: Realizam procedimento semelhante ao das linha 1,2 e 3. Calcula os valores do módulo de distância para o modelo f . Para isso utiliza a função **mu** com os parâmetros obtidos através da função **sol_f_sne.x**, onde **sol_f_sne.x[0]** é w_0 , **sol_f_sne.x[1]** é w_1 e **sol_f_sne.x[2]** é Ω_{0m} .

Através desse conjunto de dados, podemos plotar o gráfico do módulo da distância-luminosidade $\mu(z)$ em função do redshift z , com o intuito de comparar os parâmetros que melhor se ajustam ao conjunto de dados. Para isso, foi plotado o limite Λ CDM, com resultado de Ω_{0m} aproximadamente igual a 0,315. O código utilizado para cria o gráfico está representado na figura 14.

Figura 14 – Gerando o gráfico da distância luminosa.

```

1 plt.figure(figsize=(10,8))
2 plt.errorbar(z_union2s, mu_z_union2s, yerr=sigma_mu_z_union2s,
3             fmt=".k", capsize=0, alpha=0.2)
4 plt.plot(z_union2s, model_LCDM_sne, label=r'$\Lambda$CDM', color='y')
5 plt.plot(z_union2s, model_f_sne, label='DE model', color='b', alpha=0.6)
6
7 plt.title('Supernovae')
8 plt.xlabel(r'$z$')
9 plt.ylabel(r'$\mu(z)$')
10 plt.legend()
11 plt.show()

```

Fonte: (CASTILLO, 2021).

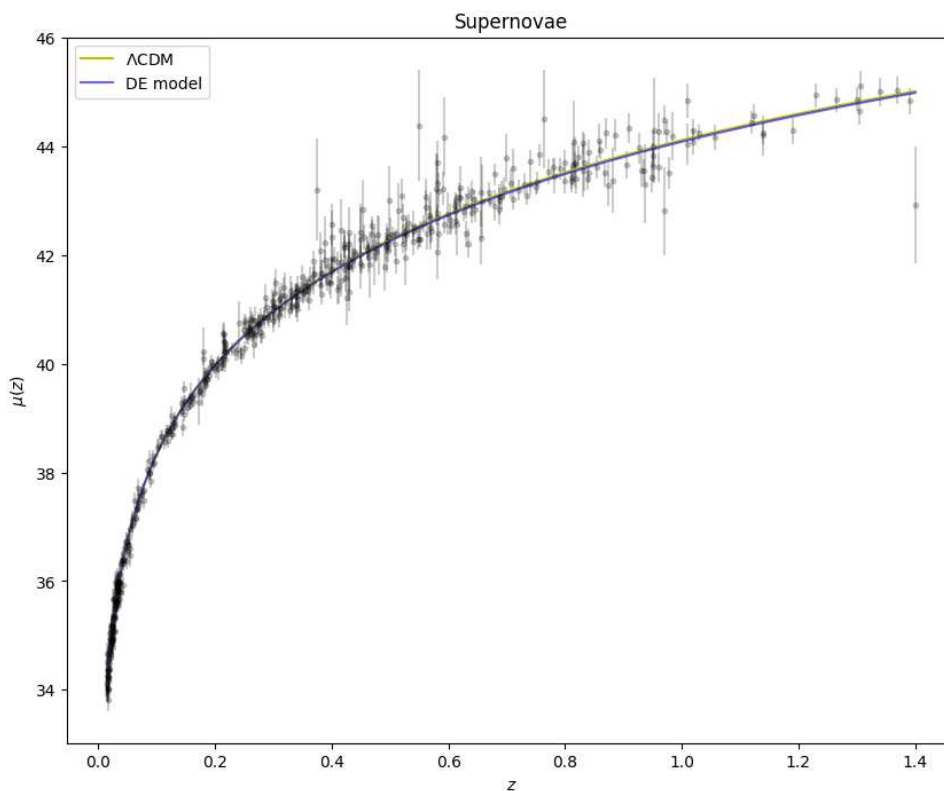
Comentando as linhas de código

- Linha 1: Cria uma para plotar o gráfico, com a seguinte dimensão 10x8 polegadas.
- Linha 2 e 3: Utilizamos **plt.errorbar()** para cria um gráfico de barra de erro.
- Linha 4 e 5: Usamos **plt.plot()** para traçar as linhas do gráfico, para os dados do modelo Λ CDM e do modelo específico de energia escura.

- Linha 7: Define o título do gráfico.
- Linha 8: Rotula o eixo x como z .
- Linha 9: Rotula o eixo y como $\mu(z)$.
- Linha 10 e 11: Adiciona a legenda e exibe o gráfico.

Como resultado obtemos a Figura 11. Através do gráfico, podemos observar um excelente ajuste do modelo $f(z)$ estudado em relação ao modelo Λ CDM. Para redshifts menores, nosso resultado se aproxima melhor do Λ CDM. No entanto, quando avançamos para redshifts maiores, podemos perceber uma pequena diferença.

Figura 15 – Distância-luminosidade.



Fonte: (CASTILLO, 2021).

3.6 Método MCMC

A seguir utilizaremos o método de Monte Carlo de Cadeias de Markov (MCMC) para inferir parâmetros de interesse para o modelo Λ CDM e nosso modelo de energia escura. Mais especificamente, utilizaremos o *sampler* **emcee**, implementado em Python, na tentativa de gerar uma cadeia de amostras no espaço de parâmetros do modelo. Essas amostras serão então usadas para estimar a distribuição a posteriori dos parâmetros.

Primeiramente, vamos definir a nossa distribuição logarítmica da priori para o modelo Λ CDM, conforme mostra as linhas de código da Figura 16.

Figura 16 – Definindo a distribuição a priori.

```
1 def log_prior_LCDM_sne(theta):
2     Om = theta
3     if 0.25 < Om < 0.35:
4         return 0.0
5     return -np.inf
```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Define a função `log_prior_LCDM_sne` que calcula o logaritmo da distribuição a priori.
- Linha 2: Atribui o valor do parâmetro Ω_{0m} ao parâmetros `theta` da função que está sendo definida..
- Linha 3: Verifica se o parâmetro Ω_{0m} está dentro do intervalo de 0,25 à 0,35, levando em consideração que no modelo Λ CDM o valor esperado para Ω_{0m} é 0,3.
- Linha 4: Retorna zero para Ω_{0m} dentro do intervalo específico. Quando fazemos isso, estamos assumindo que a densidade de probabilidade dentro do intervalo é constante e não igual a zero. Ou seja, estamos dizendo que a densidade de probabilidade é uniforme e que os valores dentro do intervalo têm a mesma probabilidade de ocorrer.
- Linha 5: Retorna menos infinito se o valor de Ω_{0m} estiver fora do intervalo, indicando que os valores fora deste intervalos são muito improváveis.

Definidas as nossas funções logarítmicas log-prior (Fig. 16) e log-verossimilhança (Fig. 12), podemos calcular o logaritmo da probabilidade a posteriori para o modelo cosmológico Λ CDM. As linhas de código com esse objetivo são apresentadas na Figura 17.

- Linha 1: Define a função `log_probability_LCDM_sne` que calcula o logaritmo da probabilidade para o modelo Λ CDM.
- Linha 2: Calcula o log-prior para os parâmetros do modelo usando a função definida acima, `log_prior_LCDM_sne`.

Figura 17 – Definindo a distribuição a posteriori.

```

1 def log_probability_LCDM_sne(theta, z, mu_obs, sig_mu):
2     lp = log_prior_LCDM_sne(theta)
3     if not np.isfinite(lp):
4         return -np.inf
5     return lp + log_likelihood_LCDM_sne(theta, z,
6         mu_obs, sig_mu)

```

Fonte: (CASTILLO, 2021).

- Linha 3: Verifica se o log-prior é finito. Caso não seja, isso indica que os parâmetros do modelo estão fora do intervalo de interesse.
- Linha 4: Retorna menos infinito ($-\infty$) para os parâmetros que estão fora do modelo, indicando uma probabilidade infinitamente pequena.
- Linha 5: Retorna a soma do logaritmo da priori e do logaritmo da verossimilhança para os parâmetros do modelo. Isso representa o logaritmo da probabilidade a posteriori, pois se o logaritmo da priori for finito, a função calculará o logaritmo da verossimilhança e retornará a soma de ambos.

Na sequência, a Figura 18 define a nossa distribuição logarítmica a priori para o modelo de energia escura estudado.

Figura 18 – Definindo a distribuição a priori.

```

1 def log_prior_f_sne(theta):
2     w0, w1, Om = theta
3     if -1.2 < w0 < -0.8 and 0.3 < w1 < 0.7 and 0.25 < Om < 0.35:
4         return 0.0
5     return -np.inf

```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Define a função `log_prior_f_sne` que calcula o logaritmo da priori para os parâmetros do modelo de energia escura.
- Linha 2: O `theta` representa o conjunto de parâmetros w_0 , w_1 e Ω_m . Essa linha desacopla esses parâmetros contido em `theta`.
- Linha 3: Verifica se os parâmetros estão dentro de intervalos aceitáveis. Esses intervalos baseiam-se nos valores obtidos pelo método de máxima verossimilhança.

- Linha 4: Assim como no código anterior, essa linha retorna zero se os parâmetros estiverem dentro dos intervalos especificados.
- linha 5: Retorna menos infinito se os parâmetros estiverem fora dos intervalos especificados.

Agora vamos calcular nossa distribuição logarítmica a posteriori, utilizando a distribuição a priori `log_prior_f_sne` (Figura 18) e a verossimilhança `log_likelihood_f_sne` (Figura 12). As linhas de código são apresentadas na Figura 19.

Figura 19 – Definindo a distribuição a posteriori.

```
1 def log_probability_f_sne(theta, z, mu_obs, sig_mu):
2     lp = log_prior_f_sne(theta)
3     if not np.isfinite(lp):
4         return -np.inf
5     return lp + log_likelihood_f_sne(theta, z, mu_obs, sig_mu)
```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Define a função `log_probability_f_sne` que será responsável por calcular o logaritmo da probabilidade para o modelo de energia escura.
- Linha 2: Chama a função `log_prior_f_sne` que calcula o logaritmo da priori, de `lp` para depois utilizar na função.
- Linhas 3 e 4: Verifica se o logaritmo da priori não é finito, pois, se for infinito, os parâmetros do modelo estão fora do intervalo de interesse definido pela função `log_prior_f_sne` (Figura 18).
- Linha 5: Retorna a soma do logaritmo da priori e do logaritmo da verossimilhança para os parâmetros do modelo, ou seja, calcula o logaritmo da posteriori.

Na próxima célula (Figura 20), vamos implementar o método MCMC usando a biblioteca `emcee` para o modelo Λ CDM com 32 *walkers*. *Walkers* são essencialmente “caminhantes” através do espaço de parâmetros, onde cada “walker” representa um “salto” de um estado a outro na construção de uma cadeia de Markov independente que usamos para explorar o espaço de parâmetros (REGNIER et al., 2023).

Figura 20 – Implementando um cálculo de MCMC.

```

1 pos_LCDM_sne = sol_LCDM_sne.x + 1e-4 * np.random.randn(32, 1)
2 nwalkers_LCDM_sne, ndim_LCDM_sne = pos_LCDM_sne.shape
3
4 sampler_LCDM_sne = emcee.EnsembleSampler(nwalkers_LCDM_sne, ndim_LCDM_sne,
5 log_probability_LCDM_sne, args=(z_union2s, mu_z_union2s, sigma_mu_z_union2s))
6
7 sampler_LCDM_sne.run_mcmc(pos_LCDM_sne, 500, progress=True);

```

100% ██████████ 500/500 [40:53<00:00, 4.91s/it]

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Gera um conjunto inicial de estados da cadeia e armazena como `pos_LCDM_sne`, onde:
 - `sol_LCDM_sne.x` é o valor do parâmetro Ω_{0m} encontrado anteriormente (Figura 11).
 - `1e-4` que é equivalente a 10^{-4} , e representa um pequeno valor que adicionamos aos deslocamentos para garantir que a cadeia nunca inicie do mesmo lugar.
 - `np.random.randn(32, 1)` gera uma matriz com 32 linhas e 1 coluna, como os números aleatoriamente distribuídos em uma distribuição normal padrão.
- Linha 2: Define um conjunto bidimensional de parâmetros, formado pelo número de *walkers* e um dado número de dimensões.
- Linha 4 e 5: Cria um objeto `EnsembleSampler` da biblioteca `emcee` (SALVATIER; WIECKI; FONNESBECK, 2016), que realizará a amostragem MCMC, onde:
 - `nwalkers_LCDM_sne` é o número de caminhantes.
 - `log_probability_LCDM_sne` é a função que calcula o logaritmo da posterior expressa na Figura 17.
 - `args=(z_union2s, mu_z_union2s, sigma_mu_z_union2s)` são os argumentos que serão adicionados.
- Linha 7: Realiza a amostragem MCMC utilizando o método `run_mcmc`.
 - `pos_LCDM_sne` são as posições iniciais dos caminhantes.
 - `500` é o número de passos a serem realizados pelo MCMC.
 - `progress=True` mostrar o progresso da amostragem a medida que ela está sendo executada.

As linhas de código seguintes (Figura 21) extraem e formatam as amostras geradas no código acima, de maneira que todas as cadeias de Markov serão organizadas em uma lista unidimensional.

Figura 21 – Organizando as amostras.

```
1 flat_samples_LCDM_sne = sampler_LCDM_sne.get_chain(discard=5,
2           thin=1, flat=True)
3
4 print(flat_samples_LCDM_sne.shape)

(15840, 1)
```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linhas 1 e 2: Utilizamos a função `get_chain()` a `sampler_LCDM_sne` para extrair as cadeias de Markov geradas durante a amostragem MCMC.
 - `discard=5` é usado para descartar os primeiros passos. Fazemos isso para que a cadeia possa atingir a distribuição estacionária antes que as amostras sejam consideradas válidas.
 - `thin=1` indica que todas as amostras devem ser mantidas, pois o `emcee` sempre salta amostras para reduzir a autocorrelação entre amostras.
 - `flat=True` indica que serão retornadas em uma matriz unidimensional.
- Linha 4: Imprime o formato do dessa matriz, indicando o número de linhas e colunas. Como podemos observar, nossa `array flat_samples_LCDM_sne` tem **15840** linhas e **1** coluna.

Através dessas amostras podemos gerar o grafia de contorno dessa distribuição por meio do código abaixo(Figura 22):

Figura 22 – Gerando o gráfico das distribuições.

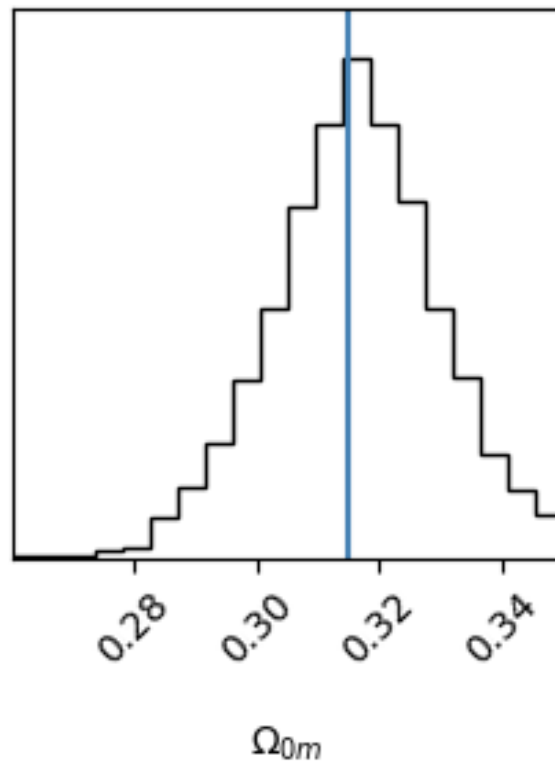
```
1 fig_LCDM_sne = corner.corner(flat_samples_LCDM_sne,
2   labels=[r'$\Omega_{0m}$'], truths=[0.315]);
```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1 e 2: Utilizamos a função `corner.corner(...)` para gerar um gráfico de contorno, onde `flat_samples_LCDM_sne` é o conjunto de dados utilizados; `labels=[r'Ω0m']` rotula o gráfico e `truths=[0.315]` representa o valor do parâmetro atual.

Figura 23 – Distribuição a posterior para Ω_{0m} durante a amostragem MCMC.



Fonte: (CASTILLO, 2021).

No gráfico da Figura 23, podemos visualizar a distribuição a posteriori para o parâmetro Ω_{0m} do modelo Λ CDM. A linha em azul indica o seu valor esperado com base em dados de observações realizadas pelo satélite Planck em 2018. A linha azul indica o valor mais provável, ou seja, o valor esperado para Ω_{0m} (CASTILLO, 2021). Nosso cálculo converge para o valor mais aceito de Ω_{0m} atualmente, com um pequeno grau de incerteza associado, comprovando que o modelo Λ CDM se ajusta bem aos dados observados. Os valores detalhados de Ω_{0m} e suas incertezas são calculados nas linhas de código da Figura 24.

Figura 24 – Estimativa do valor de Ω_{0m} do modelo Λ CDM.

```

1 labels_LCDM_sne = [r'\Omega_{0m}']
2 results_LCDM_sne = []
3 for i in range(ndim_LCDM_sne):
4     mcmc_LCDM_sne = np.percentile(flat_samples_LCDM_sne[:, i], [16, 50, 84])
5     q_LCDM_sne = np.diff(mcmc_LCDM_sne)
6     txt_LCDM_sne = "\mathrm{{{3}}} = {0:.3f}_{-{1:.3f}}^{{{2:.3f}}}"
7     txt_LCDM_sne = txt_LCDM_sne.format(mcmc_LCDM_sne[1], q_LCDM_sne[0],
8         q_LCDM_sne[1], labels_LCDM_sne[i])
9     display(Math(txt_LCDM_sne))
10    results_LCDM_sne.append(mcmc_LCDM_sne[1])

```

$$\Omega_{0m} = 0.316_{-0.013}^{0.013}$$

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Define uma lista `labels_LCDM_sne` que contém o rótulo do parâmetro Ω_{0m} .
- Linha 2: Inicia uma lista vazia que será utilizada para armazenar os resultados.
- Linha 3: Inicia um *loop* que será responsável por iterar todas as amostras obtidas durante a amostragem MCMC.
- Linha 4: calcula o percentis 16, 50 e 84 para amostras dos i -ésimo parâmetro do conjunto `flat_samples_LCDM_sne`, fornecendo um intervalo de credibilidade de 68% para o parâmetro Ω_{0m} .
- Linha 5: Calcula a diferença entre os percentis calculados anteriormente, ou seja, os limites do intervalo de credibilidade.
- Linha 6: Define uma *string* de formato para formatar os resultados, de maneira que o texto que está dentro da chave será lido em LaTeX.
- Linha 7: Formata a *string* de resultado com os valores calculados e os rótulos dos parâmetros.
- Linha 8: Exibe os resultados que foram formatados utilizando a função `display()`.
- Linha 9: Adiciona o valor médio de Ω_{0m} aos resultados finais que foram armazenados na lista `results_LCDM_sne`.
- Resultado: Como resultado obtemos o valor médio de $\Omega_m = 0,315$ e um intervalo de credibilidade tal que $0,315 - 0,014 \leq \Omega_{0m} \leq 0,315 + 0,014$.

Em seguida, utilizamos o código descrito na Figura 25 para executar uma amostragem, através do método MCMC, para estimar os parâmetros w_0, w_1 e Ω_m do modelo de energia escura. Esse código é semelhante ao código da Figura 20.

Figura 25 – Implementando um cálculo de MCMC.

```

1 pos_f_sne = sol_f_sne.x + 1e-4 * np.random.randn(32, 3)
2 nwalkers_f_sne, ndim_f_sne = pos_f_sne.shape
3
4 sampler_f_sne = emcee.EnsembleSampler(nwalkers_f_sne, ndim_f_sne,
5   log_probability_f_sne, args=(z_union2s, mu_z_union2s, sigma_mu_z_union2s))
6 sampler_f_sne.run_mcmc(pos_f_sne, 500, progress=True);

100%|██████████| 500/500 [12:45<00:00, 1.53s/it]

```

Fonte: (CASTILLO, 2021).

Comentando as linhas de código

- Linha 1: Primeiramente geramos um conjunto de inicial de valores para o MCMC, utilizando os resultados dos parâmetros obtidos na função `sol_f_sne.x`. Uma matriz com os números aleatórios de uma distribuição normal, com 32 linhas e 3 colunas é gerada com `np.random.randn(32, 3)`.
- Linha 2: Define o número de “saltos” dentro da cadeia e o número de dimensões a serem amostrados a partir da função inicial `pos_f_sne.shape`.
- Linha 3: cria uma objeto `EnsembleSampler` que será responsável pela amostra de MCMC, onde:
 - `nwalkers_f_sne` é o numero de “saltos” na cadeia.
 - `ndim_f_sne` é o numero parâmetros.
 - `log_probability_f_sne` é uma função que calcula o logaritmo da probabilidade a posteriori.
 - `args=(z_union2s, mu_z_union2s, sigma_mu_z_union2s)` são argumentos adicionais.
- Linha 4: Executa a amostragem MCMC usando o método `run_mcmc`, em que:
 - `pos_f_sne` indica as posições iniciais da cadeia.
 - `500` é o número de passos a serem realizados pelo MCMC.
 - `progress=True` Exibe o progresso da amostragem enquanto ela está sendo executada.

Posteriormente, utilizamos o código descrito na Figura 26 para extrair e formatar as amostras obtidas durante a amostragem MCMC, e expressá-las como matriz, indicando o número linhas e colunas.

Figura 26 – Organizando as amostras.

```
1 flat_samples_f_sne = sampler_f_sne.get_chain(discard=5,
2       thin=1, flat=True)
3 print(flat_samples_f_sne.shape)
```

(15840, 3)

Fonte: (CASTILLO, 2021).

Esse código é semelhante ao código apresentado na Figura 21. Os passos são, portanto, os mesmos, dispensando comentá-los novamente.

Agora vamos analisar as amostras obtidas e a partir delas exibir os intervalos de credibilidade para os parâmetros do modelo, w_0 , w_1 e Ω_{0m} . Isso é feito nas linhas de código da Figura 27. A explicação dessa célula é a mesma da Figura 24; a única diferença é que, no código atual, estão presentes dois parâmetros a mais em relação à figura citada, que são w_0 e w_1 . Por isso não serão explicadas as linhas desse código; apenas analisaremos o resultado.

Figura 27 – Estimativa de valores dos parâmetros do modelo de energia escura.

```
1 labels_f_sne = [r'w_0', r'w_1', r'\Omega_{0m}']
2 results_f_sne = []
3 for i in range(ndim_f_sne):
4     mcmc_f_sne = np.percentile(flat_samples_f_sne[:, i], [16, 50, 84])
5     q_f_sne = np.diff(mcmc_f_sne)
6     txt_f_sne = "\mathrm{{{3}}} = {{0:.3f}}_{{-{{1:.3f}}}}^{{{{2:.3f}}}}"
7     txt_f_sne = txt_f_sne.format(mcmc_f_sne[1], q_f_sne[0],
8                                 q_f_sne[1], labels_f_sne[i])
9     display(Math(txt_f_sne))
10    results_f_sne.append(mcmc_f_sne[1])
```

$$w_0 = -0.906_{-0.072}^{0.058}$$

$$w_1 = 0.513_{-0.121}^{0.125}$$

$$\Omega_{0m} = 0.284_{-0.023}^{0.033}$$

Fonte: (CASTILLO, 2021).

Como resultado obtivemos os seguintes intervalos de credibilidade:

- Para w_0 , o valor médio é $-0,903$, e

$$-0,093 - 0,078 \leq w_0 \leq -0,093 + 0,054.$$

- Para w_1 , o valor médio é 0,497,

$$0,497 - 0,132 \leq w_1 \leq 0,497 + 0,130.$$

- Para Ω_{0m} , o valor médio é 0,283, e

$$0,283 - 0,034 \leq \Omega_m \leq 0,283 + 0,023.$$

Por meio do código da Figura 28 é possível gerar o gráfico de contorno dessas amostras para os parâmetros w_0 , w_2 e Ω_{0m} .

Figura 28 – Organizando as amostras.

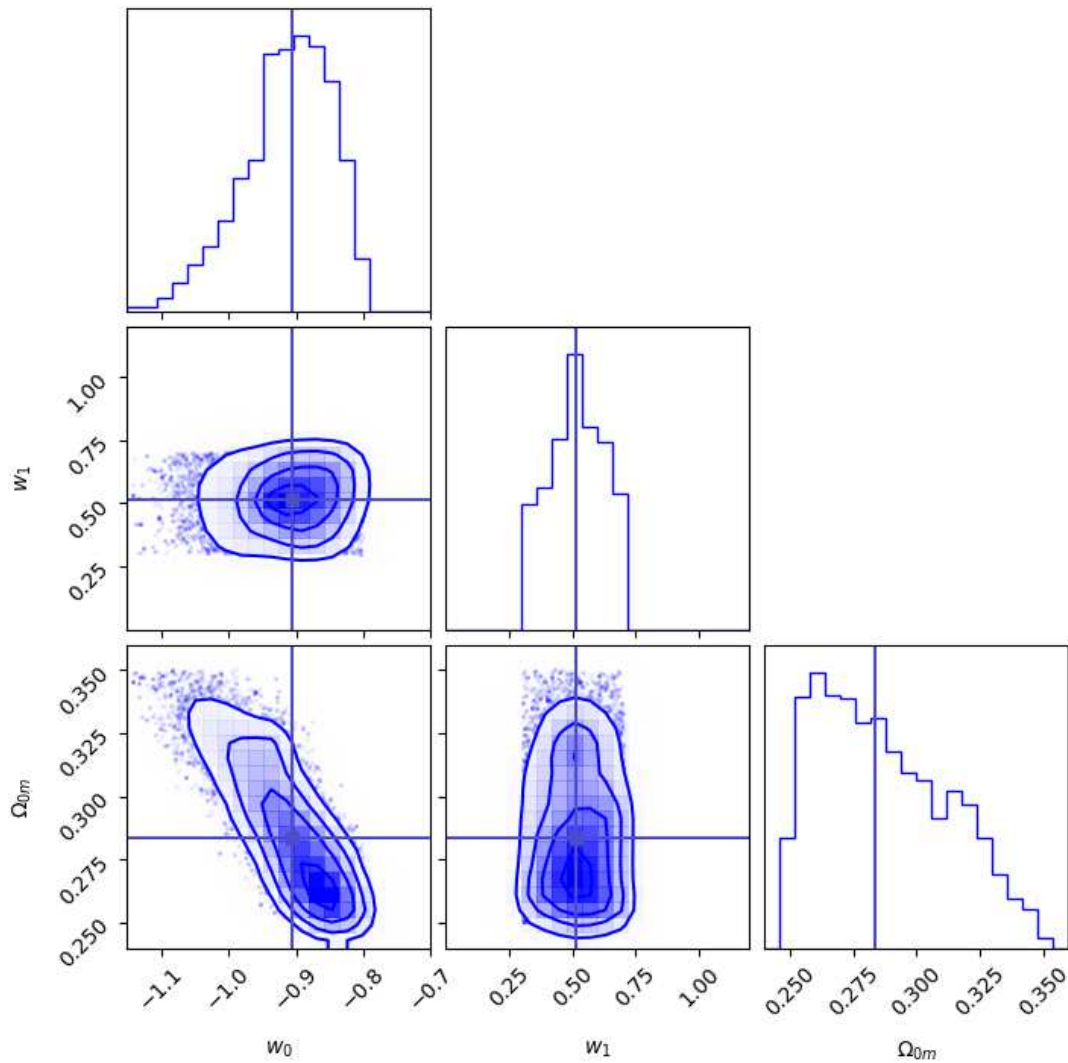
```
1 fig_f_sne = corner.corner(flat_samples_f_sne, range=[(-1.15, -0.7),
2 (0, 1.2), (0.24, 0.36)], truth_color='#4e4dc9', smooth=0.9, color='b',
3 labels=[r'$w_0$', r'$w_1$', r'$\Omega_{0m}$'],
4 truths=[results_f_sne[0], results_f_sne[1], results_f_sne[2]]);
5 fig_f_sne.savefig("contoursSN.png", bbox_inches='tight', dpi=500)
```

Fonte: (CASTILLO, 2021).

Por fim, apresentamos o gráfico de contornos na Figura 29, em que os parâmetros w_0 , w_1 e Ω_{0m} são estimados. As linhas azuis representam os valores mais prováveis desses parâmetros, calculados através do método de máxima verossimilhança, o que serviu de ponto inicial para o método de MCMC.

Os valores dos parâmetros w_0 , w_1 e Ω_{0m} são obtidos conforme Figura 30:

Figura 29 – Distribuição a posteriori unidimensional e contornos bidimensionais para o espaço de parâmetros (w_0, w_1, Ω_{0m}) referentes ao modelo de energia escura 3.3.



Fonte: (CASTILLO, 2021).

Figura 30 – Valores dos parâmetros w_0, w_1 e Ω_m .

```
1 results_f_sne
```

```
[-0.9027152148831092, 0.4971806159671268, 0.2825598293069979]
```

Fonte: (CASTILLO, 2021).

Considerações finais

O presente trabalho apresenta os principais fundamentos cosmológicos atuais, desde as equações de Friedmann até a análise de dados de supernovas do Tipo Ia. A partir das equações de Friedmann, pudemos entender como a teoria da Relatividade de Einstein descreve a expansão do espaço-tempo e nos fornece uma estrutura teórica para entender como ocorre essa expansão. A lei de Hubble também é um dos fundamentos cosmológicos importantes, ela relaciona a velocidade de recessão das galáxias e suas distâncias, é um dos pilares cosmológicos que oferecem evidências da expansão do universo. O modelo Λ CDM, aceito como modelo padrão, descreve a expansão do universo, incorporando uma constante cosmológica e a matéria escura.

Além disso, é expressa também uma introdução aprofundada à estatística Bayesiana, mostrando como é realizado o tratamento de dados utilizando este paradigma estatístico. Ao longo deste estudo, aprendemos como a estatística Bayesiana se diferencia da estatística frequentista, pois ela considera os conhecimentos prévios (a priori) e realiza a atualização da crença através das distribuições a posteriori.

A estatística Bayesiana oferece ferramentas que são fundamentais para a análise de dados cosmológicos contemporâneos, como o Método de Monte Carlo por Cadeias de Markov (MCMC). Ao aplicar esse método em Python, foi possível visualizar e entender melhor como os paradigmas Bayesianos são utilizados na prática. O MCMC se provou ser uma ferramenta poderosa e eficaz, e os resultados obtidos demonstraram a capacidade do MCMC de estimar parâmetros cosmológicos de maneira eficiente, fornecendo estimativas confiáveis, assim como a presença de suas incertezas associadas.

Ao compreender esses conceitos, foi possível comprovar a importância da estatística Bayesiana na interpretação de dados complexos, como os cosmológicos. De maneira que podemos concluir que, utilizando essa abordagem, conseguimos obter resultados/informações mais profundos e contextualizados sobre os fenômenos cosmológicos, levando a uma compreensão mais significativa do universo.

Fizemos a aplicação do MCMC a um modelo específico de energia escura, o qual fornece resultados semelhantes ao modelo Λ CDM para o tempo atual, ilustrando a diversidade de abordagens na cosmologia e como podemos utilizar diversos modelos para buscar entender o universo. Esperamos que este trabalho sirva de base para investigações posteriores e contribua para o avanço da ciência e da cosmologia. Além disso, acreditamos que outras parametrizações e novos modelos podem ser explorados em aplicações subsequentes.

Referências

AL, R. van de Schoot et. *Bayesian statistics and modelling*. [s.n.], 2021. Disponível em: <<https://doi.org/10.1038/s43586-020-00001-2>>. Citado na página 49.

ALPHER, R. A.; BETHE, H.; GAMOW, G. The origin of chemical elements. *Physical Review*, v. 73, 4 1948. ISSN 0031-899X. Citado na página 25.

ALPHER, R. A.; HERMAN, R. C. Theory of the origin and relative abundance distribution of the elements. *Reviews of Modern Physics*, v. 22, 4 1950. ISSN 0034-6861. Citado na página 25.

ALVES, L. D. O. Abordagem bayesiana para tratamento de dados faltantes com aplicação em um modelo logístico. *Universidade Federal do Ceará*, 2019. Citado na página 52.

AMENDOLA, L.; TSUJIKAWA, S. *Dark energy: theory and observations*. [S.l.]: Cambridge University Press, 2010. Citado 6 vezes nas páginas 27, 28, 29, 31, 35 e 38.

ANDREON S.; WEAVER, B. *Bayesian Methods for the Physical Sciences*. v. 4. ISBN 978-3-319-15286-8.: [S.l.]: Springer International Publishing, 2015. Citado na página 41.

ANGELOTTI, W. F. D. et al. Uma abordagem simplificada do método monte carlo quântico: da solução de integrais ao problema da distribuição eletrônica. *Química Nova*, SciELO Brasil, v. 31, p. 433–444, 2008. Citado na página 59.

AWARI. *Gerando números aleatórios com Python: Dicas e truques*. 2023. Disponível em: <<https://awari.com.br/gerando-numeros-aleatorios-com-python-dicas-e-truques/>>. Citado na página 63.

BATISTA, B. D. de O. Inferência bayesiana. *Universidade Federal de Lavras*, 2021. Citado na página 52.

BAUMANN, D.; MCALISTER, L. *Inflation and string theory*. [S.l.]: Cambridge University Press, 2015. ISBN 978-1-107-08969-3. Citado na página 26.

BROWN, D. van R. P. C. S. D. *A simple introduction to Markov Chain Monte–Carlo sampling*. [s.n.], 2018. Disponível em: <<https://doi.org/10.3758/s13423-016-1015-8>>. Citado na página 59.

CASTILLO, A. J. G. Statistical analysis of a bidimensional dark energy parametrization. Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, 2021. Disponível em: <https://github.com/Aldo-Gamboa/Bidimensional-Dark-Energy-Model/blob/main/Bidimensional_DE_models.ipynb>. Citado 22 vezes nas páginas 63, 64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84 e 85.

CHANG, Z.; ZHAO, D.; ZHOU, Y. Constraining the anisotropy of the universe with the pantheon supernovae sample. *Chinese Physics C*, v. 43, n. 12, p. 125102, 2019. Disponível em: <<http://hepnp.ihep.ac.cn/en/article/doi/10.1088/1674-1137/43/12/125102>>. Citado na página 64.

- CHIB, S.; GREENBERG, E. Understanding the metropolis-hastings algorithm. *The american statistician*, Taylor & Francis, v. 49, n. 4, p. 327–335, 1995. Citado na página 60.
- DIMOPOULOS, K. *Introduction to cosmic inflation and dark energy*. [S.l.]: CRC Press, 2021. ISBN 9781351174862. Citado 8 vezes nas páginas 26, 27, 28, 31, 32, 33, 34 e 35.
- DODELSON, S.; SCHMIDT, F. *Modern Cosmology*. [S.l.]: Academic Press, 2021. ISBN 978-0-12-815948-4. Citado 2 vezes nas páginas 25 e 26.
- DOWNEY, A. B. *Think Bayes*. Gravenstein Highway North, Sebastopol, CA: O’Reilly Media, Inc, 2021. Citado 6 vezes nas páginas 42, 47, 49, 51, 56 e 57.
- DOWNEY, A. B. *Think Bayes: Bayesian statistics in python*. 2. ed. [S.l.]: O’Reilly Media,, 2021. Citado 2 vezes nas páginas 42 e 47.
- DURRER, R. *The Cosmic Microwave Background*. [S.l.]: Cambridge University Press, 2020. ISBN 9781316471524. Citado na página 26.
- ESCAMILLA-RIVERA, C. Status on bidimensional dark energy parameterizations using sne ia jla and bao datasets. *Galaxies*, MDPI, v. 4, n. 3, p. 8, 2016. Citado 2 vezes nas páginas 38 e 40.
- ETZ, A. *Introduction to the Concept of Likelihood and Its Applications*. [s.n.], 2018. v. 1. 60-69 p. Disponível em: <<https://doi.org/10.1177/2515245917744314>>. Citado na página 53.
- EVANS, R. *The Cosmic Microwave Background*. [S.l.]: Springer International Publishing, 2015. ISBN 978-3-319-09927-9. Citado na página 25.
- FARIAS, A. M. L. de; KUBRUSLY, J. Q.; SOUZA, M. A. de O. Teoria das probabilidades ii variáveis aleatórias unidimensionais. Universidade Federal Fluminense, Departamento de Estatística, 2016. Disponível em: <<https://www.professores.uff.br/jessica/wp-content/uploads/sites/137/2017/09/GET00143-0.pdf>>. Citado 4 vezes nas páginas 43, 44, 45 e 46.
- FASTERCAPITAL. *Espaco de evento probabilidade incondicional o significado dos espacos de eventos*. 2024. Disponível em: <<https://fastercapital.com/pt/contente/Espaco-de-evento--probabilidade-incondicional--o-significado-dos-espacos-de-eventos.html>>. Citado na página 41.
- FERREIRA, S. M. *Bioestatística Básica*. Rio de janeiro, RJ: [s.n.], 2021. Disponível em: <<http://www.lampada.uerj.br/bioestatisticabasica>>. Citado 5 vezes nas páginas 43, 44, 45, 46 e 47.
- FIXSEN, D. J. The temperature of the cosmic microwave background. *The Astrophysical Journal*, v. 707, 12 2009. ISSN 0004-637X. Citado na página 25.
- FLEISCH, D. *A student’s guide to vectors and tensors*. 1. ed. [S.l.]: Cambridge University Press, 2017. ISBN 9780521171908. Citado na página 29.
- GAMOW, G. Expanding universe and the origin of elements. *Physical Review*, v. 70, 10 1946. ISSN 0031-899X. Citado na página 25.

GELMAN, A. et al. *Bayesian data analysis*. [S.l.]: Chapman and Hall/CRC, 1995. Citado 2 vezes nas páginas 51 e 54.

GIOVANNINI, M. *A primer on the physics of the cosmic microwave background*. [S.l.]: World Scientific Publishing, 2008. Citado na página 26.

GUNDERSEN, G. *Why Metropolis–Hastings Works*. 2019. Disponível em: <<https://gregorygundersen.com/blog/2019/11/02/metropolis-hastings/>>. Citado na página 60.

HASHIMOTO, M. aki et al. *Big-Bang Nucleosynthesis*. [S.l.]: Springer Singapore, 2018. ISBN 978-981-13-2934-0. Citado 2 vezes nas páginas 25 e 36.

HELM, J. A new version of the lambda-cdm cosmological model, with extensions and new calculations. *Journal of Modern Physics*, Scientific Research Publishing, v. 15, n. 2, p. 193–238, 2024. Citado 2 vezes nas páginas 36 e 37.

JONES, B. J. T. *Precision Cosmology The First Half Million Years*. [S.l.]: Cambridge university press, 2017. Citado 3 vezes nas páginas 36, 55 e 58.

JONES, E. et al. *SciPy: Open source scientific tools for Python*. 2001–. Disponível em: <<http://www.scipy.org/>>. Citado na página 64.

KARAMANIS, M. Bayesian computation in astronomy: Novel methods for parallel and gradient-free inference. *arXiv preprint arXiv:2303.16134*, 2023. Citado 2 vezes nas páginas 57 e 58.

KRUSCHKE, J. *Doing bayesian data analysis: A tutorial with r, jags, and stan*. Academic Press, 2014. Citado na página 60.

KRUSCHKE, J. K. Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, v. 6, n. 3, p. 299–312, 2011. PMID: 26168520. Disponível em: <<https://doi.org/10.1177/1745691611406925>>. Citado na página 48.

LI, B.; KOYAMA, K. *Modified gravity: progresses and outlook of theories, numerical techniques and observational tests*. World Scientific, 2019. Citado na página 36.

MAIOLI, M. Inferência bayesiana como um procedimento de decisão. *Monografia (PIBIC/CNPq)*. UNICAMP, 2014. Citado na página 51.

MARCH, M. C. *Advanced Statistical Methods for Astrophysical Probes of Cosmology*. [S.l.]: Springer Berlin Heidelberg, 2013. ISBN 978-3-642-35059-7. Citado na página 48.

MARTIN, O. A.; KUMAR, R.; LAO, J. *Bayesian Modeling and Computation in Python*. Boca Raton: [s.n.], 2021. ISBN 978-0-367-89436-8. Citado 2 vezes nas páginas 48 e 50.

MEYER, P. L. *probabilidade: Aplicações à Estatística*. Rio de Janeiro, RJ: LTC- Livros Técnicos e Científicos Editora S.A, 1965. Citado 2 vezes nas páginas 43 e 45.

MORESCO, M. Raising the bar: new constraints on the Hubble parameter with cosmic chronometers at $z \sim 2$. *Monthly Notices of the Royal Astronomical Society: Letters*, v. 450, n. 1, p. L16–L20, 04 2015. ISSN 1745-3925. Disponível em: <<https://doi.org/10.1093/mnras/51/lv037>>. Citado na página 64.

- MOURA, Í. K. F. *Testes cosmológicos aplicados a modelos de energia escura*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, 2016. Citado na página 39.
- MUKHANOV, V. F. *Physical foundations of cosmology*. [S.l.]: Cambridge university press, 2005. Citado na página 28.
- MYUNG, I. J. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, v. 47, n. 1, p. 90–100, 2003. ISSN 0022-2496. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022249602000287>>. Citado na página 53.
- OLIPHANT, T. E. *Guide to Numpy*. 2006. Disponível em: <<https://archive.org/details/NumPyBook>>. Citado na página 63.
- PADILLA, L. E. et al. Cosmological parameter inference with bayesian statistics. *Universe*, MDPI, v. 7, n. 7, p. 213, 2021. Citado 2 vezes nas páginas 39 e 40.
- PARDO, P. N. M. et al. *Estrutura lagrangiana para fluidos compressíveis não barotrópicos em dimensão dois*. Tese (Doutorado) — Thesis, IMECC-UNICAMP, 2013. Citado na página 33.
- PAULINO, C. D. e. a. *Estatística bayesiana*. [S.l.]: Fundação Calouste Gulbenkian,, 2018. v. 2. Citado na página 48.
- PENZIAS, A. A.; WILSON, R. W. A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, v. 142, 7 1965. ISSN 0004-637X. Citado na página 25.
- PETTINARI, G. W. *The Intrinsic Bispectrum of the Cosmic Microwave Background*. [S.l.]: Springer International Publishing, 2016. ISBN 978-3-319-21881-6. Citado na página 25.
- REGNIER, L. et al. Universal exploration dynamics of random walks. *Nature Communications*, 2023. Disponível em: <<https://doi.org/10.1038/s41467-023-36233-5>>. Citado na página 77.
- RIESS, A. G. et al. Milky way cepheid standards for measuring cosmic distances and application to <i>gaia</i> dr2: Implications for the hubble constant. *The Astrophysical Journal*, v. 861, 7 2018. ISSN 1538-4357. Citado na página 26.
- ROCCA, J. Bayesian inference problem, mcmc and variational inference. *Medium*, 2019. Disponível em: <<https://towardsdatascience.com/bayesian-inference-problem-mcmc-and-variational-inference-25a8aa9bce29>>. Citado na página 59.
- RODRIGUES, D. d. P. F. et al. Aspectos clássicos da cosmologia inflacionária. Universidade Federal da Paraíba, 2011. Citado na página 28.
- SALVATIER, J.; WIECKI, T. V.; FONNESBECK, C. Probabilistic programming in python using pymc3. *PeerJ Journals*, 2016. Citado 2 vezes nas páginas 64 e 78.
- SCHOOT, R. Van de et al. A gentle introduction to bayesian analysis: Applications to developmental research. *Child development*, Wiley Online Library, v. 85, n. 3, p. 842–860, 2014. Disponível em: <<https://doi.org/10.1111/cdev.12169>>. Citado 4 vezes nas páginas 49, 50, 52 e 54.

SCOLNIC, D. M. et al. The complete light-curve sample of spectroscopically confirmed sne ia from pan-starrs1 and cosmological constraints from the combined pantheon sample. *The Astrophysical Journal*, The American Astronomical Society, v. 859, n. 2, p. 101, may 2018. Disponível em: <<https://dx.doi.org/10.3847/1538-4357/aab9bb>>. Citado na página 64.

SILVA, J. L. de Castro e; FERNANDES, M. W.; ALMEIDA, R. L. F. de. *Matemática: Estatística e Probabilidade*. [S.l.]: Fortaleza - CE, 2015. Citado na página 41.

SILVA, M. *Domine a Visualização de Dados com Matplotlib em Python*. 2023. Disponível em: <<https://maykosilva.com/blog/domine-a-visualizacao-de-dados-com-matplotlib-em-python/>>. Citado na página 63.

STOROPOLI, J. *Estatística Bayesiana com R e Stan*. [s.n.], 2021. Disponível em: <<https://storopoli.github.io/Estatistica-Bayesiana>>. Citado 3 vezes nas páginas 42, 43 e 44.

SULLIVAN, M. et al. Snls3: Constraints on dark energy combining the supernova legacy survey three-year data with other probes. *The Astrophysical Journal*, v. 737, 8 2011. ISSN 0004-637X. Citado na página 26.

VANDEKERCKHOVE, A. E. J. *Introduction to Bayesian Inference for Psychology*. [s.n.], 2018. Disponível em: <<https://doi.org/10.3758/s13423-017-1262-3>>. Citado 3 vezes nas páginas 55, 56 e 57.

WEINBERG, S. *Foundations of Modern Physics*. [S.l.]: Cambridge University Press, 2021. ISBN 9781108894845. Citado na página 35.

YU, H.; RATRA, B.; WANG, F.-Y. Hubble parameter and baryon acoustic oscillation measurement constraints on the hubble constant, the deviation from the spatially flat Λ cdm model, the deceleration–acceleration transition redshift, and spatial curvature. *The Astrophysical Journal*, The American Astronomical Society, v. 856, n. 1, p. 3, mar 2018. Disponível em: <<https://dx.doi.org/10.3847/1538-4357/aab0a2>>. Citado na página 64.

ZIMDAHL, W. Big bang & energia escura: problemas atuais da cosmologia. *Cadernos de Astronomia*, v. 2, n. 1, p. 106–106, 2021. Citado na página 38.