Report on Project 1: Cosmological models

Erik Karlsson Öhman, Hampus Hansen

August 5, 2025

1 Introduction

In this report, we will infer the model parameters of the cosmological model ΛCDM, and compare it to its extension, the wCDM model, given data from type Ia supernova observations. The observed redshift is indicative of the expansion or contraction of the universe. The result of the redshift is characterized by the *distance modulus*

$$\mu = 5\log(d_L) + 25, \tag{1}$$

where d_L is the *luminosity distance*. This observable can be used to test different models of cosmology. In this report, both the ΛCDM model, characterized by Hubble parameter

$$H(z) = H_0\sqrt{\Omega_{M,0}(1+z)^3 + \Omega_{\Lambda,0}} \equiv H_0\sqrt{E_\Lambda(z)} \tag{2}$$

and its extended wCDM model, characterized by

$$E_w(z) \equiv \sqrt{\Omega_{M,0}(1+z)^3 + \Omega_{\Lambda,0}(1+z)^{3(1+w)}}, \tag{3}$$

will be compared. In the case of $w = -1$ the wCDM model simplifies to the original ΛCDM model. A flat universe will be considered, where the Hubble parameter is connected to the luminosity distance through

$$d_L(z) = c(1+z)\int_0^z \frac{dz'}{H(z')}. \tag{4}$$

For $z << 1$, we can Taylor expand $E_\Lambda(z)$, and we will approximate that

$$d_L(z) \approx \frac{c}{H_0}\left(z + \frac{1}{2}(1-q_0)z^2\right) \quad \text{for } z < 0.5, \tag{5}$$

$$d_L(z) \approx z\frac{c}{H_0} \quad \text{for } z < 0.05, \tag{6}$$

where q_0 is the deceleration parameter. For an accelerating universe q_0 is negative, and it is positive for a decelerating universe.

In the report a joint probability distribution of H_0 and q_0 is extracted in the small-z regime ($z < 0.5$) and a predictive posterior distribution for the ΛCDM model is made. The ΛCDM and wCDM models are also compared in the full z region by calculating their respective Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores. Lastly, a posterior probability distribution for $\Omega_{M,0}$ in the full ΛCDM model is extracted. In the case of a flat universe the matter and dark energy density parameters are related by $1 = \Omega_{M,0} + \Omega_{\Lambda,0}$, consequently a value for $\Omega_{\Lambda,0}$ is also obtained. [1]

2 Method

The SCP 2.1 dataset consist of observations from 833 type Ia supernovas, containing the name of the supernova, redshift, distance modulus and the measurement error of the distance modulus. The data likelihood is set to a normal distribution,

$$p(\boldsymbol{\mathcal{D}}|\mu, \sigma^2) \sim \mathcal{N}(\mu, \sigma^2), \tag{7}$$

centered around the modeled distance modulus μ with variance σ^2. In task one, the model parameters are H_0, q_0 and σ^2, where $\sigma^2 = \frac{\sigma^2_{\text{model}}}{w_i}$, where w_i are weights inversely proportional to the measurement errors and normalized such that $\sum_{i=1}^{N_d} w_i = N_d$. In task 2 the model parameters are $\{\Omega_{M,0}\}$ for ΛCDM and $\{\Omega_{M,0}, w\}$ for wCDM and the value of the Hubble constant is now constrained to $H_0 = 70$km/s/Mpc. For task 2, the data likelihood function is still a normal distribution centered around the respective model distance modulus, but now with variance $\sigma^2 = 1/w_i$. This assumes that the overall scale of the errors is $\sigma_{\text{model}} = 1$.

A normally distributed likelihood is chosen since the data comes from experimental observations, which are prone to errors. These errors arise from many independent factors. The total contribution of a large sum of independent errors tend to be normally normally distributed, according to the Central Limit Theorem. The choice of prior distribution for the two tasks are discussed below in their respective parts.

2.1 Task 1

The small-z approximation ($z < 0.5$) means that we can use Equation (5) to formulate $d_L(z < 0.5)$, and thus the model distance modulus

$$\mu = 5 \log \left(\frac{c}{H_0} \big(z + \frac{1}{2}(1 - q_0) z^2 \big) \right) + 25. \tag{8}$$

This region consists of 412 data points. The smaller region ($z < 0.05$) consists of 140 data points. Both regions are deemed to contain a sufficient amount of data for the respective analyses. The priors of the model parameters H_0 and q_0 are uniform,

$$\begin{aligned} p(H_0) &\sim \mathcal{U}(0, 100) \\ p(q_0) &\sim \mathcal{U}(-5, 5), \end{aligned} \tag{9}$$

where the limits are chosen as broad as possible to hopefully include the maximum a posteriori (MAP). The unknown error scale σ^2_{model} is modeled by an inverse-gamma distribution

$$p(\sigma^2_{\text{model}}) \sim \mathcal{IG}(\alpha = 1, \beta = 1). \tag{10}$$

Setting $\alpha = \beta = 1$ results in a relative uninformative prior with a long tail, something that captures our initial lack of knowledge about the error scale. We use the `pymc`-library to construct the uniform priors, inverse-gammas and normal distributions. The MAP is found using the built in function `find_MAP()`, which uses the Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimization algorithm [2]. The posterior sampling was done using `pymc`'s built in sequential Monte Carlo (SMC) sampler `sample_smc()`, with four chains of 2000 iterations each.

2.2 Task 2

To asses wether the ΛCDM or the wCDM is preferred to model the SCP 2.1 dataset the respective AIC

$$\text{AIC} = 2\ln p(\mathcal{D}|\theta_\star) - 2N_p \tag{11}$$

and BIC

$$\text{BIC} = 2\ln p(\mathcal{D}|\theta_\star) - N_p \ln N_d \tag{12}$$

scores are calculated. N_d and N_p denote the number of data points and the number of model parameters. A higher score for one model in relation to another model indicates that the model with the highest score is more favored by the data [3]. To find the maximum likelihood estimators $\Omega_{M,0\star}$ and $w_\star$, the `scipy.optimize.minimze` optimization algorithm is performed on the respective negative log-likelihoods of the two models. The bounds of the optimization were set to `[0,1.0]` for $\Omega_{M,0}$ and `[-2,2]` for w, which only appears in the wCDM model. This makes implicit the assumption that $0 < \Omega_{M,0} < 1$, which is reasonable since $\Omega_{M,0} + \Omega_{\Lambda,0} = 1$ and the densities $\Omega_{M,0}, \Omega_{\Lambda,0}$ ought to not be negative. Since wCDM simplifies to ΛCDM for $w = -1$ it is reasonable to assume that w should be in the vicinity of this value, which motivates the search bounds for w. The likelihood function used is once again a normal distribution $\mathcal{N}(\mu, 1/w_i)$ where the error scale is given by $1/w_i$ and w_i are inversely proportional to the measurement errors of the data. The integral in Equation (4), which contributes to the likelihood, is evaluated numerically.

To extract the posterior distribution for $\Omega_{M,0}$ a Markov chain Monte Carlo (MCMC) simulation is performed with the `emcee` python library, with four walkers and 2000 iterations. This is done with a uniform prior, $\Omega_{M,0} \sim \mathcal{U}(0,1)$ [1], motivated in a similar fashion as the bounds of optimization presented above.

3 Results and discussion - task 1

The SMC sampling results in the probability distributions of the parameters seen in Figure 1, with inferred parameter values $H_0 = 69.69^{+0.50}_{-0.47}$ km/s/Mpc, $q_0 = -0.43 \pm 0.08$ and $\sigma^2 = 0.040 \pm 0.003$, where the uncertainty denotes the 68% credible interval. The value of the Hubble constant is within the 68% credible interval of what others studies have shown [4]. The inferred deceleration parameter is larger than the value reported a similar study [5], $q_0 \approx -0.55$. This small disagreement is somewhat mitigated since they also reported a slightly larger H_0. From the joint probability distribution of H_0 and q_0 we see a negative correlation, meaning that H_0 increases with larger negative values for q_0. This is expected, since the former implies a faster acceleration of the Universe, which corresponds to a larger negative value of the deceleration parameter. The fact that all of the probability mass of q_0 is located in the negative region is convincing evidence that the universe is indeed accelerating. The error σ^2 is uncorrelated with both q_0 and H_0, which is desirable. Had this not been the case and then one could suspect that the model was ill conditioned or not suited to the dataset we are modeling.

The predictive posterior is seen in Figure 2a. From the figure it is apparent that the posterior predictive distribution accurately models the data distribution for $z < 0.5$. In the $z < 0.5$ region there are some outlier datapoints that are outside the posterior predictive region, though these likely have high individual measurement errors and thus downweighted. For $z > 0.5$, outside of the validity of Equation (5) we note a subtle downward drift of the datapoints which the posterior predictive distribution fail to accurately capture in its entirety. To validate the the extraction of H_0 we can disregard

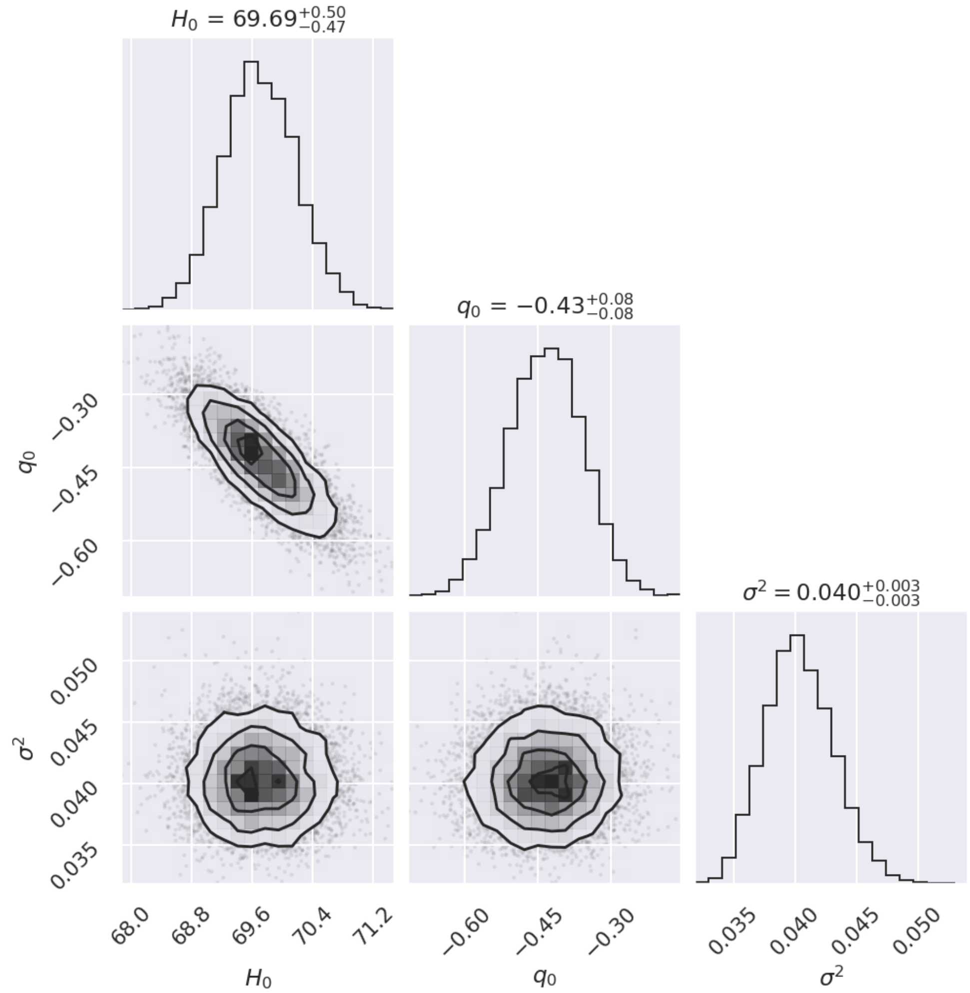
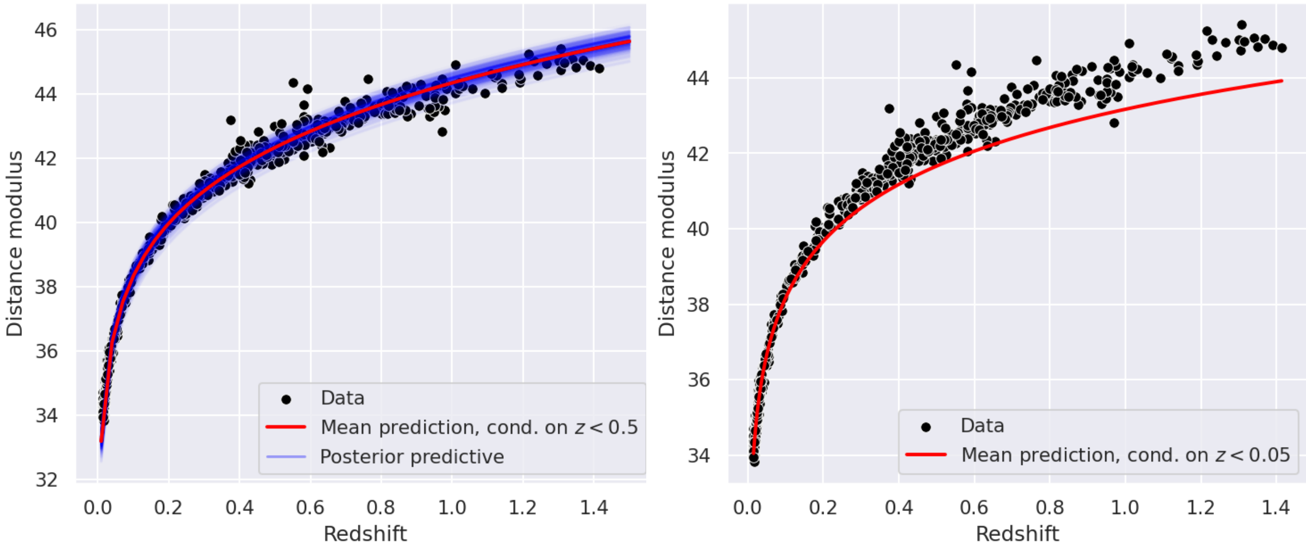


Figure 1: Joint probability distributions of the Hubble parameter H_0, the deceleration parameter q_0 and the error σ^2.

the term quadratic in z, resulting in Equation (6). By using this to formulate our model distance modulus and finding the MAP, we get the value $H_0 = 68.1\,\mathrm{km/s/Mpc}$, which is within four standard deviations of the extracted Hubble constant. Advantageous with this estimate is that H_0 is now inferred independently of q_0. As seen in Figure 2b, the model conditioned only on $z < 0.05$ will, unsurprisingly, fit the data well for small z, but starts to diverge more than than the model conditioned on $z < 0.5$, as seen in Figure 2a.



(a) Posterior predictive plot, represented by the blue lines. Each line is a draw from the posterior distribution. The red line is the mean prediction, and the black dots are data points.

(b) Mean prediction of distance modulus as a function of redshift, conditioned on the region $z < 0.05$. The data is represented as the black dots.

Figure 2: Comparison of predictive posterior conditioned on data $z < 0.5$ and the model conditioned on data $z < 0.05$.

The inferences of H_0 and q_0 could be improved by including higher order terms in the Taylor expansion of Equation (4) or for instance by extending the dataset with new data. Furthermore the inferences could also be improved by choosing more informative priors. The current priors are uniform, so a majority of the probability density in parameter space is located in very low-likelihood regions.

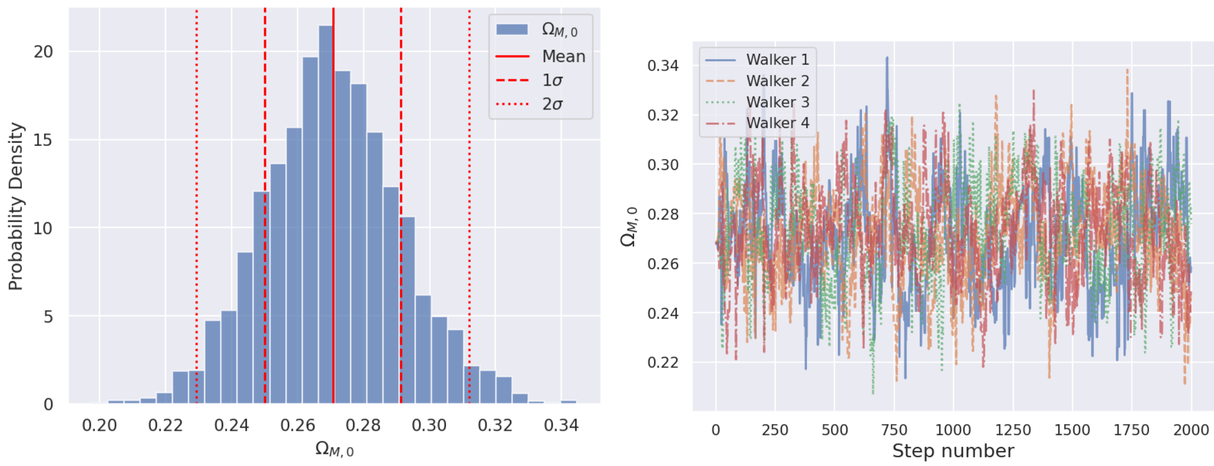
4 Results and discussion - task 2

The obtained maximum likelihood estimators for the respective models and the information criterion scores are presented in Table 1. The value of the matter density parameter has been measured to be $\Omega_{M,0} = 0.31 \pm 0.01$ [4]. Out of the two models examined wCDM gives the maximum likelihood estimator that is closest to this value, however this is expected since wCDM is a more complex model than ΛCDM and thus more prone to overfitting. The value of $w_\star$ is -1.10 which is very close to values of w obtained in similar studies [6]. That $w_\star$ is not equal -1, the value for which wCDM simplifies to ΛCDM, means that we cannot directly deem the wCDM model to be redundant. This implies that the wCDM may capture some complexity of nature that ΛCDM does not, or that the model overfits the data to a higher degree than ΛCDM. To gauge which model is actually favoured by the data it is more instructive to compare the respective AIC and BIC scores of the models. There we see that both the BIC and AIC scores are higher for the ΛCDM model, thus we can conclude that the ΛCDM is more favoured by the data than the wCDM model. We may also note that ΔAIC ≈ 2 and ΔBIC ≈ 6, the difference in score between the models is highest when calculating the BIC score. This is expected since BIC penalizes complex models more severely than AIC.

Table 1: Obtained parameter values and information criterion scores. The values of AIC and BIC are calculated without the constant factor of $(2\pi\sigma^2)^{-N_d/2}$ in the likelihood functions

Model	$\Omega_{M,0\star}$	$w_\star$	$\Omega_{\Lambda,0}$	BIC	AIC
ΛCDM	0.27	-	0.73	-30.33	-25.97
wCDM	0.31	-1.10	0.69	-36.67	-27.95

Figure 3a shows the extracted posterior distribution for $\Omega_{M,0}$ that is obtained by MCMC sampling. The walkers were initiated to $\Omega_{M,0} = \Omega^{\Lambda}_{M,0\star} = 0.268$, the maximum likelihood estimator for the ΛCDM model. This choice of initialization ensures that all of the walkers are sampling in a relevant region from the start, which negates the need to set a burn-in period during the beginning of the sampling. The soundness of this approach is further validated by the fact that the traces of the walkers, seen in Figure 3b, are relatively stable around the initialization value and show no indication of drift or consistent movement towards a new region of the parameter space. From the posterior distribution we obtain the following value for the matter density parameter $\Omega_{M,0} = 0.28 \pm 0.02$, where the uncertainty denotes one standard deviation from the mean. We may note that both the experimentally obtained value and the value obtained from the maximum likelihood estimation for $\Omega_{M,0}$ lay within two standard deviations of the value extracted from the posterior distribution [7].



(a) Posterior distribution of $\Omega_{M,0}$, where the dashed and dotted lines are the 1σ- and 2σ-levels respectively.

(b) Traces of the four walkers from sampling of $\Omega_{M,0}$ for 2000 iterations.

Figure 3: Posterior distribution of $\Omega_{M,0}$ obtained from MCMC sampling and corresponding traces of walkers.

References

[1] A. Ekström, *Advanced Simulation and Machine Learning: Project 1*. Gothenburg, Sweden: Chalmers University of Technology, Oct. 28, 2024.

[2] A. Ekström, *Advanced Simulation and Machine Learning: Lab 2*. Gothenburg, Sweden: Chalmers University of Technology, Oct. 28, 2024.

[3] P. Erhart, A. Ekström, and A. Gonoskov, *Advanced Simulation and Machine Learning*. Gothenburg, Sweden: Chalmers University of Technology, Oct. 28, 2024.

[4] P. A. R. Ade, N. Aghanim, M. Arnaud, *et al.*, "Planck2015 results: Xiii. cosmological parameters," *Astronomy amp; Astrophysics*, vol. 594, A13, Sep. 2016, ISSN: 1432-0746. DOI: 10.1051/0004-6361/201525830. [Online]. Available: http://dx.doi.org/10.1051/0004-6361/201525830.

[5] D. Camarena and V. Marra, "Local determination of the hubble constant and the deceleration parameter," *Phys. Rev. Res.*, vol. 2, p. 013 028, 1 Jan. 2020. DOI: 10.1103/PhysRevResearch.2.013028. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevResearch.2.013028.

[6] M. J. Mortonson, D. H. Weinberg, and M. White, *Dark energy: A short review*, 2013. arXiv: 1401.0046 [astro-ph.CO]. [Online]. Available: https://arxiv.org/abs/1401.0046.

[7] J. An, B.-R. Chang, and L.-X. Xu, "Cosmic constraints to the w cdm model from strong gravitational lensing," *Chinese Physics Letters*, vol. 33, no. 7, p. 079 801, Jul. 2016, ISSN: 1741-3540. DOI: 10.1088/0256-307x/33/7/079801. [Online]. Available: http://dx.doi.org/10.1088/0256-307X/33/7/079801.