PARAMETER ESTIMATION IN PHYSICAL MODELS: A COMPARISON BETWEEN BAYESIAN AND FREQUENTIST METHODS

by

Oscar M. Aguilar

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Applied Mathematics & Statistics).

Golden, Colorado

Date _____

Signed: _____
Oscar M. Aguilar

Signed: _____
Dr. Luis Tenorio
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____
Dr. Willy Hereman
Professor and Head
Department of Applied Mathematics & Statistics

ABSTRACT

A mathematical/physical model is a mathematical description of a system or physical experiment. A model may help to explain a system or experiment, or to make future predictions. Most of the real-life mathematical/physical models depend on parameters that need to be estimated. Parameter estimation techniques use mathematical/physical model of the system under consideration and physical measurements to infer numerical values of the parameters of interest. In this project, we are interested in estimating the gravity constant and the specific coefficient of air resistance in the equation describing a falling body. A physical experiment was conducted at Texas A&M University. In the experiment, a body was dropped from a known height and the free fall was recorded for one second using a video camera. The video recording was analyzed frame by frame to obtain the distance the body had fallen as a function of time, as well as measures of uncertainty in the data. We consider two approaches to estimate the parameters of interest: *Bayesian* and non-*Bayesian* techniques. The goal of this thesis is to complement Allmaras et al. (2012) work by providing a careful validation analysis of the assumptions and to show the use of non-Bayesian methods for parameter estimation.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First, I would like to thanks to my first alma mater *Universidad Nacional de Ingenieria del Peru*; the place where I learned all the basics. I have reached this far in my career because of the education that I have received there (2004 -2006).

Thanks to my thesis advisor, Dr. Luis Tenorio, for his countless hours of advice, for his friendship, and for constantly pushing me to be a better statistician and to solve problems in one line. Luis' favorite quote: *This is a one-liner!*

Thanks to my second alma mater the Department of Applied Math & Statistics at the Colorado School of Mines, including all the faculty and staff who have helped me during my time here (Spring 2010 - Spring 2013). In particular, thanks to Dr. Willy Hereman, who gave me the opportunity to work on a REU project during summer 2011. Also thanks to my thesis committee Drs. William Navidi and Amanda Hering both of whom have taught me more statistics than I care to admit.

Last but not least thanks to my parents Victor Aguilar and Roxana Aguilar for their unconditionally support all the time.

To my parents Victor and Roxana

CHAPTER 1
INTRODUCTION

The main goal of mathematical modeling is to understand the behavior of a system and make future predictions. Many mathematical models depend on unknown parameters that need to be estimated by conducting experiments. Parameter estimation techniques use a mathematical/physical model of the system under consideration that provides information about the parameters of interest, and physical measurements to infer numerical values of these parameters.

In this project, we are interested in estimating g (gravity constant) and C (specific coefficient of air resistance) using the data from a physical experiment carried out by Allmaras et al. (2012) at Texas A&M University. In the experiment, a body is dropped from a known height, and the fall is recorded for one second by a video camera.

We consider two approaches to estimate the parameters of interest: *Bayesian* inversion and non-*Bayesian* inversion methods that include maximum likelihood and nonlinear regression. Bayesian inversion has became very popular in its application to inverse problems. The framework of Bayesian inversion combines any prior information that one has about the model before observing the data (prior probability model) with the information contained in the observed data to update the prior information. The updated distribution is the posterior conditional model distribution given the data. Most of the time, the posterior distribution is summarized by the posterior mean and posterior covariance matrix. However, very seldom are we able to compute all the posterior estimates analytically; we often employ Markov Chain Monte Carlo methods to approximate the posterior estimates (see, for example, [13]). However, we do not use Markov Chain Monte Carlo methods to compute the posterior dis-

tribution because we want to avoid convergence issues.

Even though Bayesian inversion has became very popular, Bayesian methods are not the only available tools in parameter estimation. On the other hand, the frequentist framework is based on repeated sampling from the distribution of the data conditioned on the unknown parameters. Comparisons between Bayesian and frequentist methods in inverse problems can be found in [15, 17]. Bayesian estimators can be evaluated from a frequentist perspective and vice versa (see, for example, [17]). In this project, we study the frequentist coverage of the Bayesian credible intervals.

The thesis is organized as follows: Chapter 2 summarizes how the experimental data were obtained and derive the physical model that will be used throughout. Chapter 3 discusses the Bayesian inversion approach to estimate the parameters of interest. This chapter summarizes what was done by Allmaras et al. (2012) and also discusses frequentist properties of the Bayesian intervals that were not included in Allamaras et al. (2012). Chapter 4 discusses maximum likelihood estimates of the parameters of interest and provides a comparison with the results of Chapters 3. Chapter 5 discusses the assumptions made by Allmaras et al. (2012). Specifically, we discuss if the likelihood function for the data assumed by Allmaras et al. (2012) is adequate or not. Chapter 6 discusses the nonlinear regression analysis approach to estimate the parameters of interest. Finally, Chapter 7 discusses an alternative approach to approximate confidence intervals for the parameter of interest suggested by P. B. Stark (see, for example, [16]).

CHAPTER 2

THE EXPERIMENTAL DATA AND THE MODEL

In the first part of this chapter we summarize how the experimental data were obtained. In the second part we derive the physical model that will be used throughout. The experiment was conducted by Allmaras et al. (2012) at Texas A&M University. From now on we will refer to Allmaras et al. (2012) as "ALA12". In this experiment, a body was dropped from a known height and the free fall was recorded using a video camera. The video recording was analyzed frame by frame to obtain the distance the body had fallen as a function of time, and to derive measures of uncertainty in the data. The data were obtained at certain times t_i by analyzing frames of the video that recorded the falling of a body in one second. The camera records thirty frames per second, so if the initial time t_0 is fixed and i is the number of the video frame through which the body is seen falling, then the time corresponding to frame i is defined as $t_i = t_0 + i/30$ (s). The distance, d_i, that the body fell in each frame was also recorded. In real-life experiments there are always measurement errors in the data. Two possibles types of measurement errors in the experiment are:

- The limited resolution of the video camera.
- The rotation of the body as it falls, which makes it difficult to obtain an accurate reading of the fallen distance.

The major source of error in reading the traveled distance by the body was its rotation as shown in Figure 2.1.

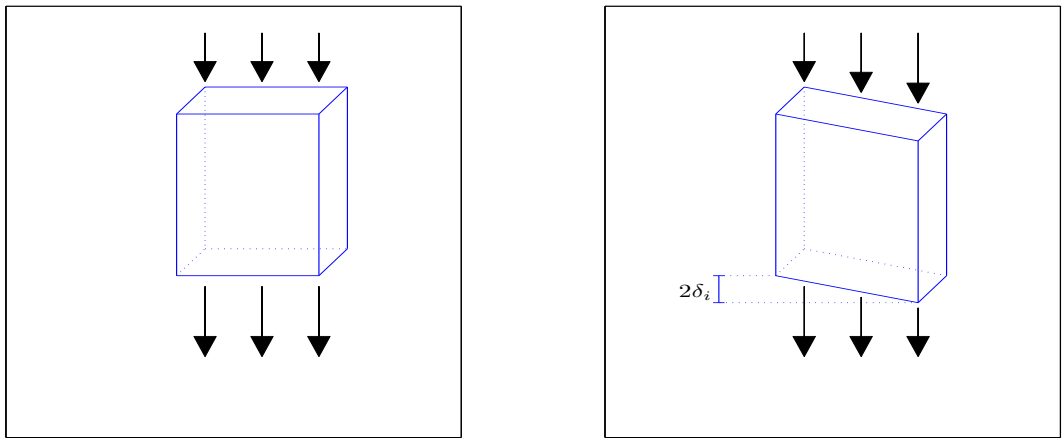

Figure 2.1: Graphical representation of the body as it falls.

Figure 2.1 depicts the falling body. The left panel in the figure represents an ideal situation when there is no rotation. On the other hand, the right panel represents one of the video recording frames, where the body rotates as it falls. ALA12 determined the uncertainty of each measurement, δ_i, and the fallen distance of the body, d_i, from each video frame. Figure 2.2 shows the traveled distance of the body, d_i, in each video frame and its uncertainty, $\pm\delta_i$, represented as error bars. Note that the uncertainties vary with time.

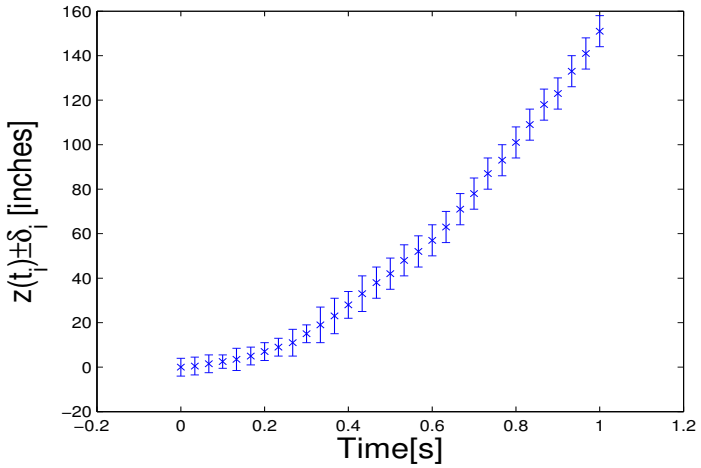

Figure 2.2: Raw data showing the distance the body has fallen (d_i) at a given time with its uncertainty ($\pm\delta_i$).

We next proceed to derive the mathematical model that will be used.

2.1 The Model for a Falling Body

The free fall of a body follows Newton's second law of motion,

$$F_{\text{total}} = m\frac{\mathrm{d}v}{\mathrm{d}t}, \tag{2.1}$$

where F_{total} is the sum of all the forces acting on the body, m is its mass, and $\mathrm{d}v/\mathrm{d}t$ is its acceleration. The are two forces acting in opposite directions on the body: the force of gravity, which can be expressed as mg, where g is the acceleration of the body due to gravity, and the frictional force caused by air resistance that makes the body decelerate. The air resistance is proportional to the square of its velocity, v. The constant of proportionality, k, (called the coefficient of air resistance), can be computed from a number of more fundamental parameters such as the density and viscosity of air and the cross-section of the object. In this project we are not interested in neither the mass of the body nor k. For simplicity, we define the specific coefficient of air resistance $C = k/m$ (m^{-1}).

The parameters of interest are g and C. Using some basic definitions from physics and calculus, we derive the equations that will be used to estimate these parameters using experimental data. First, let us focus on the total force acting on the body

$$F_{\text{total}} = mg - Cmv^2. \tag{2.2}$$

Using Newton's second law of motion (2.1) and (2.2), we find

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g - Cv^2, \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = v. \tag{2.3}$$

The initial conditions defined by assuming that the body is released at time t_0 with zero initial velocity are given by:

$$z(t_0) = 0, \qquad v(t_0) = 0. \tag{2.4}$$

Using equations (2.3) and (2.4), we obtain the following solutions:

$$v(t) = \sqrt{\frac{g}{C}}\tanh[\sqrt{gC}(t - t_0)], \qquad z(t) = \frac{1}{C}\log\cosh[\sqrt{gC}(t - t_0)], \tag{2.5}$$

where $v(t)$ and $z(t)$ are the velocity of the body and the traveled distance by the body at time t, respectively. We will now proceed to use the experimental data and the physical model (2.5) to estimate the parameters of interest.

CHAPTER 3

BAYESIAN INVERSION

In this chapter, we summarize what was done by ALA12. Let $\mathbf{m}$=(g, C) be the parameters to be estimated, and let $\mathbf{d}$ be the measurements of the physical model. ALA12 used a Bayesian approach to estimate the parameters of interest; they find the *posterior* probability density $\sigma(\mathbf{m}|\mathbf{d})$. This probability density function is calculated taking into account the physical model, the measured data, and any prior knowledge we may have about the parameters that we want to estimate.

As mentioned before, the goal of Bayesian inversion is to find the posterior probability distribution, but in some cases we are interested in some specific characteristic of this distribution such as a "maximum a posteriori" or "MAP", the posterior mean, and the posterior standard deviation. Sometimes a "MAP" or the posterior mean is used as a point estimate of the parameter of interest. The MAP is defined as:

$$\mathbf{m}_{\text{MAP}} = \arg\max_{\mathbf{m}} \sigma(\mathbf{m}|\mathbf{d}), \tag{3.1}$$

provided that the posterior distribution has a single mode. The posterior mean is defined as:

$$\mathbb{E}(\mathbf{m}|\mathbf{d}) = \int \mathbf{m}\, \sigma(\mathbf{m}|\mathbf{d}) \mathrm{d}\mathbf{m}. \tag{3.2}$$

The posterior mean tells us where the posterior density is centered. We may also be interested in the scatter of the posterior around the ith component of $\mathbf{m}$. This information is given by the posterior standard deviation defined as:

$$\text{std}(m_i|\mathbf{d}) = \sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |m_i - \mathbb{E}(\mathbf{m}_i|\mathbf{d})|^2\, \sigma(\mathbf{m}|\mathbf{d})\, \mathrm{d}m_1\, \mathrm{d}m_2}. \tag{3.3}$$

In order to use Bayesian methods, we need to define a prior distribution and a likelihood. We next define the prior distribution and the likelihood used by ALA12.

3.1 The Prior Distribution

The prior distribution expresses one's knowledge about the parameters before the data are taken into account. Let $\rho(\mathbf{m})$ be the chosen prior pdf of $\mathbf{m}$. Note that it would not make sense to assume that g or C could take negative values. ALA12 use the prior constraints: $7 \leq \ g \leq 11$ m/s^2 and $0 \leq \ C \leq 0.35$ m^{-1}, and choose to model g and C as independent random variables with g is uniformly distributed on $[7,\ 11$ m/s$^2]$ and C uniformly distributed on $[0,\ 0.35$ m$^{-1}]$. Hence, the prior pdf is given by

$$\rho(\mathbf{m}) = \rho(g, C) \ \propto \ \chi[7,\ 11 \text{ m/s}^2](g)\, \chi[0,\ 0.35 \text{ m}^{-1}](C), \tag{3.4}$$

where χ_A is the characteristic function of the set A.

3.2 The Likelihood

For a fixed parameter $\mathbf{m}$, we can predict how the system will behave. Let us define the vector $\mathbf{f}(\mathbf{m})$ with entries $f_i(\mathbf{m}) = z(t_i)$, where $z(t_i)$ is the distance the body has fallen at time t_i as given by (2.5), for the fixed $\mathbf{m}$. The experimental data are modeled as:

$$\mathbf{d} = \mathbf{f}(\mathbf{m}) + \epsilon, \tag{3.5}$$

where ϵ is a vector of measurement errors that introduces random uncertainty in the data. ALA12 define a pdf for ϵ_i as an inverted parabola centered around zero with support $[-\delta_i, \delta_i]$. The probability density of ϵ_i is

$$\rho_i(\epsilon_i) = K_i \left(1 - \frac{\epsilon_i^2}{\delta_i^2}\right) \chi[-\delta_i\ ,\ \delta_i](\epsilon_i), \tag{3.6}$$

where $K_i = 3/4\delta_i$ is the normalization constant. By doing a rescaling and shifting of the beta distribution, $Beta(2, 2)$, we obtain (3.6):

$$\frac{\epsilon_i}{\delta_i} \ \sim \ 2Beta(2, 2) - 1. \tag{3.7}$$

In Figure 3.1, we see how we obtain the distribution (3.7) after a rescaling and shifting of a $Beta(2,2)$ distribution.

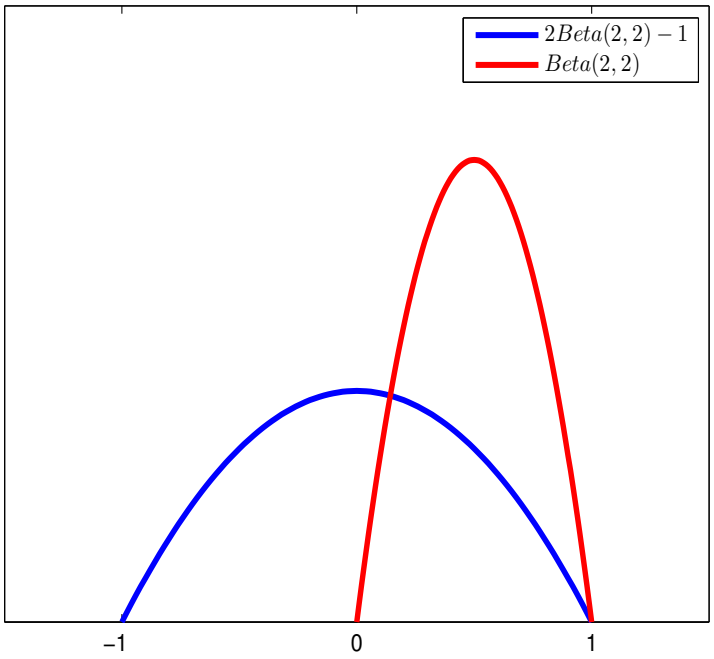

Figure 3.1: Black line represents a beta distribution with parameters $\alpha = 2$ and $\beta = 2$, and the blue line is a rescaling and shifting of the beta distribution with parameters $\alpha = 2$ and $\beta = 2$.

From (3.5), we note that the pdf of $\mathbf{d}$ for a fixed $\mathbf{m}$ is the pdf of ϵ shifted by $f(\mathbf{m})$. The random behavior of $\mathbf{d}$, conditioned on the fixed parameters g and C, is described by the conditioned probability density $\rho(\mathbf{d}|\mathbf{m})$. This also defines the likelihood function of $\mathbf{m}$ with a fixed $\mathbf{d}$. From (3.5) and (3.6), we obtain

$$\rho_i(d_i|\mathbf{m}) = K_i \left(1 - \frac{|d_i - f_i(\mathbf{m})|^2}{\delta_i^2}\right) \chi[f_i(\mathbf{m}) - \delta_i\ ,\ f_i(\mathbf{m}) + \delta_i](d_i). \tag{3.8}$$

Since the measurements were extracted from individual frames, ALA12 assumed that the measurement errors are independent. Therefore, the probability density that describes all

the measurements is given by the product of the densities for each measurement.

$$\rho(\mathbf{d}|\mathbf{m}) = \prod_i \rho_i(d_i|\mathbf{m}). \tag{3.9}$$

3.3 Finding the Posterior Distribution

The goal of parameter estimation in Bayesian inversion is to compute the probability density $\sigma(\mathbf{m}|\mathbf{d}) = \sigma(g, C|\mathbf{d})$ using the mathematical model based on the physics (2.5), the likelihood function (3.8), the prior distribution (3.4), and the experimental data. By Bayes' theorem,

$$\sigma(\mathbf{m}|\mathbf{d}) \ \propto \ \rho(\mathbf{m})\, \rho(\mathbf{d}|\mathbf{m}). \tag{3.10}$$

In order to compute $\sigma(\mathbf{m}|\mathbf{d})$, ALA12 select a grid of points for the parameter $\mathbf{m}$, and for each value of $\mathbf{m}$ they compute the value of the prior density (3.4). Then, they calculate the predicted measurements $\mathbf{f}(\mathbf{m})$ using the forward model (2.5) and evaluate the corresponding likelihood $\rho(\mathbf{d}|\mathbf{m})$ (3.8).

3.4 First Results

A straightforward way to visualize $\sigma(g, C|\mathbf{d})$ is to evaluate it at a number of points $(g_0 + i \triangle g,\, C_0 + j \triangle C)$ and plot a function that interpolates at these points. Such a plot of the non-normalized probability density for different values of the parameter $\mathbf{m} = (g, C)$ is shown in Figure 3.2. ALA12 use a grid of 401 points in the range $7 \text{ m/s}^2 \leq g \leq 11 \text{ m/s}^2$ and 351 points in the range of $0 \text{ m}^{-1} \leq C \leq 0.35 \text{ m}^{-1}$. It has been verified that this grid of points is fine enough to ignore the quadrature error. The posterior density σ is zero outside the ranges for g and C. In Figure 3.2, the contours with the highest values are shown in red, and the contours with the lower values are shown in blue. The posterior expected values for g and C are computed using (3.2) by approximating the integrals with appropriately weighted sums (quadrature) over the same grid of points from which the posterior distribution was computed. The posterior means are:

$$\mathbb{E}(g|\mathbf{d}) = 8.82 \text{ m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.116 \text{ m}^{-1}. \tag{3.11}$$

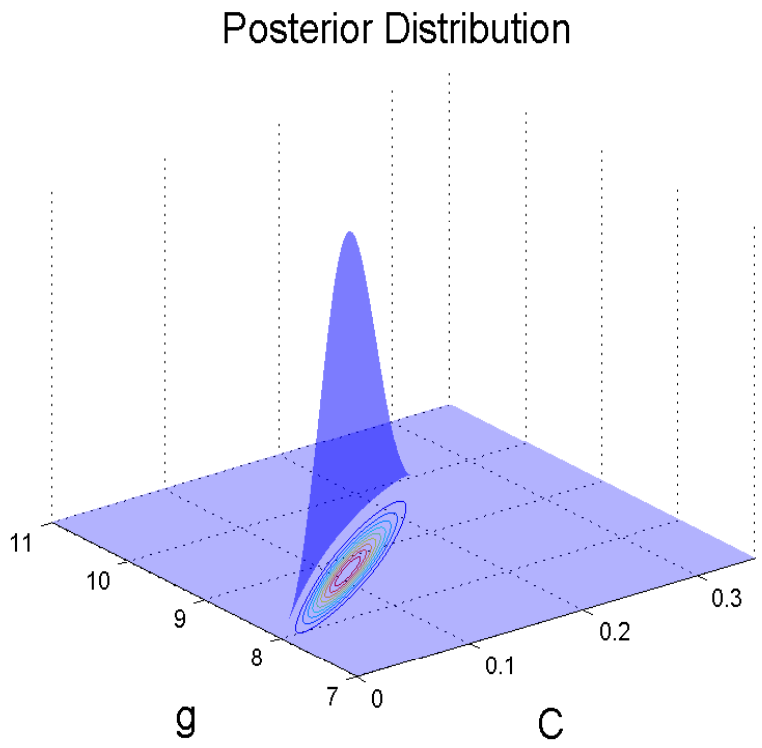

Figure 3.2: Visualization of the non-normalized posterior probability density $\sigma(g, C|\mathbf{d})$.

ALA12 used $\mathbb{E}(g|\mathbf{d})$ as an estimate of g. While they have no value to compare to C (as it depends on the viscosity and density of air as well as the weight and cross section of the dropped body), the value of the gravity constant at College Station, is approximately 9.79 $\mathrm{m/s^2}$. The estimated value for g (3.11) seems to underestimates its true value. But the Bayesian solution is the posterior distribution which provides more information than just the posterior mean (3.11). Therefore, we can look at the posterior probability density for g given by

$$\sigma(g|\mathbf{d}) = \int_{-\infty}^{\infty} \sigma(g, C|\mathbf{d}) \, \mathrm{d}C. \tag{3.12}$$

Again, the integrals are approximated by weighted sums over the same grid of points from which the posterior distribution was computed. Figure 3.3 shows the marginal posterior of g.

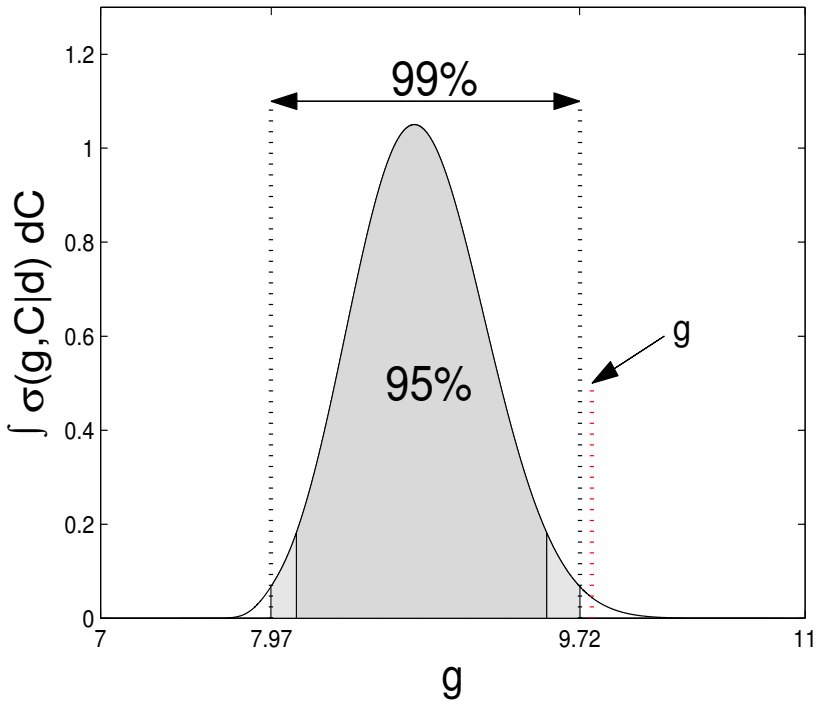

Figure 3.3: Marginal posterior density for gravity $\sigma(g|\mathbf{d})$.

In Figure 3.3, the darker area in gray represents the shortest 95% credible interval, [8.11 m/s^2 , 9.53 m/s^2], and the clearer area in gray represents the shortest 99% credible interval, [7.97 m/s^2 , 9.72 m/s^2]. From these results, we see that the true value of the gravity constant ($\approx$ 9.79 m/s^2) lies outside the 95% and 99% shortest credible intervals. This implies that the true value for g is highly unlikely ($P(g \geq 9.79 \mid \mathbf{d}) = 0.006$) under the posterior of the proposed model and available data.

3.5 Adding a Third Parameter: Starting Time

The results in Section 3.4 were computed treating t_0 as a fixed number ($t_0 = 0$). However, while ALA12 were reading the traveled distance of the body frame-by-frame, they realized that it was not trivial to estimate the starting time when the body was released. The time $\triangle t$ between each frame is 1/30 of a second, and the distance that the body fell in the first frame after release is $(1/2)g \triangle t^2 \approx 0.5$ cm. After flipping back and forth

between the frames a number of times, they concluded that the body was definitely moving between frames 42 and 43, and that it was not moving between frames 40 and 41. They concluded that the starting time was between frame 41 and 42; that is $41 \triangle t \ \leq \ t_0 \ \leq \ 42 \triangle t$.

ALA12 also used various values for t_0 to compute the results in Section 3.4, and realized that $\sigma(\mathbf{m}|\mathbf{d})$ was very sensitive to the value of t_0. It was then decided that t_0 is another unknown parameter that needs to be included in the model. The new parameter vector for the physical model (2.5) is then $\mathbf{m} = (g, C, t_0)$. The prior probability density for $\mathbf{m}$ in Section 3.1 is modified by ALA12 by assuming that t_0 is independent of g and C and has a probability density function given by

$$\psi(t_0) = \begin{cases} 1 & \text{for } 41 \triangle t \ \leq \ t_0 \ \leq \ 42 \triangle t, \\ 1 - (t_0 - 42 \triangle t)/ \triangle t & \text{for } 42 \triangle t \ \leq \ t_0 \ \leq \ 43 \triangle t, \\ 0 & \text{for } t_0 < 42 \triangle t \text{ or } 43 \triangle t < t_0 \end{cases} \tag{3.13}$$

And the graphical representation of the prior distribution of t_0 is shown in Figure 3.4.

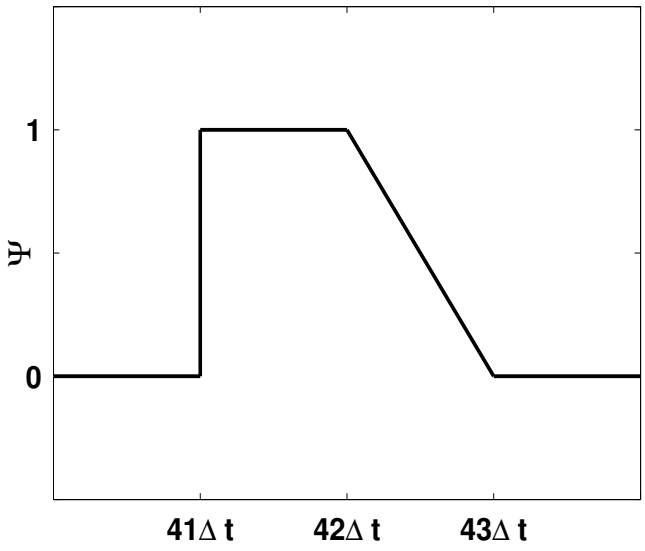

Figure 3.4: Prior distribution of t_0.

The new prior probability density function for $\mathbf{m}$ is then

$$\rho(\mathbf{m}) = \rho(g, C, t_0) \ \propto \ \chi[6, \ 12 \text{ m/s}^2](g) \, \chi[-0.2, \ 0.5 \text{ m}^{-1}](C) \, \psi(t_0). \tag{3.14}$$

By combining the likelihood (3.8) and the new prior distribution (3.14), the posterior distribution $\sigma(g, C, t_0|\mathbf{d})$ is computed. ALA12 computed the posterior distribution over a grid of 601 points in the range 6 m/s^2 $\leq g \leq$ 12 m/s^2, 141 points in the range -0.2 m^{-1} $\leq C \leq$ 0.5 m^{-1}, and 101 points in the range $42 \triangle t - 0.05$ s $\leq t_0 \leq 42 \triangle t + 0.05$ s. Again, it was verified that this mesh of points is fine enough to ignore the quadrature error. Using (3.2) and the same grid of points from which the posterior density function was computed, the posterior means were found to be

$$\mathbb{E}(g|\mathbf{d}) = 8.64 \text{ m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.106 \text{ m}^{-1}, \qquad \mathbb{E}(t_0|\mathbf{d}) = 1.395 \text{ s}. \tag{3.15}$$

These results show that the addition of the third parameter (t_0) has a significant impact on the posterior mean of g; the new posterior mean (3.14) is farther from the true value of the gravity constant ($\approx$ 9.79 m/s^2) than the posterior expected value for g computed treating t_0 as a fixed number (3.11). It seems that the addition of t_0 did not lead to any improvement. However, the addition of t_0 causes a bigger uncertainty in g (a wider posterior density) when compared to the results in Section (3.4). We can see the impact of the addition of t_0 as another parameter in the model in the posterior density of g:

$$\sigma(g|\mathbf{d}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma(g, C, t_0|\mathbf{d}) \mathrm{d}t_0 \mathrm{d}C. \tag{3.16}$$

Again, the integrals are approximated by weighted sums over the same grid of points from which the posterior distribution was computed.

In Figure 3.5, the darker area in gray represents the shortest 95% credible interval, [7.43 m/s^2 , 9.93 m/s^2], and the clearer area in gray represents the shortest 99% credible interval, [7.33 m/s^2 , 10.31 m/s^2]. As a consequence of adding t_0 as the third parameter, the true value of the gravity constant lies in both the shortest 95% and 99% credible intervals and is a more plausible value for g under the posterior.

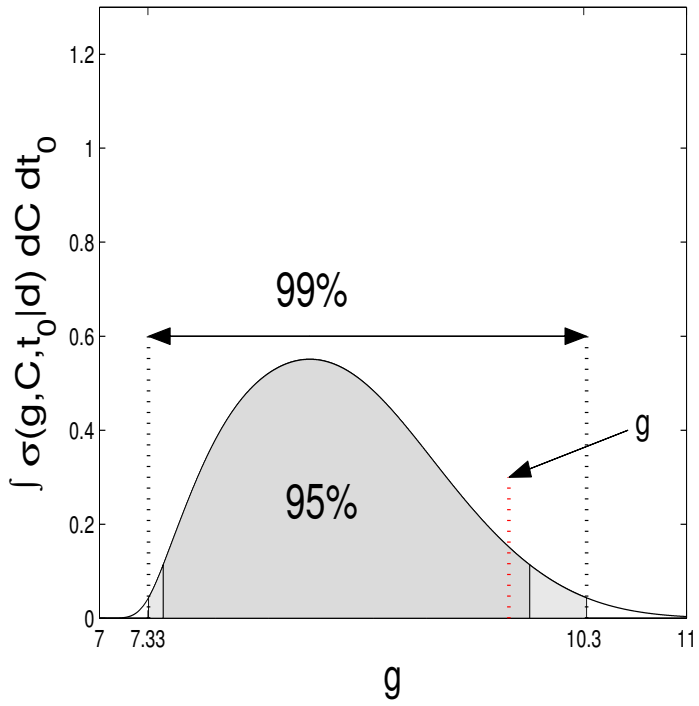

Figure 3.5: Marginal posterior density for gravity $\sigma(g|\mathbf{d})$.

The results (3.15) were obtained using (3.14) as the prior distribution. However, ALA12 did not justify why that particular prior distribution of t_0 was chosen. In ALA12's paper, they mentioned that the body was not moving between frame 40 and 41, and they concluded that the starting time was between frame 41 and 42. Then, the question is why did ALA12 use (3.13) as the prior distribution of t_0 instead of a uniform distribution between $41 \triangle t$ and $42 \triangle t$ as the prior distribution of t_0?

3.5.1 A Uniform Prior Distribution for the Starting Time

If we use a uniform distribution between $41 \triangle t$ and $42 \triangle t$ as the prior distribution of the starting time (t_0), we obtain the following posterior means:

$$\mathbb{E}(g|\mathbf{d}) = 8.28 \text{ m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.086 \text{ m}^{-1}, \qquad \mathbb{E}(t_0|\mathbf{d}) = 1.386 \text{ s}. \tag{3.17}$$

We see that the final results are sensitive to the choice of the prior of t_0 as shown in Figure 3.6. We next compute the posterior density of g using (3.16).

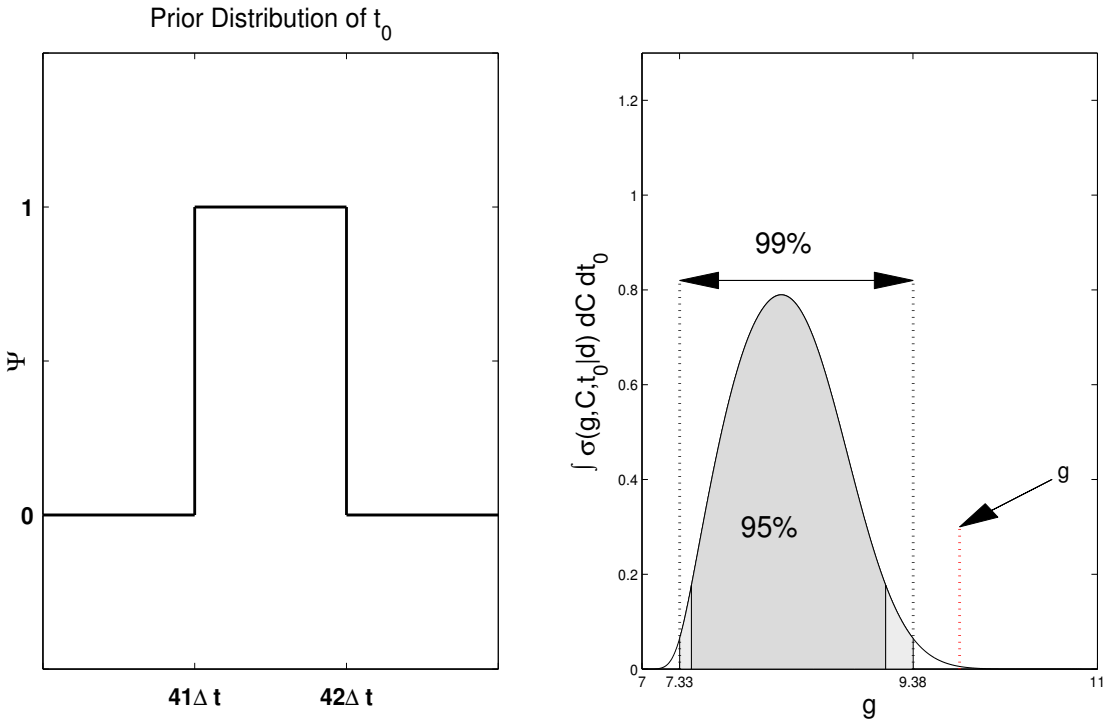

Figure 3.6: *Left*: prior distribution of t_0. *Right*: Marginal posterior distribution of g using a uniform distribution as the prior of t_0.

In Figure 3.6 , we see that the shape of the prior distribution of t_0 has a big effect in the final results. By using a uniform distribution between $41 \triangle t$ and $42 \triangle t$ as the prior distribution of t_0, neither the shortest 95%, [7.43 m/s^2 , 9.14 m/s^2], nor 99%, [7.33 m/s^2 , 9.38 m/s^2], credible intervals contain the true value of the gravity constant.

This concludes the results obtained by ALA12. In the next sections, we will consider the frequentist properties of the Bayesian credible intervals, and will also conduct a validation study of the assumptions made by ALA12.

3.6 Frequentist Coverage of the Bayesian Credible Intervals

In this section, we examine the frequentist performance of the Bayesian credible interval (e.g., [17]). To be precise, we examine the frequentist coverage probability of the Bayesian credible intervals obtained in Sections 3.4 and 3.5. We conduct simulation studies, in order to find the frequentist coverage probability of the Bayesian credible intervals. The most important part of this simulation study is how the synthetic data are generated. Recall that the data are modeled as $\mathbf{d} = \mathbf{f}(\mathbf{m}) + \epsilon$, where ϵ is the vector of measurement errors that introduces random uncertainty in the data. The random behavior of $\mathbf{d}$, for the fixed parameters g and C, is described by the probability density (3.8). Therefore, we can generate synthetic data by sampling errors from (3.6) and adding them to $f_i(\widehat{\mathbf{m}})$ (2.5) using an estimate $\widehat{\mathbf{m}}$ of $\mathbf{m}$. In Section 3.2, we explained that $\epsilon_i/\delta_i \sim 2Beta(2,2) - 1$; therefore, we can simulate errors as follows:

- generate $x \sim Beta(2,2)$.
- define $\epsilon_i = \delta_i(2x - 1)$.

After generating the errors, we use the posterior means as the estimates $\widehat{\mathbf{m}}$ of the parameters of interest to compute $\mathbf{f}(\widehat{\mathbf{m}})$. Then, we add the errors to $\mathbf{f}(\widehat{\mathbf{m}})$.

We consider four cases:

- First simulation: $g = 8.82$ m/s^2 and $C = 0.116$ m^{-1}.
- Second simulation: $g = 9.79$ m/s^2 and $C = 0.116$ m^{-1}.
- Third simulation: $g = 8.64$ m/s^2, $C = 0.106$ m^{-1}, and $t_0 = 1.395$ s.
- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.106$ m^{-1}, and $t_0 = 1.395$ s.

Using the estimates of parameters of each simulation study, we generate 1000 data sets for each simulation study by randomly generating ϵ_i as described above. First, we compute

the posterior distribution using (3.4) and (3.8), or (3.14) and (3.8) considering g and C or g, C and t_0 as the parameters in the model respectively. We then compute the marginal distribution of g in two ways using (3.12) considering only g and C as the parameters in the model, or using (3.16) considering g, C, and t_0 as the parameters in the model. We find the shortest 95% Bayesian credible interval of g for each data set. Finally, we check if these intervals contain the corresponding true value (i.e., the parameters used to generate the synthetic data). We construct Agresti-Coull confidence intervals for proportion of intervals that contain the true value (i.e., the frequentist coverage) [2]. Let $\widehat{p}$ be the sample proportion and n the sample size. Define

$$\tilde{n} = n + z_{\alpha/2}^2, \qquad \tilde{p} = \frac{n}{\tilde{n}}\widehat{p} + \frac{z_{\alpha/2}^2}{2\tilde{n}}. \tag{3.18}$$

The Agresti-Coull $100(1-\alpha)\%$ confidence interval for p is defined as

$$\tilde{p} \pm z_{\alpha/2}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}. \tag{3.19}$$

The results of the simulation studies are:

	$\mathbb{E}(g\|\mathbf{d})$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	95.13 ± 0.01	0.83 ± 0.01	94.63 ± 0.01	0.84 ± 0.01
g, C, t_0	97.12 ± 0.01	2.03 ± 0.02	97.02 ± 0.01	2.2 ± 0.02

The results show that when we consider only g and C as the parameters in the model (2.5), the frequentist coverage of the Bayesian credible interval is almost 95% as we would like. However, when t_0 is considered as another parameter in the model (2.5), the frequentist coverage increases in both cases: using the posterior mean of g as the estimate and using the true value for the gravity constant. Note that the average length of the Bayesian intervals increase when we consider t_0 as another parameter in the model. These results are based on the assumption that (3.8) is the adequate likelihood for the data. This assumption will checked in Chapter 5.

CHAPTER 4

MAXIMUM LIKELIHOOD

In Chapter 3, we summarized the approach of ALA12 to estimate the parameters g and C in the physical model (2.5) using Bayesian methods. In this chapter, we use maximum likelihood [7, 8, 11] to estimate these parameters. For the moment, we continue to assume that the likelihood (3.8) is appropriate for the data but we will check this assumption in Chapter 5.

4.1 Maximun Likelihood

We find the maximum likelihood estimates of g and C by finding the set of values that maximize the likelihood function over the same grid of points that were used to to estimate g and C in Section 3.4. The estimates are:

$$\widehat{g} = 8.76 \text{ m/s}^2 \qquad \widehat{C} = 0.11 \text{ m}^{-1} \tag{4.1}$$

We see that these estimates are vey similar to the posterior means (3.11). In the next section, we derive confidence intervals for g and C based on the asymptotic theory of maximum likelihood estimators (e.g., [3, 11, 18]).

4.2 Fisher Information Matrix and Confidence Intervals

Fisher information is a key tool in statistical inference; it is defined in the following manner: Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be a random variable, and let $f(\boldsymbol{X} \,|\mathbf{m})$ denote its probability density function parametrized by a vector $\mathbf{m} = (m_1, m_2, \ldots, m_k)$. Then, under regularity conditions, the Fisher information matrix $I(\mathbf{m})$ is defined by the $k \times k$ symmetric matrix whose ij-th entry is given by [3]:

$$I(\mathbf{m})_{i,j} = -\mathbb{E}\left[\frac{\partial^2 \ln f(\boldsymbol{X}|\mathbf{m})}{\partial m_i \partial m_j}\right] \tag{4.2}$$

However, sometimes it is not possible to compute the expected value, and one uses instead the *observed Fisher information* matrix $J(\mathbf{m})$ defined as

$$J(\mathbf{m})_{i,j} = -\sum_{k=1}^{n} \frac{\partial^2}{\partial m_i \partial m_j} \ln f(X_k|\mathbf{m}). \tag{4.3}$$

An approximate $100(1-\alpha)\%$ confidence interval for m_j is given by

$$\widehat{m}_j \pm z_{\alpha/2}\sqrt{(I(\widehat{\mathbf{m}})^{-1})_{jj}}, \tag{4.4}$$

where $z_{\alpha/2}$ is the appropriate z critical value (for instance, 1.96 for 95% confidence or 1.645 for 90% confidence). Similarly, an approximate $100(1-\alpha)\%$ confidence interval using the *observed Fisher Information* is given by

$$\widehat{m}_j \pm z_{\alpha/2}\sqrt{(J(\widehat{\mathbf{m}})^{-1})_{jj}}. \tag{4.5}$$

Note that $(J(\hat{\mathbf{m}})^{-1})_{jj}$ is the jjth entry of the inverse of the observed Fisher information matrix. This is different from taking the jj component of the observed Fisher information matrix, which is a scalar, and then inverting it.

The observed Fisher information matrix for our problem is given by

$$J(\mathbf{m}) = -\begin{bmatrix} \sum_{i=1}^{n} \frac{\partial^2}{\partial g^2} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial g \partial C} \ln f(X_i|\mathbf{m}) \\ \\ \sum_{i=1}^{n} \frac{\partial^2}{\partial C \partial g} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial C^2} \ln f(X_i|\mathbf{m}) \end{bmatrix}, \tag{4.6}$$

where $f(\boldsymbol{X}|\mathbf{m})$ is the likelihood function given by (3.8). Efron and Hinkley [5] suggested the use the observed Fisher information matrix to obtain approximate confidence intervals. We used the observed Fisher information matrix to obtain approximate 95% confidence intervals for the parameters g and C. Using (4.1), (4.5) and (4.6), we obtained the following approximate 95% confidence intervals for the parameters g and C:

$$\begin{array}{ll} g: & [7.98 \text{ m/s}^2 \text{ , } 9.54 \text{ m/s}^2] \\ C: & [0.015 \text{ m}^{-1} \text{ , } 0.2 \text{ m}^{-1}]. \end{array} \tag{4.7}$$

And the approximate 99% confidence intervals for the parameters g and C:

$$\begin{array}{ll} g: & [7.73 \text{ m/s}^2 \text{ , } 9.79 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} \text{ , } 0.24 \text{ m}^{-1}]. \end{array} \tag{4.8}$$

Since we know the true value for the gravity constant, we will only focus on this parameter. From the above results, we see that the approximate 95% confidence interval of g does not contain the true value for the gravity constant ($\approx$ 9.79 m/s^2). On the other hand, the approximate 99% confidence interval just barely contains it. The coverage and the average length of the approximate 95% confidence interval is discussed in Section 4.5.

4.3 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include t_0 as another parameter in the model (2.5). We will use maximum likelihood to estimate the parameters g, C, and t_0 with the likelihood function (3.8). We found these estimates by finding the set of values that maximize the likelihood function over the same grid of points that were used to to estimate g, C, and t_0 in Section 4.5. The results are:

$$\widehat{g} = 8.43 \text{ m/s}^2 \qquad \widehat{C} = 0.09 \text{ m}^{-1} \qquad \widehat{t_0} = 1.392 \text{ s} \tag{4.9}$$

If we compare the estimates of g given by (4.1) and (4.9), we see that the addition of t_0 did not improve the point estimate of g. In both, Bayesian and maximum likelihood methods, the addition of t_0 does not improve the estimate of g, but it makes the intervals wider. Moreover, the new estimates (3.15) and (4.9) are farther away from the true value of g ($\approx$ 9.79 m/s^2). Now, we proceed to compute approximate confidence intervals for g, C, and t_0. However, in the case of three parameter, the observed Fisher information matrix is slightly different from (4.6). The observed Fisher information matrix for three parameter is given by

$$J(\mathbf{m}) = -\begin{bmatrix} \sum_{i=1}^{n} \frac{\partial^2}{\partial g^2} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial g \partial C} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial g \partial t_0} \ln f(X_i|\mathbf{m}) \\ \sum_{i=1}^{n} \frac{\partial^2}{\partial C \partial g} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial C^2} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial C \partial t_0} \ln f(X_i|\mathbf{m}) \\ \sum_{i=1}^{n} \frac{\partial^2}{\partial t_0 \partial g} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial t_0 \partial C} \ln f(X_i|\mathbf{m}) & \sum_{i=1}^{n} \frac{\partial^2}{\partial t_0^2} \ln f(X_i|\mathbf{m}) \end{bmatrix} \quad (4.10)$$

where $f(\boldsymbol{X}|\mathbf{m})$ is the likelihood function given by (3.8). Using (4.5), (4.9) and (4.10), we obtained the following approximate 95% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [6.16 \text{ m/s}^2 \text{ , } 10.7 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} \text{ , } 0.25 \text{ m}^{-1}] \\ t_0: & [1.34 \text{ s , } 1.45 \text{ s}]. \end{array} \quad (4.11)$$

And the approximate 99% confidence intervals for the parameters g, C, and t_0:

$$\begin{array}{ll} g: & [5.46 \text{ m/s}^2 \text{ , } 11.4 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} \text{ , } 0.3 \text{ m}^{-1}] \\ t_0: & [1.32 \text{ s , } 1.46 \text{ s}]. \end{array} \quad (4.12)$$

By adding t_0 as another parameter, the approximate 95% confidence interval for g (4.11) is wider than the approximate 95% confidence interval for g obtained in Section 4.3 (4.7). We also see that the true value for the gravity constant ($\approx$ 9.79 m/s^2) lies in both the approximate 95% and 99% confidence interval of g. The coverage and the average length of the approximate 95% confidence interval of g, considering three parameters in the model, is discussed in the next section.

4.4 Checking True Coverage

Since the confidence intervals in Sections 4.3 and 4.4 are approximations, we need to check their actual coverage. In order to find the true coverage of these confidence intervals, we conduct a simulation study. We generate synthetic datasets just as the synthetic datasets generated in Section 3.6.

We consider four cases:

- First simulation: $g = 8.76$ m/s^2 and $C = 0.11$ m^{-1}.

- Second simulation: $g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.
- Third simulation: $g = 8.43$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.
- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.

Using the estimates of the parameters of each simulation study, we generate 1000 data sets for each simulation study. We then compute the approximate 95% confidence intervals for g in two ways using (4.5) and (4.6) considering g and C as the parameters in the model or using (4.5) and (4.10) considering g, C, and t_0 as the parameters in the model. Finally, we check if these approximate 95% confidence intervals for g contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute confidence intervals for the coverage of the approximate confidence intervals. The results of the simulations are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	98.02 ± 0.01	0.99 ± 0.01	97.92 ± 0.01	1.01 ± 0.01
g, C, t_0	98.41 ± 0.01	3.16 ± 0.04	98.61 ± 0.01	3.22 ± 0.04

These results show that the approximate 95% confidence intervals in Sections 4.3 and 4.4 are conservative. However, these results are based on the assumption that (3.8) is the correct likelihood for the data. By comparing these results to the results obtained in Section 3.6, we conclude that we obtain better coverage and shorter interval lengths using Bayesian techniques than maximum likelihood provided the likelihood assumption holds.

CHAPTER 5

VALIDATION ANALYSIS

All the analyses done in Chapters 3 and 4 were based on the likelihood (3.8) proposed in ALA12. However, the are some natural questions: why did ALA12 choose the likelihood (3.8)? and, do the results change dramatically if we choose a different likelihood? We need to check the sensitivity of the results to the choice of likelihood. To this end, we chose five different shapes for the likelihood and computed the posterior densities of g and their corresponding shortest 95% credible intervals as shown in Figure 5.1.

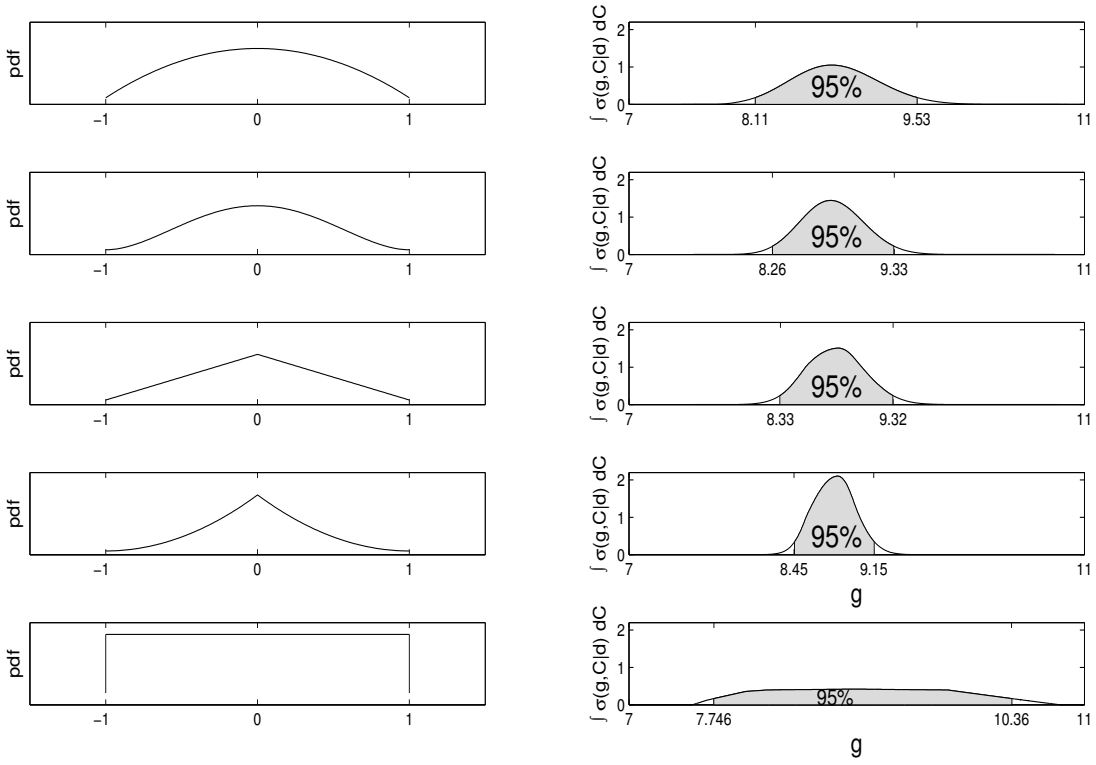

Figure 5.1: *Left*: Different shapes for the likelihood. *Right*: Shortest 95% credible intervals for g using different shapes for the likelihood (considering g and C as the only parameters in the model).

In Figure 5.1, we see that the final results depend on the shape of the likelihood. For instance, there is only one shortest 95% credible interval that contains the true value for the gravity constant (bottom right). We also see that the length of the shortest 95% credible interval varies up to a factor of 3.7. We now include t_0 as another parameter in the model and compute the posterior densities of g and their corresponding shortest 95% credible intervals as shown in Figure 5.2.

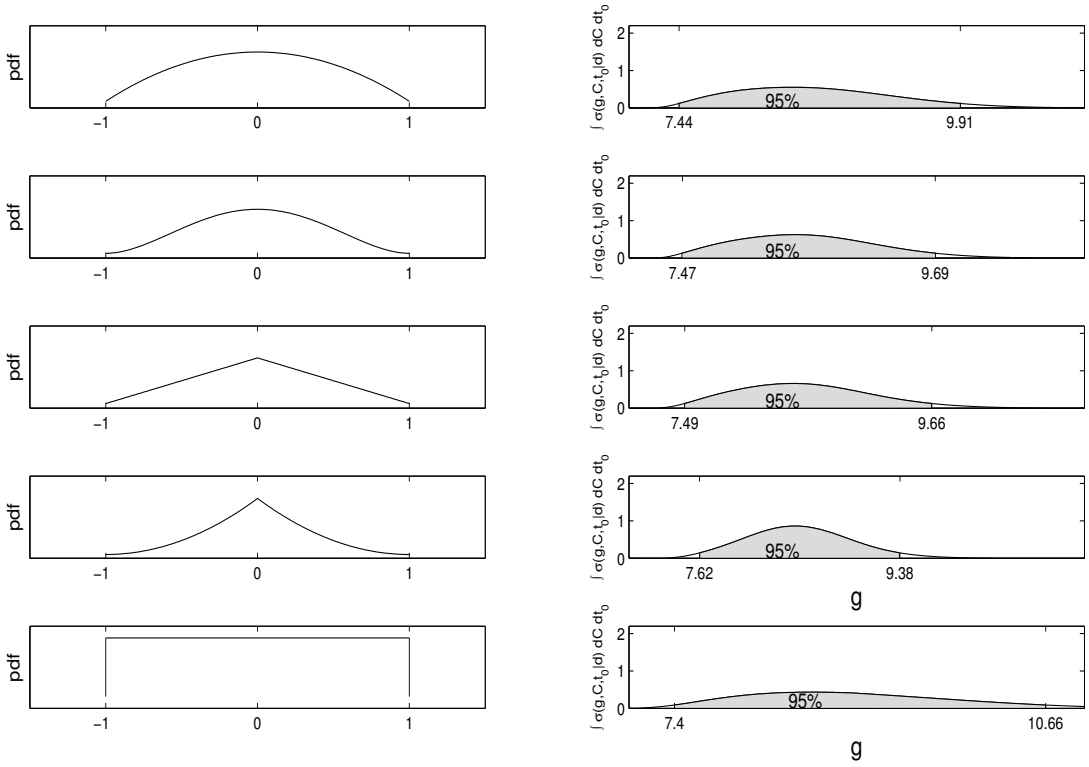

Figure 5.2: *Left*: Different shapes for the likelihood. *Right*: Shortest 95% credible intervals for g using different shapes for the likelihood (considering g, C, and t_0 the parameters in the model).

In Figure 5.2, we still see some differences in the length of the 95% credible intervals; however, the posterior densities, considering g, C and t_0 as parameters, are more alike. Now the length of the shortest 95% credible interval varies up to a factor of 1.8. We also see that there are only two 95% credible interval that contains there true value for the gravity

constant (top right and bottom right). From these results, we conclude that the final results depend on the choice of likelihood. Then, the question is why ALA12 chose the likelihood (3.8). Since, we can not tell why the likelihood (3.8) was chosen, we do not know if it is the adequate likelihood for the data set. Therefore, we need to check how well the likelihood (3.8) fits the data.

5.1 Validation of the Likelihood

Since ALA12 did not provide a justification for the choice of likelihood (3.8); we need check how well this likelihood fits the data. Recall that the data are modeled as $\mathbf{d} = \mathbf{f}(\mathbf{m}) + \epsilon$. The likelihood of the data is determined by the distribution of the corrected errors (ϵ_i/δ_i). In other words, we need to check the distribution of the corrected errors. In Figure 2.2, we see that there is a clear quadratic relationship between z (traveled distance at time time t by the body) and t. Therefore, we conduct a parametric fit and compare the cdf of corrected errors under the proposed noise distribution (3.7) to the cdf of the residuals from the model.

We thus model the data as:

$$d_i = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \epsilon_i, \quad \mathbb{E}(\epsilon_i) = 0, \quad \mathbb{V}\mathrm{ar}(\epsilon_i) = \delta_i^2 \sigma^2, \tag{5.1}$$

and find the weighted least-squares estimates of β_i by minimizing

$$\sum_{i=1}^{31} \left(\frac{d_i - \beta_0 - \beta_1 t_i - \beta_2 t_i^2}{\delta_i} \right)^2,$$

where d_i, t_i and δ_i (for $i = 1, 2, \ldots, 31$) represent the measurement, time and uncertainty in each measurement respectively as explained in Chapter 2.

Before we fit a least-squares model, we need to review the statistics of least-squares. Let the data y be modeled as

$$y = K\beta + \epsilon, \tag{5.2}$$

where K is an $n \times m$ matrix of rank $m < n$. The least-square estimate of β is

$$\hat{\beta} = \arg\min_b ||y - Kb||^2 = K^\dagger y, \tag{5.3}$$

where $K^\dagger = (K^t K)^{-1} K^t$. The residual vector of the least-squares fit is defined as

$$r = y - \hat{y} = y - K\hat{\beta} = Qy, \tag{5.4}$$

where $Q = I - KK^\dagger$ is the projection matrix onto the orthogonal complement of the column-space of K. *If the noise vector ϵ has mean zero and covariance matrix $\sigma^2 I$, then:*

- $\mathbb{E}(r) = 0$ and $\mathbb{V}\text{ar}(r) = \sigma^2 Q$.
- if ϵ is Gaussian, then $r \sim N(0, \sigma^2 Q)$.

The residuals provide an important tool to check the validity of the model but we have to account for their correlations and heteroscedasticity (e.g., [10]). Using (5.2), we know that

$$r = Qy = Q(K\beta + \epsilon)$$

Since $K\beta$ belongs to the column space of K and Q projects on its complement $QK\beta = 0$. Thus,

$$r = Q\epsilon. \tag{5.5}$$

Let UDU^t be the SVD of Q, then

$$Q = UDU^t = (U_1 U_2) \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} U_1^t \\ U_2^t \end{pmatrix} = U_1 D U_1^t, \tag{5.6}$$

where D is a $n - m$ diagonal matrix with ones in the diagonal. Then using (5.5) and (5.6), we get the following result:

$$U_1^t r = U_1^t \epsilon, \tag{5.7}$$

which shows that $U_1^t r$ and $U_1^t \epsilon$ have the same distribution. The idea is then to compare the distribution of $U_1^t r$ to the distribution of $U_1^t \epsilon$, when ϵ has the assumed distribution for the model, to see how well the assumed distribution fits the data. Figure 5.3 shows the raw data

and the fit obtained from model (5.1).

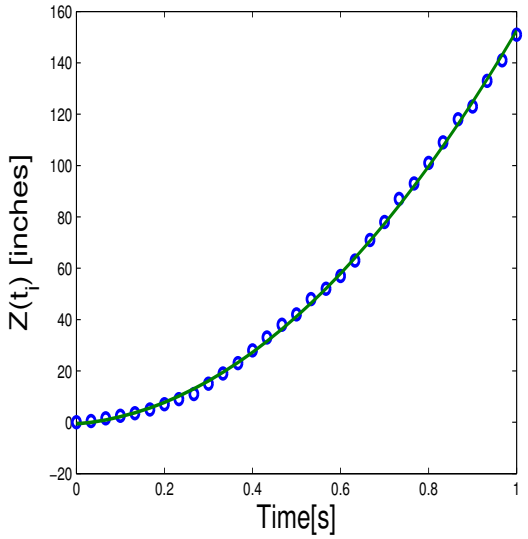

Figure 5.3: Raw data with fit obtained from model (5.1).

We next proceed to compare the distribution of $U_1^t r$, where r is the residuals from the model (5.1), to the distribution of $U_1^t \epsilon$, where $\epsilon \sim 2Beta(2,2) - 1$ as assumed in ALA12, as shown in Figure 5.4.

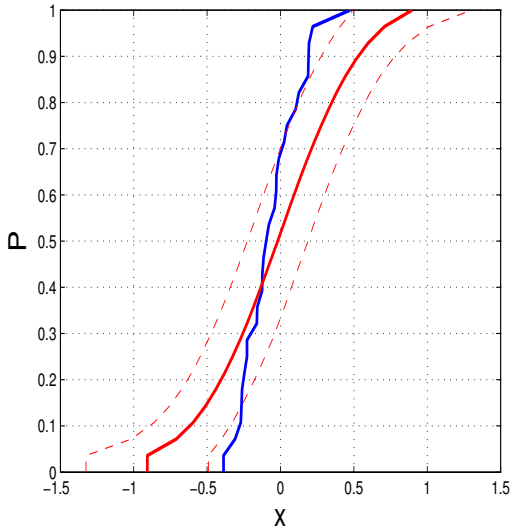

Figure 5.4: Comparison between $U_1^t r$, where r is the residuals from the model (5.1), and $U_1^t \epsilon$, where $\epsilon \sim 2Beta(2,2) - 1$.

In Figure 5.4, the blue line represents the empirical cumulative distribution of $U_1^t r$. The solid red line represent the cumulative distribution of $U_1^t \epsilon$, where $\epsilon \sim 2Beta(2,2) - 1$. The red dashed lines represent a point-wise 95% confidence interval of $U_1^t \epsilon$. We see that $U_1^t r$ does not behave as $U_1^t \epsilon$, so the assumption made by ALA12 that the corrected errors follow a $2Beta(2,2) - 1$ does not seem reasonable. We get a better fit when we assume that $\epsilon \sim \mathcal{U}[-0.5\ ,\ 0.5]$ as shown in Figure 5.5.

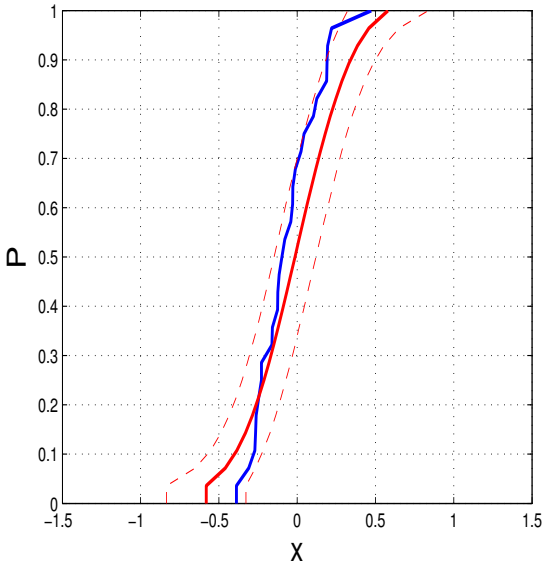

Figure 5.5: Comparison between $U_1^t r$, where r is the residuals from the model (5.1), and $U_1^t \epsilon$, where $\epsilon \sim \mathcal{U}[-0.5\ ,\ 0.5]$.

In Figure 5.5, the blue line represents the empirical cumulative distribution of $U_1^t r$. The solid red line represent the cumulative distribution of $U_1^t \epsilon$, where $\epsilon \sim \mathcal{U}[-0.5\ ,\ 0.5]$. The red dashed lines represent a point-wise 95% confidence interval of $U_1^t \epsilon$. We see that the assumption that the corrected errors follow a $\mathcal{U}[-0.5\ ,\ 0.5]$ seems more reasonable. After trying a couple of different distribution for the corrected errors, we found a distribution for the errors than fits the data even better than $\mathcal{U}[-0.5\ ,\ 0.5]$ as shown in Figure 5.6 .

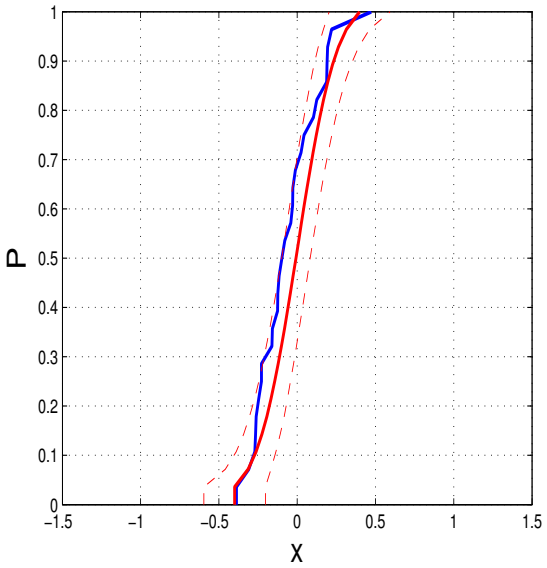

Figure 5.6: Comparison between $U_1^t r$, where r is the residuals from the model (5.1), and $U_1^t \epsilon$, where $\epsilon \sim N(0, (0.2)^2)$.

In Figure 5.6, the blue line represents the cumulative distribution of $U_1^t r$. The solid red line represents the cumulative distribution of $U_1^t \epsilon$, where $\epsilon \sim N(0, (0.2)^2)$. The standard deviation of ϵ was obtained from computing the standard deviation of $U_1^t r$. That is, $\text{std}(U_1^t r) = 0.2$. We see that the assumption that the corrected errors follow a $N(0, (0.2)^2)$ is reasonable. To be more precise with the new findings, we conduct a hypothesis test to see if the $U_1^t r$ comes from a normal distribution. The p-value of the hypothesis test is 0.3441, so at a 5% significance level we can not reject the null hypothesis. Therefore, the assumption that $U_1^t r$ comes from a Gaussian distribution is plausible. Hence, we propose a new likelihood defined by

$$\rho_i(\mathbf{d}_i|\mathbf{m}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(d_i - f_i(\mathbf{m}))^2/2\sigma^2\delta_i^2}, \tag{5.8}$$

where $\sigma = 0.2$. In the following Section, we redo the calculations from Chapters 3 and 4 using the new likelihood (5.8).

5.2 Bayesian Inversion with Gaussian Likelihood

After finding a better likelihood for the data, we need to recompute the results of Bayesian and maximum likelihood inversion using (5.8) as the likelihood. First, we start with Bayesian inversion. We compute posterior means of g and C using (3.4) as the prior and (5.8) as the likelihood over the same grid of points that was used in Section 3.4. We obtain

$$\mathbb{E}(g|\mathbf{d}) = 8.77 \text{ m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.11 \text{ m}^{-1}. \tag{5.9}$$

If we compare these results with those results obtained using ALA12's likelihood (3.11), we see that there is no major change in the posterior mean of C, however; the new posterior mean of g is farther away from true value of the gravity constant ($\approx$ 9.79 m/s^2) than the posterior mean of g given in (3.11). The new posterior density of g using (3.12) and the new shortest 95% and 99% credible intervals are shown in Figure 5.7

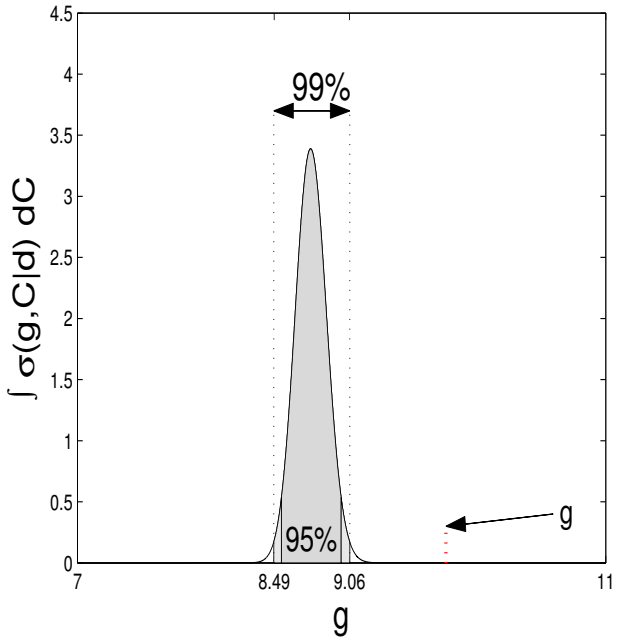

Figure 5.7: Posterior density for g using (5.8) as the likelihood (considering g and C as the only parameters in the model).

We see that neither the shortest 95% credible interval, which is [8.54 m/s^2, 8.99 m/s^2], nor the shortest 99% credible interval contain the true value of the gravity constant. Moreover,

this new shortest 95% credible interval is narrower than the shortest 95% credible interval obtained in Section 3.4. We next include the starting time (t_0) as another parameter in the model.

5.2.1 Adding a Third Parameter: Starting Time

As we did in Section 3.5, we now include t_0 as another parameter in the model. We compute the posterior means of g, C, and t_0 using (3.14) as the prior and (5.8) as the likelihood over the same grid of points that was used in Section 4.5. We obtain

$$\mathbb{E}(g|\mathbf{d}) = 8.474 \text{ m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.093 \text{ m}^{-1}, \qquad \mathbb{E}(t_0|\mathbf{d}) = 1.392 \text{ s}. \tag{5.10}$$

The posterior means of C and t_0, computed using (5.8) as the likelihood, are slightly different from the posterior means of C and t_0 obtained in Section 3.5. However, the new posterior mean of g is very different form the one obtained in Section 3.5. The posterior density of g using (3.16) and the new shortest 95% and 99% credible intervals are shown in Figure 5.8

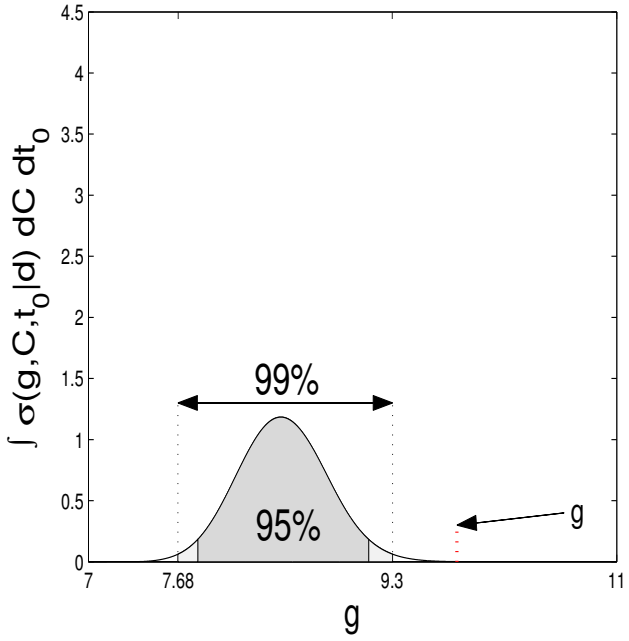

Figure 5.8: Posterior density for g using (5.8) as the likelihood and including t_0 as another parameter in the model.

We see that neither the shortest 95% credible interval, which is [7.83 m/s^2 , 9.12 m/s^2], nor the shortest 99% credible interval contain the true value of the gravity constant. On the other hand, the shortest 95% credible interval obtained in Section 3.5 contains the true value of the gravity constant. We next proceed to examine the frequentist properties of the Bayesian intervals using (5.8) as the likelihood.

5.2.2 Frequentist Coverage of the Bayesian Credible Intervals

We now proceed to examine the frequentist coverage of the Bayesian intervals obtained in Sections 5.2 and 5.2.1. We conduct four simulation studies as in Section 3.6. However, we now generate synthetic data by sampling errors from a $N(0, (0.2)^2)$ and adding them to $f_i(\widehat{\mathbf{m}})$ (2.5) using an estimate $\widehat{\mathbf{m}}$ of $\mathbf{m}$. In other words, we generate errors as follows:

- generate $\quad x \sim N(0, (0.2)^2)$
- define $\quad \epsilon_i = \delta_i x$

After generating the errors, we use the posterior means as the estimates $\widehat{\mathbf{m}}$ of the parameters of interest to compute $\mathbf{f}(\widehat{\mathbf{m}})$. Then, we add the errors to $\mathbf{f}(\widehat{\mathbf{m}})$.

- First simulation: $\quad g = 8.77$ m/s^2 and $C = 0.11$ m^{-1}.
- Second simulation: $\quad g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.
- Third simulation: $\quad g = 8.474$ m/s^2, $C = 0.093$ m^{-1}, and $t_0 = 1.392$ s.
- Fourth simulation: $\quad g = 9.79$ m/s^2, $C = 0.093$ m^{-1}, and $t_0 = 1.392$ s.

Using the estimates of the parameters of each simulation, we generate 1000 data sets for each simulation study by randomly generating ϵ_i as described above. First, we compute the posterior distribution using (3.4) and (5.8), or (3.14) and (5.8) considering g and C or g, C and t_0 as the parameters in the model respectively. We then compute the marginal distribution of g in two ways using (3.12) considering only g and C as the parameters in

the model or using (3.16) considering g, C, and t_0 as the parameters in the model. We find the shortest 95% Bayesian credible interval of g for each data set. Finally, we check if these intervals contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The simulation results are:

	$\mathbb{E}(g\|\mathbf{d})$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	95.72 ± 0.01	0.453 ± 0.001	95.62 ± 0.01	0.462 ± 0.001
g, C, t_0	94.43 ± 0.01	1.264 ± 0.006	94.53 ± 0.01	1.315 ± 0.005

These results show that when we consider only g and C as the parameters in the model (2.5), we obtain almost the same frequentist coverage of the Bayesian interval obtained in Section 3.6 using ALA12's likelihood (3.8). However, when we include t_0 as another parameter in the model (2.5), there is no improvement in the frequentist coverage of the Bayesian interval. Note that the average length of the Bayesian interval is not as wide as before (using (3.8) as the likelihood).

5.3 Maximum Likelihood with Gaussian Likelihood

Now we compute the MLE of g and C using the likelihood function (5.8). We found the MLEs by finding the set of values that maximizes the likelihood function over the same grid of points that was used in Section 3.4. The estimates are:

$$\widehat{g} = 8.77 \text{ m/s}^2 \qquad \widehat{C} = 0.11 \text{ m}^{-1}, \tag{5.11}$$

which are almost the same results we found using ALA12's likelihood (4.1). In addition, we compute approximate 95% confidence intervals for g and C using (5.11), (4.5) and (4.6). We obtain

$$\begin{array}{ll} g: & [8.54 \text{ m/s}^2 \text{ , } 9.001 \text{ m/s}^2] \\ C: & [0.083 \text{ m}^{-1} \text{ , } 0.139 \text{ m}^{-1}] \end{array} \tag{5.12}$$

And the approximate 99% confidence intervals for the parameters g, C:

$$\begin{array}{ll} g: & [8.46\ \text{m/s}^2\ ,\ 9.07\ \text{m/s}^2] \\ C: & [0.07\ \text{m}^{-1}\ ,\ 0.148\ \text{m}^{-1}] \end{array} \tag{5.13}$$

Even though the estimates found using (3.8) as the likelihood and (5.8) as the likelihood are almost the same, the approximate 95% confidence intervals are very different. The new approximate 95% confidence interval for g (5.12) is not as wide as the approximate 95% confidence interval found in Section 4.3 (4.7).

5.3.1 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include t_0 as another parameter in the model. We found the MLEs by finding the set of values that maximizes the likelihood function (5.8) over the same grid of points that was used in Section 3.5. The estimates are:

$$\widehat{g} = 8.44\ \text{m/s}^2 \qquad \widehat{C} = 0.09\ \text{m}^{-1} \qquad \widehat{t_0} = 1.392\ \text{s}. \tag{5.14}$$

Note that we get the same estimates for g, C, and t_0 using ALA12's likelihood (3.8) and the new likelihood (5.8). We then proceed to compute approximate 95% confidence intervals for the parameters using (5.14), (4.5) and (4.9). We obtain

$$\begin{array}{ll} g: & [7.76\ \text{m/s}^2\ ,\ 9.11\ \text{m/s}^2] \\ C: & [0.04\ \text{m}^{-1}\ ,\ 0.137\ \text{m}^{-1}] \\ t_0: & [1.38\ \text{s}\ ,\ 1.41\ \text{s}]. \end{array} \tag{5.15}$$

And the approximate 99% confidence intervals for the parameters g, C, and t_0:

$$\begin{array}{ll} g: & [7.55\ \text{m/s}^2\ ,\ 9.31\ \text{m/s}^2] \\ C: & [0.02\ \text{m}^{-1}\ ,\ 0.15\ \text{m}^{-1}] \\ t_0: & [1.37\ \text{s}\ ,\ 1.41\ \text{s}]. \end{array} \tag{5.16}$$

We see that the new approximate 95% confidence interval for g does not contain the true value of the gravity constant. Moreover, this new approximate 95% confidence interval is not as wide as the approximate 95% confidence interval found in Section 4.4 (4.11).

5.3.2 Checking True Coverage

Since the confidence intervals of Sections 5.3 and 5.3.1 are approximations, we need to find their true coverage. We generate synthetic datasets just as the synthetic datasets that

were generated in Section 5.2.2. We consider four cases as in Section 5.2.2.

- First simulation: $g = 8.77$ m/s^2 and $C = 0.11$ m^{-1}.
- Second simulation: $g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.
- Third simulation: $g = 8.44$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.
- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.

We generate 1000 data sets for each simulation study. We then compute the approximates 95% confidence intervals for g in two ways using (4.5) and (4.6) considering g and C as the parameters in the model or using (4.5) and (4.10) considering g, C, and t_0 as the parameters in the model. Finally, we check if these approximate 95% confidence intervals for g contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The results of the simulation studies are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	94.83 ± 0.01	0.460 ± 0.001	94.83 ± 0.01	0.470 ± 0.001
g, C, t_0	94.33 ± 0.01	1.326 ± 0.005	94.04 ± 0.01	1.365 ± 0.005

We see that the approximate 95% confidence intervals computed in Sections 5.3 and 5.3.1 have almost the target coverage. If we compare these results (considering g and C as the parameters) to the results obtained in Section 5.2.2, we see that we obtain slightly better coverage using maximum likelihood than Bayesian techniques. However, If we compare these results (considering g, C, t_0 as the parameters) to the results obtained in Section 5.2.2, we see that we obtain slightly better coverage using Bayesian techniques than maximum likelihood.

5.4 Confidence Intervals Using Fisher Information Matrix

In Sections 5.3 and 5.3.1, we approximated confidence intervals for the parameters of interest using the observed Fisher information matrix. However, we can also approximate

confidence intervals for the parameters of interest using the Fisher information matrix (4.2). We compute approximate 95% confidence intervals for g and C using (5.11), (4.2) and (4.4). We obtain

$$\begin{array}{ll} g: & [8.54\ \mathrm{m/s^2}\ ,\ 8.99\ \mathrm{m/s^2}] \\ C: & [0.082\ \mathrm{m^{-1}}\ ,\ 0.138\ \mathrm{m^{-1}}] \end{array} \tag{5.17}$$

And the approximate 99% confidence intervals for the parameters g, C:

$$\begin{array}{ll} g: & [8.47\ \mathrm{m/s^2}\ ,\ 9.07\ \mathrm{m/s^2}] \\ C: & [0.07\ \mathrm{m^{-1}}\ ,\ 0.147\ \mathrm{m^{-1}}] \end{array} \tag{5.18}$$

From these results, we see that these approximate confidence intervals are slightly narrower than the approximates confidence intervals obtained in Section 5.3 ((5.12) and (5.13)).

5.4.1 Adding a Third Parameter: Starting Time

We now include t_0 as another parameter in the model. We compute approximate 95% confidence intervals for the parameter using (5.14), (4.2) and (4.4). We obtain

$$\begin{array}{ll} g: & [7.77\ \mathrm{m/s^2}\ ,\ 9.10\ \mathrm{m/s^2}] \\ C: & [0.04\ \mathrm{m^{-1}}\ ,\ 0.137\ \mathrm{m^{-1}}] \\ t_0: & [1.38\ \mathrm{s}\ ,\ 1.40\ \mathrm{s}]. \end{array} \tag{5.19}$$

And the approximate 99% confidence intervals for the parameters g, C, and t_0:

$$\begin{array}{ll} g: & [7.57\ \mathrm{m/s^2}\ ,\ 9.31\ \mathrm{m/s^2}] \\ C: & [0.03\ \mathrm{m^{-1}}\ ,\ 0.15\ \mathrm{m^{-1}}] \\ t_0: & [1.37\ \mathrm{s}\ ,\ 1.41\ \mathrm{s}]. \end{array} \tag{5.20}$$

From these results, we see that these approximate confidence intervals are slightly narrower than the approximates confidence intervals obtained in Section 5.3.1 ((5.15) and (5.16)).

5.4.2 Checking True Coverage of Confidence Intervals Using Fisher Information Matrix

As we did in Section 5.3.2, we need to check the true coverage of the approximate confidence intervals using Fisher information matrix. We use the same simulated data sets generated in Section 5.3.2, but we now compute approximate 95% confidence intervals for g using (4.2) and (4.4). Finally, we check if these approximate 95% confidence intervals contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true

coverage of the approximate confidence intervals. The results of the simulation studies are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	94.83 ± 0.01	0.460 ± 0.001	94.83 ± 0.01	0.470 ± 0.001
g, C, t_0	94.4 ± 0.01	1.32 ± 0.005	94.2 ± 0.01	1.365 ± 0.005

From these results, we see that the true coverage of the approximate 95% confidence interval of g, using the Fisher information matrix, is almost equal the target coverage. If we compare these results with the results obtained in Section 5.3.2, we see that we obtain slightly better coverage using the Fisher information matrix instead of the observed Fisher information matrix.

5.5 Analysis without the δ_i's

In Section 5.1, we showed that ALA12's assumption was not reasonable. Moreover, we proposed a new likelihood for the data (5.8), where $\epsilon_i/\delta_i \sim N(0, 0.2^2)$. However, we always want to make our model as simple as possible. Then, the question is if the uncertainty in each measurement, as explained in Chapter 2 (δ_i for $i = 1, 2, \ldots, 31$), should or should not be included. By excluding δ_i's from the analysis, we would not have to estimate the δ_i's from each video frame. Therefore, we conduct the same analysis as done in Section 5.1 but without the δ_i's. Now, we model the data as:

$$d_i = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \epsilon_i, \quad \mathbb{E}(\epsilon_i) = 0, \quad \mathbb{V}\text{ar}(\epsilon_i) = \sigma^2, \tag{5.21}$$

and this time we minimize

$$\sum_{i=1}^{31} \left(d_i - \beta_0 - \beta_1 t_i - \beta_2 t_i^2\right)^2,$$

where d_i, and t_i represent the measurement and time as explained in Chapter 2. We then proceed to fit the least-squares model. Figure 5.9 shows the raw data with the fit obtained from model (5.21).

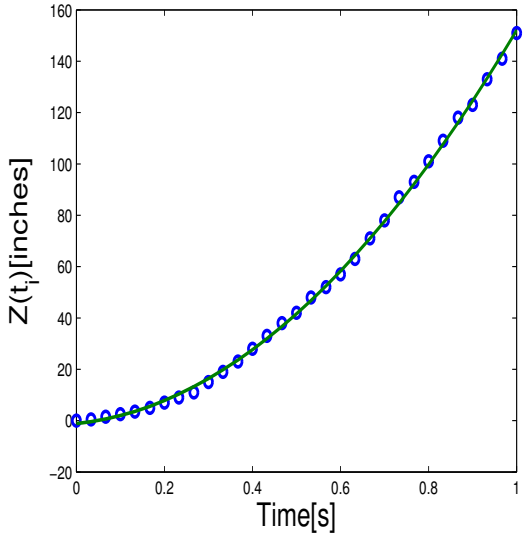

Figure 5.9: Raw data with fit obtained from model (5.21).

If we compare the fits obtained from model (5.1) with (5.21), we see that both fits fit the data well. From Section 5.1, we know that $\epsilon_i/\delta_i \sim N(0, 0.2^2)$; therefore, we now want to check if ϵ_i still follows a Gaussian distribution. As we did in Section 5.1, we next proceed to compare the distribution of $U_1^t r$, where r is the residuals from the model (5.21), to the distribution of $U_1^t \epsilon$, where ϵ follows a Gaussian distribution, as shown in Figure 5.10.

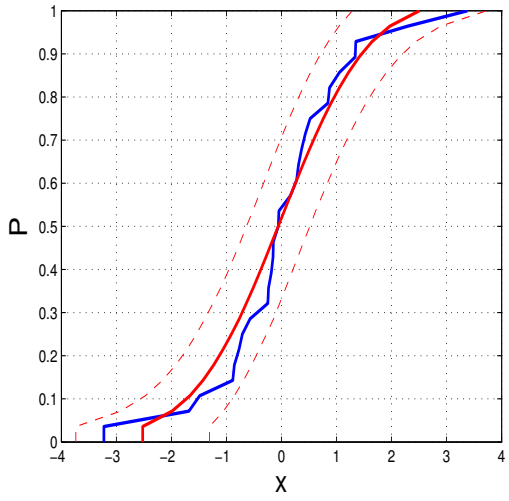

Figure 5.10: Comparison between $U_1^t r$, where r is the residuals from the model (5.21), and $U_1^t \epsilon$, where $\epsilon \sim N(0, (0.1.25)^2)$.

In Figure 5.10, the blue line represents the cumulative distribution of $U_1^t r$. The solid red line represents the cumulative distribution of $U_1^t \epsilon$, where $\epsilon \sim N(0, (0.1.25)^2)$. The standard deviation of ϵ was obtained from computing the standard deviation of $U_1^t r$. That is, $\text{std}(U_1^t r) = 1.25$. We see that the assumption that the ϵ_i follow a $N(0, (0.125)^2)$ is reasonable. To be more precise with the new findings, we conduct a hypothesis test to see if the $U_1^t r$ comes from a normal distribution. The p-value of the hypothesis test is 0.3225, so at a 5% significance level we can not reject the null hypothesis. Therefore, the assumption that $U_1^t r$ comes from a Gaussian distribution is plausible. Hence, we can propose another likelihood defined by

$$\rho_i(\mathbf{d}_i|\mathbf{m}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(d_i - f_i(\mathbf{m}))^2/2\sigma^2}, \tag{5.22}$$

where $\sigma = 0.125$. In the following section, we redo the calculations from Sections 5.2 and 5.3 using the new likelihood (5.22) to see if we obtain different results.

5.5.1 Bayesian Inversion without δ_i's

First, we start with Bayesian inversion. We compute posterior means of g and C using (3.4) as the prior and (5.22) as the likelihood over the same grid of points that was used in Section 3.4. We obtain

$$\mathbb{E}(g|\mathbf{d}) = 8.77 \text{ m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.11 \text{ m}^{-1}. \tag{5.23}$$

If we compare these results with the results obtained in Section 5.2 (5.9), we see that we obtain the same posterior means using (5.8) as the likelihood or (5.22) as the likelihood. Then, we compute the new posterior density of g using (3.12) and the new shortest 95% and 99% credible intervals are shown in Figure 5.11

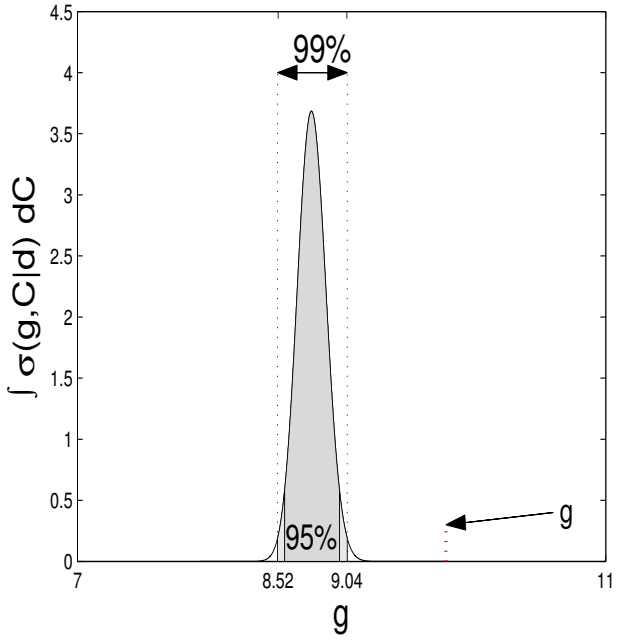

Figure 5.11: Posterior density for g using (5.22) as the likelihood (considering g and C as the only parameters in the model).

We see that neither the shortest 95% credible interval, which is $[8.57\ \text{m/s}^2, 8.98\ \text{m/s}^2]$, nor the shortest 99% credible interval contain the true value of the gravity constant. Moreover, these new shortest 95% and 99% credible intervals are narrower than the shortest 95% and 99% credible intervals obtained in Section 5.2. We next include the starting time (t_0) as another parameter in the model.

5.5.2 Adding a Third Parameter: Starting Time

As we did in Section 3.5, we now include t_0 as another parameter in the model. We compute the posterior means of g, C, and t_0 using (3.14) as the prior and (5.22) as the likelihood over the same grid of points that was used in Section 4.5. We obtain

$$\mathbb{E}(g|\mathbf{d}) = 8.478\ \text{m/s}^2, \qquad \mathbb{E}(C|\mathbf{d}) = 0.094\ \text{m}^{-1}, \qquad \mathbb{E}(t_0|\mathbf{d}) = 1.392\ \text{s}. \tag{5.24}$$

The posterior means of g, C and t_0, computed using (5.22) as the likelihood, are slightly different from the posterior means of C and t_0 obtained in Section 5.2.1. The posterior

density of g using (3.16) and the new shortest 95% and 99% credible intervals of g are shown in Figure 5.12

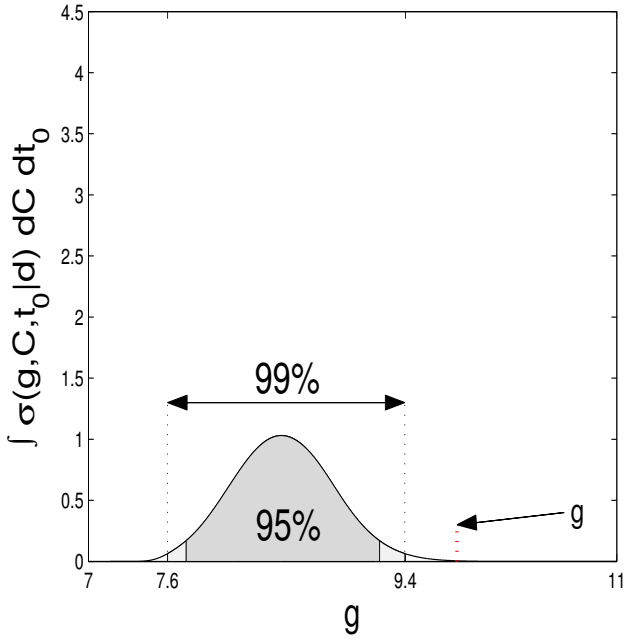

Figure 5.12: Posterior density for g using (5.22) as the likelihood and including t_0 as another parameter in the model.

We see that neither the shortest 95% credible interval, which is [7.74 m/s^2 , 9.20 m/s^2], nor the shortest 99% credible interval contain the true value of the gravity constant. We also see that the shortest 95% and 99% credible intervals are slightly wider than the shortest 95% and 99% credible intervals obtained in Section 5.2.1.

5.5.3 Frequentist Coverage of the Bayesian Credible Intervals without δ_i's

We now proceed to examine the frequentist coverage of the Bayesian intervals obtained in Sections 5.5.1 and 5.5.2. We conduct four simulation studies as in Section 5.2.2. However, We now generate synthetic data by sampling errors from a $N(0, (0.125)^2)$ and adding them to $f_i(\widehat{\mathbf{m}})$ (2.5) using an estimate $\widehat{\mathbf{m}}$ of $\mathbf{m}$. In other words, we generate errors as follows:

- generate $\quad x \sim N(0, (0.125)^2)$

- define $\epsilon_i = x$

After generating the errors, we use the posterior means as the estimates $\widehat{\mathbf{m}}$ of the parameters of interest to compute $\mathbf{f}(\widehat{\mathbf{m}})$. Then, we add the errors to $\mathbf{f}(\widehat{\mathbf{m}})$.

- First simulation: $g = 8.77$ m/s^2 and $C = 0.11$ m^{-1}.
- Second simulation: $g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.
- Third simulation: $g = 8.478$ m/s^2, $C = 0.094$ m^{-1}, and $t_0 = 1.392$ s.
- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.094$ m^{-1}, and $t_0 = 1.392$ s.

Using the estimates of the parameters of each simulation, we generate 1000 data sets for each simulation study by randomly generating ϵ_i as described above. First, we compute the posterior distribution using (3.4) and (5.22), or (3.14) and (5.22) considering g and C or g, C and t_0 as the parameters in the model respectively. We then compute the marginal distribution of g in two ways using (3.12) considering only g and C as the parameters in the model or using (3.16) considering g, C, and t_0 as the parameters in the model. We find the shortest 95% Bayesian credible interval of g for each data set. Finally, we check if these intervals contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The simulation results are:

	$\mathbb{E}(g\|\mathbf{d})$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	95.33 ± 0.01	0.416 ± 0.001	95.33 ± 0.01	0.424 ± 0.001
g, C, t_0	94.60 ± 0.01	1.393 ± 0.009	94.20 ± 0.01	1.478 ± 0.008

If we compare these results with the results obtained in Section 5.2.2, we see that we obtain slightly better frequentist coverage of the Bayesian interval when we use (5.22) as

likelihood (excluding the δ_i's). When we consider only g and C as the parameters, the average length of the Bayesian intervals is slightly narrower than the average of the Bayesian intervals obtained when we include the δ_i's in the analysis. However, when we include t_0 as another parameter in the model (2.5), the average length of the Bayesian intervals increases.

5.6 Maximum Likelihood without δ_i's

Now we compute the MLE of g and C using the new likelihood function, without δ_i's (5.22), to see if we obtain different results. We find the MLEs by finding the set of values that maximizes the likelihood function over the same grid of points that was used in Section 3.4. The estimates are:

$$\widehat{g} = 8.77 \text{ m/s}^2 \qquad \widehat{C} = 0.11 \text{ m}^{-1}, \tag{5.25}$$

which are the same estimates that we obtained using (5.8) as likelihood. In addition, we compute approximate 95% confidence intervals for g and C using (5.25), (4.5) and (4.6). We obtain

$$\begin{array}{ll} g: & [8.56 \text{ m/s}^2 \text{ , } 8.98 \text{ m/s}^2] \\ C: & [0.085 \text{ m}^{-1} \text{ , } 0.136 \text{ m}^{-1}]. \end{array} \tag{5.26}$$

And the approximate 99% confidence intervals for the parameters g, C:

$$\begin{array}{ll} g: & [8.49 \text{ m/s}^2 \text{ , } 9.05 \text{ m/s}^2] \\ C: & [0.07 \text{ m}^{-1} \text{ , } 0.144 \text{ m}^{-1}]. \end{array} \tag{5.27}$$

Even though we obtain the same estimates for g and C using (5.8) or (5.22) as the likelihood, the approximate 95% confidence intervals are slightly different. The new approximate 95% and 99% approximate confidence intervals are narrower than the approximate confidence intervals obtained in Section 5.3. We discuss the coverage of these intervals in Section 5.6.2. We next include t_0 as another parameter in the model.

5.6.1 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include t_0 as another parameter in the model. We found the MLEs by finding the set of values that maximizes the likelihood function (5.22)

over the same grid of points that was used in Section 3.5. The estimates are:

$$\widehat{g} = 8.41 \text{ m/s}^2 \qquad \widehat{C} = 0.09 \text{ m}^{-1} \qquad \widehat{t_0} = 1.391 \text{ s} \tag{5.28}$$

Note that these new estimates for g, C, and t_0 using (5.18) as the likelihood are slightly different from the estimates found in Section 5.3.1. We then proceed to compute approximate 95% confidence intervals for the parameters using (5.24), (4.5) and (4.9). We obtain

$$\begin{array}{ll} g: & [7.66 \text{ m/s}^2 \text{ , } 9.16 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} \text{ , } 0.139 \text{ m}^{-1}] \\ t_0: & [1.37 \text{ s , } 1.41 \text{ s}]. \end{array} \tag{5.29}$$

And the approximate 99% confidence intervals for the parameters g, C, and t_0:

$$\begin{array}{ll} g: & [7.43 \text{ m/s}^2 \text{ , } 9.39 \text{ m/s}^2] \\ C: & [0.02 \text{ m}^{-1} \text{ , } 0.15 \text{ m}^{-1}] \\ t_0: & [1.36 \text{ s , } 1.42 \text{ s}]. \end{array} \tag{5.30}$$

We see that neither the new approximate 95% confidence interval nor the new approximate 99% confidence interval for g do not contain the true value of the gravity constant. Moreover, these new approximate 95% and 99% confidence intervals are wider than the approximate 95% and 99% confidence intervals obtained in Section 5.3.1. In the next section, we discuss the coverage of these intervals.

5.6.2 Checking True Coverage

Since the confidence intervals of Sections 5.6 and 5.6.1 are approximations, we need to find the true coverage of those intervals. We consider four cases as in Section 5.4.3.

- First simulation: $g = 8.77 \text{ m/s}^2$ and $C = 0.11 \text{ m}^{-1}$.
- Second simulation: $g = 9.79 \text{ m/s}^2$ and $C = 0.11 \text{ m}^{-1}$.
- Third simulation: $g = 8.41 \text{ m/s}^2$, $C = 0.09 \text{ m}^{-1}$, and $t_0 = 1.391$ s.
- Fourth simulation: $g = 9.79 \text{ m/s}^2$, $C = 0.09 \text{ m}^{-1}$, and $t_0 = 1.391$ s.

We generate synthetic datasets just as the synthetic datasets that were generated in Section 5.4.3. We generate 1000 data sets for each simulation study. We then compute the approximates 95% confidence intervals for g in two ways using (4.5) and (4.6) considering g and C as the parameters in the model or using (4.5) and (4.10) considering g, C, and t_0 as the parameters in the model. Finally, we check if these approximate 95% confidence intervals for g contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The results of the simulation studies are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	94.40 ± 0.01	0.422 ± 0.001	94.40 ± 0.01	0.431 ± 0.001
g, C, t_0	93.60 ± 0.01	1.572 ± 0.008	94.00 ± 0.01	1.604 ± 0.007

If we compare these results with the results obtained in Section 5.3.2, we see that we obtain almost the same coverage with and without the δ_i's in the likelihood. Note that the average length of the approximate 95% confidence intervals (g and C as the parameters) is narrower than the average length of the approximate 95% confidence intervals obtained in Section 5.3.2. However, when t_0 is include as another parameter, the average length of the approximate 95% confidence intervals is wider than the average length of the approximate 95% confidence intervals obtained in Section 5.3.2.

5.7 Confidence Intervals Using Fisher Information Matrix without δ_i's

As we did in Section 5.4, we now proceed to approximate confidence intervals for the parameters using Fisher information matrix (4.2). Using (5.25), (4.2) and (4.4), we approximate 95% confidence intervals for g and C. We obtain

$$\begin{array}{ll} g: & [8.56 \text{ m/s}^2 \text{ , } 8.98 \text{ m/s}^2] \\ C: & [0.085 \text{ m}^{-1} \text{ , } 0.135 \text{ m}^{-1}]. \end{array} \tag{5.31}$$

And the approximate 99% confidence intervals for the parameters g, C:

$$\begin{array}{ll} g: & [8.49 \text{ m/s}^2 \text{ , } 9.05 \text{ m/s}^2] \\ C: & [0.08 \text{ m}^{-1} \text{ , } 0.143 \text{ m}^{-1}]. \end{array} \tag{5.32}$$

From these results, we see that we obtain the same approximate confidence intervals for g, but the approximate confidence intervals for C are slightly narrower than the approximate confidence intervals obtained in Section 5.6.

5.7.1 Adding a Third Parameter: Starting Time

We now proceed to include t_0 as another parameter in the model. We compute approximate 95% confidence intervals for the parameter using (5.28), (4.2) and (4.4). We obtain

$$\begin{array}{ll} g: & [7.63 \text{ m/s}^2 \text{ , } 9.20 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} \text{ , } 0.14 \text{ m}^{-1}] \\ t_0: & [1.37 \text{ s} \text{ , } 1.41 \text{ s}]. \end{array} \tag{5.33}$$

And the approximate 99% confidence intervals for the parameters g, C, and t_0:

$$\begin{array}{ll} g: & [7.38 \text{ m/s}^2 \text{ , } 9.44 \text{ m/s}^2] \\ C: & [0.02 \text{ m}^{-1} \text{ , } 0.16 \text{ m}^{-1}] \\ t_0: & [1.36 \text{ s} \text{ , } 1.42 \text{ s}]. \end{array} \tag{5.34}$$

From these results, we see that these approximate confidence intervals are wider than the approximates confidence intervals obtained in section 5.6.1 ((5.29) and (5.30)).

5.7.2 Checking True Coverage of Confidence Intervals Using Fisher Information Matrix without δ_i's

As we did in Section 5.4.2, we need to check the true coverage of the approximate confidence intervals using Fisher information matrix without δ_i's. We use the same simulated data sets generated in Section 5.6.2, but we now compute approximate 95% confidence intervals for g using (4.2) and (4.4). Finally, we check if these approximate 95% confidence intervals contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The results of the simulation studies are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	94.60 ± 0.01	0.422 ± 0.001	94.40 ± 0.01	0.431 ± 0.001
g, C, t_0	93.70 ± 0.01	1.57 ± 0.007	94.00 ± 0.01	1.60 ± 0.006

From these results, we see that the true coverage of the approximate confidence interval of g, using the Fisher information matrix, is almost equal the target coverage. If we compare these results with the results obtained in Section 5.6.2, we see that we obtain slightly better coverage using the Fisher information matrix instead of the observed Fisher information matrix. We next proceed to estimate the parameters of interest using nonlinear regression techniques.

CHAPTER 6

NONLINEAR REGRESSION

Recall the equation for a falling body (2.5):

$$z(t) = \tfrac{1}{C}\log\cosh[\sqrt{gC}(t - t_0)],$$

where t is the time, t_0 is the starting time, $z(t)$ is the traveled distance at time t, g and C are the parameters of interest. In Chapters 3 and 4, we estimated the parameters of interest using Bayesian methods and maximum likelihood respectively. These methods require, in particular, the model (2.5) and a likelihood. An alternative approach to estimate the parameters of interest using only the model (2.5) is to carry out a nonlinear regression analysis.

6.1 Nonlinear Least Squares

Suppose that we have n observations (x_i, y_i), $i = 1, 2, \ldots, n$, from a fixed-regressor nonlinear model with a known functional relationship f. Thus,

$$y_i = f(x_i; \boldsymbol{\theta}) + \epsilon_i \qquad (i = 1, 2, \ldots, n), \tag{6.1}$$

where the ϵ_i are i.i.d. $N(0, \sigma^2)$, and $\boldsymbol{\theta}$ is the vector of p parameters. The least-square estimate of $\boldsymbol{\theta}$, denoted by $\widehat{\boldsymbol{\theta}}$, minimizes the residual sum of squares

$$S(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} [y_i - f(x_i; \widehat{\boldsymbol{\theta}})]^2 \tag{6.2}$$

In our case $\epsilon_i \sim N(0, \sigma_i^2)$. So, we use nonlinear weighted least-squares

$$S_w(\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} \left[\frac{y_i}{\delta_i} - \frac{f(x_i; \widehat{\boldsymbol{\theta}})}{\delta_i} \right]^2, \tag{6.3}$$

where δ_i is the uncertainty of the ith measurement as explained in Chapter 2. The weighted least-square estimate of $\boldsymbol{\theta}$, denoted by $\widehat{\boldsymbol{\theta}}$, minimizes (6.3). A difficulty with nonlinear least square arises in trying to solve for $\widehat{\boldsymbol{\theta}}$ that minimizes (6.3).

6.2 Computational Methods for Nonlinear Least Squares

Iterative numerical methods are used to solve for $\widehat{\boldsymbol{\theta}}$. These methods require initial guesses, or starting values, for the parameters of interest; the starting values are labeled as $\boldsymbol{\theta}^0$. The starting values are substituted for $\boldsymbol{\theta}$ to compute the residual sum of squares and to compute the adjustment to $\boldsymbol{\theta}^0$ that will reduce the sum of squares and move $\boldsymbol{\theta}^0$ closer to the least squares solution. The most common methods to find the least square solution are Gauss-Newton Method, the method of steepest descent, and Levenberg-Marquardt. We next briefly describe how these methods work. For a more detailed explanation and implementation of these methods see [12, 14, 19].

6.2.1 Gauss-Newton Method

The Gauss-Newton method uses a Taylor's expansion of $f(x_i;\boldsymbol{\theta})$ about the starting value $\boldsymbol{\theta}^0$ to obtain a linear approximation of the model in the region near the starting values. That is, $f(x_i;\boldsymbol{\theta})$ is replaced by

$$f(x_i;\boldsymbol{\theta}) \doteq f(x_i;\boldsymbol{\theta}^0) + \sum_{j=1}^{p} \left(\frac{\partial f(x_i;\boldsymbol{\theta}^0)}{\partial \theta_j} \right) (\theta_j - \theta_j^0) \tag{6.4}$$

or

$$f(\boldsymbol{\theta}) \doteq f(\boldsymbol{\theta}^0) + F(\boldsymbol{\theta}^0)(\boldsymbol{\theta} - \boldsymbol{\theta}^0), \tag{6.5}$$

where $F(\boldsymbol{\theta}^0)$ is the $n \times p$ matrix of partial derivatives, evaluated at $\boldsymbol{\theta}^0$ and the n data points x_i. The matrix $F(\boldsymbol{\theta}^0)$ has the form

$$F(\boldsymbol{\theta}^0) = \begin{bmatrix} \frac{\partial [f(x_1;\boldsymbol{\theta}^0)]}{\partial \theta_1} & \frac{\partial [f(x_1;\boldsymbol{\theta}^0)]}{\partial \theta_2} & \cdots & \frac{\partial [f(x_1;\boldsymbol{\theta}^0)]}{\partial \theta_p} \\ \frac{\partial [f(x_2;\boldsymbol{\theta}^0)]}{\partial \theta_1} & \frac{\partial [f(x_2;\boldsymbol{\theta}^0)]}{\partial \theta_2} & \cdots & \frac{\partial [f(x_2;\boldsymbol{\theta}^0)]}{\partial \theta_p} \\ \vdots & \vdots & & \vdots \\ \frac{\partial [f(x_n;\boldsymbol{\theta}^0)]}{\partial \theta_1} & \frac{\partial [f(x_n;\boldsymbol{\theta}^0)]}{\partial \theta_2} & \cdots & \frac{\partial [f(x_n;\boldsymbol{\theta}^0)]}{\partial \theta_p} \end{bmatrix}. \tag{6.6}$$

Linear least squares is used on the linearized model (6.5) to estimate the shift in the parameters; that is,the update on the starting values. New values of the parameters are obtained by adding the estimated shift to the starting values. Then, the model is linearized about the

new values of the parameters and linear least squares is again applied to find the second set of adjustments; the process continues the same way until convergence is satisfied.

6.2.2 The Method of Steepest Descent

This is a method that finds the path to update the initial estimates of the parameters, so that it achieves the largest decrease in the residual sum of squares (as approximated by the linearization (6.5)). After each update in the parameters values, the residual sum of squares is approximated again in the vicinity of the new solution and a new steepest descent path is determined.

6.2.3 Levenberg-Marquadt Methods

This third class of methods is designed to capitalize on the best features of the Gauss-Newton method and the steepest descent methods. The adjustment computed by Levenberg-Marquadt's method tends toward the Gauss-Newton adjustment if the residual sum of squares is reduced at each step, and toward the method of steepest descent adjustment if the residual sum of squares increases in any step.

6.3 Confidence Intervals for Nonlinear Models

Confidence intervals for parameters in a nonlinear model are based on the approximate distribution of the nonlinear least squares estimator. The familiar properties of liner least squares apply only approximately to nonlinear least squares. It has been shown that if $\epsilon \sim N(0, \sigma^2 I)$, then

$$\widehat{\boldsymbol{\theta}} \approx N(\boldsymbol{\theta}\ ,\ (F'F)^{-1}\sigma^2), \tag{6.7}$$

where the symbol " $\approx$ " is read " approximately disturbuted ", F is the $n \times p$ matrix of partial derivatives (6.6)(e.g., [7]) . In practice, $F(\boldsymbol{\theta})$ is approximated as $F(\widehat{\boldsymbol{\theta}})$, which is labeled as $\widehat{F}$ for brevity, and σ^2 is estimated with $s^2 = S(\widehat{\boldsymbol{\theta}})/(n-p)$, so that the estimated variance-covariance matrix for $\widehat{\boldsymbol{\theta}}$ is

$$s^2(\widehat{\boldsymbol{\theta}}) = s^2(\widehat{F}'\widehat{F})^{-1}. \tag{6.8}$$

The approximate $100(1-\alpha)\%$ confidence interval estimate of θ_j is

$$\widehat{\theta}_j \pm t_{[\alpha/2,(n-p)]} \left[s^2((\widehat{F}'\widehat{F})^{-1})_{jj} \right]^{1/2}, \tag{6.9}$$

where $t_{[\alpha/2,(n-p)]}$ is the upper $\alpha/2$ percentage point of the Student's t distribution with $(n-p)$ degrees of freedom. Note that if we replace $t_{[\alpha/2,(n-p)]}$ by $z_{\alpha/2}$ and s^2 by σ^2, we obtain the same confidence interval as the one given in (4.5). A confidence region is a generalization of a confidence interval in $\mathbb{R}^n$. Often represented as an ellipsoid in $\mathbb{R}^3$ around the estimated solution to a problem, however, other shapes can be used.

6.3.1 Confidence Regions

The approximate $100(1-\alpha)\%$ confidence region for all p parameters in $\boldsymbol{\theta}$ is defined as a set of $\boldsymbol{\theta}$ for which

$$S(\boldsymbol{\theta}) - S(\widehat{\boldsymbol{\theta}}) \leq ps^2 F_{(\alpha\,;\,p\,,\,\nu)}, \tag{6.10}$$

where $S(\widehat{\boldsymbol{\theta}})$ is the residual sum of squares, $F_{(\alpha\,;\,p\,,\,\nu)}$ is the upper α percentage point of the F-distribution with p and ν degrees of freedom. Solving the inequality (6.10) leads to a p-dimensional ellipsoid that defines an approximate $100(1-\alpha)\%$ confidence region for all the parameters in the model.

6.4 First Results

We conduct a weighted nonlinear regression using the Levenberg-Maquadt algorithm. First, we only consider g and C as the only parameters in the model (2.5). We obtain the following results:

$$\widehat{g} = 8.77 \text{ m/s}^2, \qquad \widehat{C} = 0.11 \text{ m}^{-1}. \tag{6.11}$$

Note that these estimates coincide with the ML estimates (5.11) (using (5.8) as the likelihood). The reason why these estimates coincide is that when the $\epsilon_i \sim N(0, \sigma^2)$, the least-squares estimate is also the maximum likelihood estimate. Here is the explanation why the maximum likelihood estimate and the least-squares estimate coincide when we have normal

distributed data. Let $\mathcal{L}$ be the likelihood function

$$\mathcal{L} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2\sigma^2}||\mathbf{d}-f(\mathbf{m})||^2},$$

where $\mathbf{d}$ is the data, and $\mathbf{m}$ is the parameter that we want to estimate. By taking the log of the likelihood function, we obtain

$$\log(\mathcal{L}) = n\log\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2\sigma^2}||\mathbf{d} - f(\mathbf{m})||^2$$

In order to maximize the likelihood function, we need to find the estimate $\widehat{\mathbf{m}}$ that minimizes $||\mathbf{d} - f(\mathbf{m})||^2$. Similarly in nonlinear regression, we want to find the estimate $\widehat{\mathbf{m}}$ that minimizes the squared error; that is, we want to find $\widehat{\mathbf{m}}$ that minimizes $||\mathbf{d} - f(\mathbf{m})||^2$. We see that in both cases maximum likelihood and nonlinear regression, we want to find $\widehat{\mathbf{m}}$ such that $||\mathbf{d} - f(\mathbf{m})||^2$ is minimized. That is why the maximum likelihood estimate and the least-square estimate coincide. We now compute approximate 95% confidence intervals for g and C using (6.9). We obtain the following results:

$$\begin{array}{ll} g: & [8.56\ \text{m/s}^2\ ,\ 8.95\ \text{m/s}^2] \\ C: & [0.088\ \text{m}^{-1}\ ,\ 0.132\ \text{m}^{-1}] \end{array} \tag{6.12}$$

And the approximate 99% confidence intervals for g and C:

$$\begin{array}{ll} g: & [8.52\ \text{m/s}^2\ ,\ 9.01\ \text{m/s}^2] \\ C: & [0.08\ \text{m}^{-1}\ ,\ 0.14\ \text{m}^{-1}] \end{array} \tag{6.13}$$

From these results, we see that neither the approximate 95% nor 99% confidence intervals for g contain the true value of the gravity constant ($\approx 9.79\ \text{m/s}^2$). Note that these approximate confidence intervals are slightly narrower than the approximate confidence intervals obtained in Section 5.3 ((5.12) and (5.13)). We now proceed to approximate a 95% confidence region for the parameters g and C, which is represented by an ellipse using (6.10) as shown in Figure 6.1.

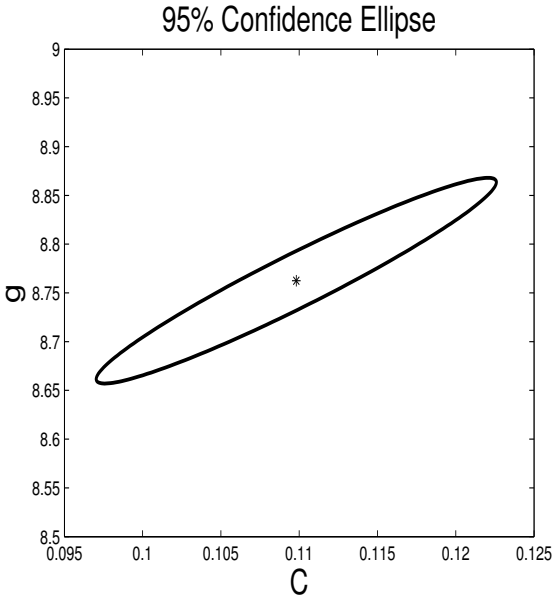

Figure 6.1: Approximate 95% confidence region for parameters g and C.

Figure 6.1 represents an approximate 95% confidence ellipse for the parameters g and C, and the least-squares estimate is represented by the star in the above plot.

6.5 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include the starting time, t_0, as another parameter in the model. Again, we carry out a weighted nonlinear regression as done in Section 6.4 but in this case we consider g, C and t_0 as the parameters in the model (2.5). We obtain the following results:

$$\widehat{g} = 8.44 \text{ m/s}^2, \qquad \widehat{C} = 0.09 \text{ m}^{-1} \qquad \widehat{t_0} = 1.392 \text{ s.} \tag{6.14}$$

Again these coincide with the ML estimates (5.14). The approximate 95% confidence intervals for g, C, and t_0 using (6.9) are:

$$\begin{array}{ll} g: & [7.92 \text{ m/s}^2 \text{ , } 8.97 \text{ m/s}^2] \\ C: & [0.05 \text{ m}^{-1} \text{ , } 0.12 \text{ m}^{-1}] \\ t_0: & [1.38 \text{ s , } 1.40 \text{ s}]. \end{array} \tag{6.15}$$

And the approximate 99% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [7.73\ \text{m/s}^2\ ,\ 9.16\ \text{m/s}^2] \\ C: & [0.04\ \text{m}^{-1}\ ,\ 0.14\ \text{m}^{-1}] \\ t_0: & [1.37\ \text{s}\ ,\ 1.41\ \text{s}]. \end{array} \tag{6.16}$$

We see that neither the approximate 95% nor 99% confidence intervals of g contain the true value of the gravity constant ($\approx 9.79\ \text{m/s}^2$). Note that these approximate confidence intervals are narrower than the approximate confidence intervals obtained in Section 5.3 ((5.15) and (5.16)). We now proceed to approximate a 95% confidence region for the parameters g, C, and t_0, which is represented by an ellipsoind using (6.10).

Figure 6.2 represents an approximate 95% confidence ellipsoid for the parameters g, C, and t_0, and the least-squares estimate is represented by the star in the below plot.

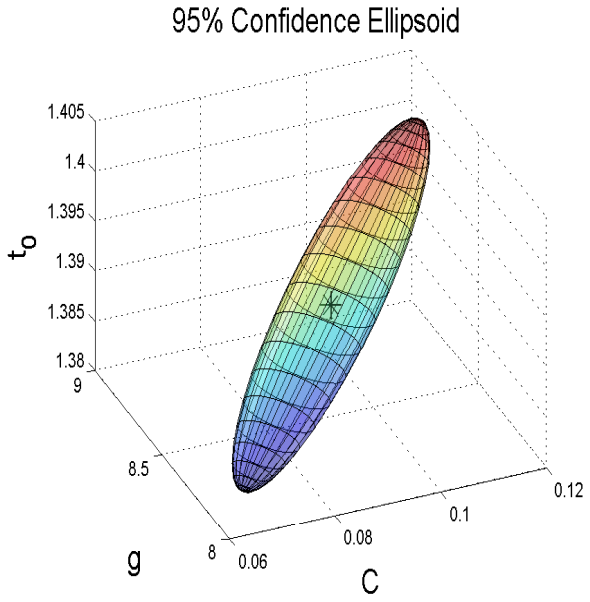

Figure 6.2: Approximate 95% confidence region for the estimated parameters.

6.6 Checking True Coverage

Since the confidence intervals computed in Sections 6.4 and 6.5 are approximations, we need to check their true coverage. We consider four cases just as in Section 5.2.2.

- First simulation: $g = 8.77$ m/s^2 and $C = 0.11$ m^{-1}.
- Second simulation: $g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.
- Third simulation: $g = 8.44$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.
- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.

We generate synthetic datasets just as the synthetic datasets that were generated in Section 5.2.2. We generate 1000 data sets for each simulation study. We then compute the approximates 95% confidence intervals for g in two ways using (6.9) and considering g and C as the parameters in the model or using (6.9) considering g, C, and t_0 as the parameters in the model. Finally, we check if these approximate 95% confidence intervals for g contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The results of the simulation studies are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	94.83 ± 0.01	0.474 ± 0.004	95.23 ± 0.01	0.483 ± 0.004
g, C, t_0	94.03 ± 0.01	1.38 ± 0.01	93.83 ± 0.01	1.41 ± 0.01

These results shows that the approximate 95% confidence intervals computed in Sections 6.4 and 6.5 have almost the target coverage. Moreover, if we compare these results to the results obtained in Section 5.2.2 and 5.3.2, we see that we obtain almost the same coverage by using frequentist techniques (ML and nonlinear regression) when we consider g and C as the parameters in the model. However, when we consider g, C, and t_0 as the parameters in the model, we obtain better coverage by using Bayesian techniques.

6.7 Bootstrap-Based Confidence Intervals

An alternative approach to obtain approximate confidence intervals without assuming a distribution are bootstrap-based confidence intervals (see, for example, [4, 6]).

6.7.1 Percentile Intervals

The quantiles of the bootstrap sampling distribution of $\hat{\theta}^*$ are used as the endpoints of the confidence interval. Let $\hat{\theta}^*_{(b)}$ represent the ordered bootstrap estimates. Then, the interval is

$$\left(\hat{\theta}^*_{(lower)}, \hat{\theta}^*_{(upper)}\right), \tag{6.17}$$

where lower$=B*\alpha/2$, upper$=B*(1-\alpha/2)$, and B is the number of bootstrap samples [6]. A better bootstrap confidence interval is Bias-Corrected intervals [6].

6.7.2 Bias-Corrected Intervals

This interval correct for the bias in the percentile interval. The interval is still based on quantiles of the bootstrap sampling distribution of $\hat{\theta}^*$ but uses quantiles different. The interval is defined as

$$\left(\hat{\theta}^*_{(lower^*)}, \hat{\theta}^*_{(upper^*)}\right), \tag{6.18}$$

where lower* $= B*\alpha_1$ and upper* $= B*\alpha_2$ and the values of α_1 and α_2 are given by

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{\alpha/2}}{1-\hat{a}(\hat{z}_0 + z_{\alpha/2})}\right), \qquad \alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{1-\alpha/2}}{1-\hat{a}(\hat{z}_0 + z_{1-\alpha/2})}\right), \tag{6.19}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function [6], z_α is the $100\alpha th$ percentile point of the standard normal distribution. $\hat{z}_0$ and $\hat{a}$ are defined as

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^*(b) < \hat{\theta}\}}{B}\right), \qquad \hat{a} = \frac{\sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^3}{6\left(\sum_{i=1}^n (\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2\right)^{3/2}}, \tag{6.20}$$

where $\Phi(\cdot)^{-1}$ indicates the inverse function of a standard normal cumulative distribution. $\#\{\hat{\theta}^*(b) < \hat{\theta}\}$ represents the number of bootstrap replications less than the original estimate. The easiest to explain $\hat{a}$ is in terms of the *jackknife* values of a statistic $\hat{\theta} = s(\mathbf{x})$. Let $\mathbf{x}_{(i)}$ be the original sample with the ith point x_i deleted, let $\hat{\theta}_{(i)} = s(\mathbf{x}_{(i)})$, and $\hat{\theta}_{(\cdot)} = \sum_{i=1}^n \hat{\theta}_{(i)}/n$. We next proceed to approximate confidence intervals for the parameters interest using the Bias-Corrected Intervals.

6.7.3 First Results

We generate bootstrap samples as follows:

1. Compute the the residuals r_i from the model (5.1).

2. Standardize the residuals. That is, $\tilde{r}_i = r_i/\sqrt{1 - H_{ii}}$, where $H = KK^{\dagger}$ (K and $K^{\dagger}$ are defined in Section 5.1).

3. Sample with replacement from $\tilde{r}_i$ to obtain a set of bootstrap residuals $r^* = (r_1^*, r_2^*, \ldots, r_{31}^*)$.

4. Define, $\epsilon = (\delta_1 r_1^*, \delta_2 r_2^*, \ldots, \delta_{31} r_{31}^*)$.

5. Obtain bootstrap data y^* by setting $y^* = f(x; \boldsymbol{\theta}) + \epsilon$, where $f(x; \boldsymbol{\theta})$ is the model (2.5) and $\boldsymbol{\theta} = (g, C)$ considering two parameters in the model or $\boldsymbol{\theta} = (g, C, t_0)$ considering three parameters in the model.

6. Use the bootstrap data to estimate $\boldsymbol{\theta}$ using weighted nonlinear least-squares.

7. Repeat steps 3-6 above B times to obtain an estimate of the bootstrap distribution.

We generate 1000 bootstrap samples as describe above to approximate 95% confidence intervals for g and C using (6.18), (6.19) and (6.20). Figure 6.3 represents the histogram of 1000 bootstrap replications of g. The approximate 95% confidence interval of g using Bias-Corrected Interval is enclosed by the vertical lines.

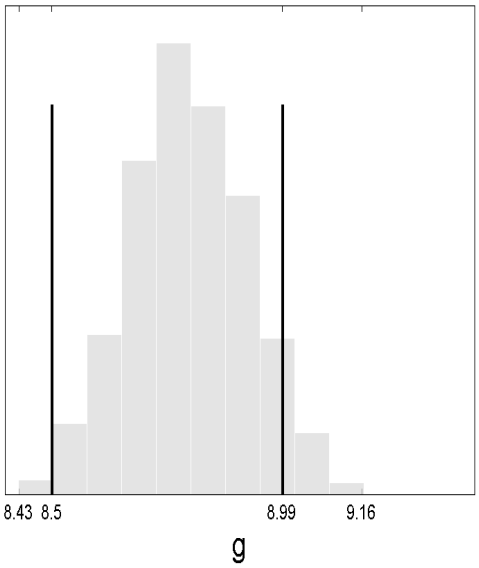

Figure 6.3: Histogram of 1000 bootstrap replications of g.

We obtain the following approximate 95% confidence intervals for g and C:

$$\begin{array}{ll} g: & [8.5 \text{ m/s}^2 , 8.99 \text{ m/s}^2] \\ C: & [0.081 \text{ m}^{-1} , 0.138 \text{ m}^{-1}] \end{array} \tag{6.21}$$

And the approximate 99% confidence intervals for g and C are:

$$\begin{array}{ll} g: & [8.43 \text{ m/s}^2 , 9.01 \text{ m/s}^2] \\ C: & [0.07 \text{ m}^{-1} , 0.148 \text{ m}^{-1}] \end{array} \tag{6.22}$$

From these results, we see that neither the approximate 95% nor the 99% confidence intervals for g contain the true value of the gravity constant. Note that these approximate confidence intervals are slightly wider than the approximate confidence intervals obtained in Section 6.4 ((6.12) and (6.13)).

6.7.4 Adding a Third Parameter: Starting Time

We now include t_0 as another parameter in the model (2.5). We generate 1000 bootstrap samples as explained in Section 6.7.3. Using (6.18), (6.19), and (6.20), we approximate confidence intervals for the parameters. We obtain the following approximate 95% confidence

intervals for g, C and t_0:

$$\begin{array}{ll} g: & [8.40\ \text{m/s}^2\ ,\ 9.88\ \text{m/s}^2] \\ C: & [0.04\ \text{m}^{-1}\ ,\ 0.14\ \text{m}^{-1}] \\ t_0: & [1.37\ \text{s}\ ,\ 1.41\ \text{s}]. \end{array} \tag{6.23}$$

And the approximate 99% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [8.17\ \text{m/s}^2\ ,\ 9.88\ \text{m/s}^2] \\ C: & [0.03\ \text{m}^{-1}\ ,\ 0.16\ \text{m}^{-1}] \\ t_0: & [1.37; \text{s}\ ,\ 1.42\ \text{s}]. \end{array} \tag{6.24}$$

From these results, we see that both the approximate 95% and 99% confidence interval contains the true value of the gravity. Note that these approximate confidence intervals are wider than the approximate intervals obtained in Section 6.5 ((6.15) and (6.16)).

6.8 Checking True Coverage of Bootstrap-Based Confidence Intervals

Since the confidence intervals computed in Sections 6.7.3 and 6.7.4 are approximations, we need to check their true coverage. We consider four cases as in Section 6.6.

- First simulation: $g = 8.77\ \text{m/s}^2$ and $C = 0.11\ \text{m}^{-1}$.
- Second simulation: $g = 9.79\ \text{m/s}^2$ and $C = 0.11\ \text{m}^{-1}$.
- Third simulation: $g = 8.44\ \text{m/s}^2$, $C = 0.09\ \text{m}^{-1}$, and $t_0 = 1.392\ \text{s}$.
- Fourth simulation: $g = 9.79\ \text{m/s}^2$, $C = 0.09\ \text{m}^{-1}$, and $t_0 = 1.392\ \text{s}$.

We generate bootstrap data as explained in Section 6.7.3. We generate 10^6 bootstrap data samples. Then, we compute the approximate 95% confidence interval for g using (6.18), (6.19) and (6.20) for every 1000 bootstrap data samples. Finally, we check if these approximate 95% confidence intervals for g contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The results of the simulation study are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	99.80 ± 0.01	0.47 ± 0.001	99.80 ± 0.01	0.48 ± 0.001
g, C, t_0	99.80 ± 0.01	1.35 ± 0.003	99.80 ± 0.01	1.38 ± 0.003

These results show that the approximate 95% confidence intervals computed in Sections 6.7.3 and 6.7.4 are conservative. If we compare these results to the results obtained in in Section 6.6, we see that the average length of the Bias-Corrected interval is slightly shorter than the average length of the approximate confidence interval obtained using (6.9) considering two and three parameters in the model.

6.9 Analysis without the δ_i's

As we did in Section 5.4, we now conduct the nonlinear regression analysis without the δ_i's.

6.9.1 First Results

We conduct a nonlinear regression analysis using the Levenberg-Maquadt algorithm. First, we only consider g and C as the parameters in the model (2.5). We obtain the following results:

$$\widehat{g} = 8.77 \text{ m/s}^2, \qquad \widehat{C} = 0.11 \text{ m}^{-1}. \tag{6.25}$$

Note that these new estimates coincide with the previous estimates (6.11) (using the uncertainty measurements). We now compute approximate 95% confidence intervals for g and C using (6.9). We obtain the following results:

$$\begin{array}{ll} g: & [8.60 \text{ m/s}^2 \text{ , } 8.94 \text{ m/s}^2] \\ C: & [0.09 \text{ m}^{-1} \text{ , } 0.131 \text{ m}^{-1}] \end{array} \tag{6.26}$$

And the approximate 99% confidence intervals for g and C:

$$\begin{array}{ll} g: & [8.53 \text{ m/s}^2 \text{ , } 9.01 \text{ m/s}^2] \\ C: & [0.08 \text{ m}^{-1} \text{ , } 0.138 \text{ m}^{-1}] \end{array} \tag{6.27}$$

We see that these new approximate confidence intervals are slightly narrower than the approximate confidence intervals obtained in Section 6.4 ((6.12) and (6.13)). We next include t_0 as another parameter in the model (2.5).

6.9.2 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include the starting time, t_0, as another parameter in the model. Again, we carry out a nonlinear regression as done in Section 6.7.1 but in this case we consider g, C and t_0 as the parameters in the model (2.5). We obtain the following results:

$$\widehat{g} = 8.44 \text{ m/s}^2, \qquad \widehat{C} = 0.09 \text{ m}^{-1} \qquad \widehat{t_0} = 1.39 \text{ s}. \tag{6.28}$$

Note that these new estimates coincide with the previous estimates (6.14) (using the δ_i's). We now compute approximate 95% confidence intervals for g, C, and t_0 using (6.9). We obtain the following results:

$$\begin{array}{ll} g: & [7.8 \text{ m/s}^2 \text{ , } 9.1 \text{ m/s}^2] \\ C: & [0.05 \text{ m}^{-1} \text{ , } 0.13 \text{ m}^{-1}] \\ t_0: & [1.38 \text{ s} \text{ , } 1.41 \text{ s}]. \end{array} \tag{6.29}$$

And the approximate 99% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [7.57 \text{ m/s}^2 \text{ , } 9.32 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} \text{ , } 0.15 \text{ m}^{-1}] \\ t_0: & [1.37 \text{ s} \text{ , } 1.41 \text{ s}]. \end{array} \tag{6.30}$$

We see that these new approximate confidence intervals are slightly wider than the approximate confidence intervals obtained in Section 6.5 ((6.15) and (6.16)).

6.9.3 Checking True Coverage

Since the confidence intervals computed in Sections 6.9.1 and 6.9.2 are approximations, we need to check their coverage. We conduct a simulation study as we did in Section 6.6. We use the Agresti-Coull interval (3.19) to compute the coverage of the true coverage of the approximate confidence intervals. The results of the simulation study are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	94.60 ± 0.01	0.433 ± 0.004	94.70 ± 0.01	0.442 ± 0.004
g, C, t_0	93.00 ± 0.01	1.611 ± 0.016	93.10 ± 0.01	1.648 ± 0.015

If we compare these results with the results obtained in Section 6.6, we see that we obtain almost the same coverage either the uncertainty measurements (δ_i for $i = 1, 2, \ldots, 31$) are included in the analysis or not. Note that the average length of the approximate 95% confidence intervals (g and C as the parameters) is narrower than the average length of the approximate 95% confidence intervals obtained in Section 6.6. However, when t_0 is include as another parameter, the average length of the approximate 95% confidence intervals is wider than the average length of the approximate 95% confidence intervals obtained in Section 6.6. We next proceed to approximate confidence intervals for the parameters of interest using bootstrap techniques.

6.10 Bootstrap-Based Confidence Intervals without δ_i's

As we did in Sections 6.7.3 and 6.7.4, we now approximate confidence intervals for the parameters using bootstrapping. However, we now exclude the δ_i's from the analysis.

6.10.1 First Results

We now generate bootstrap samples slightly different from the bootstrap samples generated in Sections 6.7.3 and 6.7.4.

1. Compute the the residuals r_i from the model (5.21).

2. Standardize the residuals. That is, $\tilde{r}_i = r_i/\sqrt{1 - H_{ii}}$, where $H = KK^{\dagger}$ as defined in Section 5.1.

3. Sample with replacement from $\tilde{r}_i$ to obtain a set of bootstrap residuals $r^* = (r_1^*, r_2^*, \ldots, r_{31}^*)$.

4. Define $\epsilon_i = (r_1^*, r_2^*, \ldots, r_{31}^*)$.

5. Obtain bootstrap data y^* by setting $y^* = f(x;\boldsymbol{\theta}) + \epsilon_i$, where $f(x;\boldsymbol{\theta})$ is the model (2.5) and $\boldsymbol{\theta} = (g, C)$ considering two parameters in the model or $\boldsymbol{\theta} = (g, C, t_0)$ considering three parameters in the model.

6. Use the bootstrap data to estimate $\boldsymbol{\theta}$ using nonlinear least-squares.

7. Repeat steps 3-6 above B times to obtain an estimate of the bootstrap distribution.

We generate 1000 bootstrap samples as describe above to approximate 95% confidence intervals for g and C using (6.18), (6.19) and (6.20). Figure 6.4 represents the histogram of 1000 bootstrap replications of g. The approximate 95% confidence interval of g using Bias-Corrected Interval is enclosed by the vertical lines.

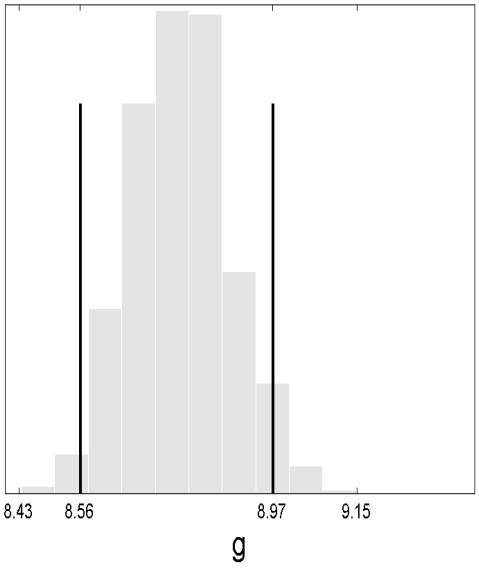

Figure 6.4: Histogram of 1000 bootstrap replications without δ_i's of g.

We obtain the following approximate 95% confidence intervals for g and C:

$$\begin{array}{ll} g: & [8.56\ \mathrm{m/s^2}\ ,\ 8.97\ \mathrm{m/s^2}] \\ C: & [0.09\ \mathrm{m^{-1}}\ ,\ 0.134\ \mathrm{m^{-1}}] \end{array} \tag{6.31}$$

And the approximate 99% confidence intervals for g and C are:

$$\begin{array}{ll} g: & [8.45\ \mathrm{m/s^2}\ ,\ 9.02\ \mathrm{m/s^2}] \\ C: & [0.08\ \mathrm{m^{-1}}\ ,\ 0.14\ \mathrm{m^{-1}}] \end{array} \tag{6.32}$$

From these results, we see that neither the approximate 95% nor the 99% confidence intervals contain the true value of g. Note that these approximate confidence intervals are wider than the approximate intervals obtained in Section 6.9.1 ((6.26) and (6.27)). We next include t_0 as another parameter in the model.

6.10.2 Adding a Third Parameter: Starting Time

We now include t_0 as another parameter in the model (2.5). We generate 1000 bootstrap samples as explained in Section 6.10.1 Using (6.18), (6.19), and (6.20), we approximate confidence intervals for the parameters. We obtain the following approximate 95% confidence intervals for g, C and t_0:

$$\begin{array}{ll} g: & [7.7 \text{ m/s}^2 , 9.26 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} , 0.14 \text{ m}^{-1}] \\ t_0: & [1.37 \text{ s} , 1.41 \text{ s}]. \end{array} \tag{6.33}$$

And the approximate 99% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [7.58 \text{ m/s}^2 , 9.46 \text{ m/s}^2] \\ C: & [0.02 \text{ m}^{-1} , 0.15 \text{ m}^{-1}] \\ t_0: & [1.36 \text{ s} , 1.42 \text{ s}]. \end{array} \tag{6.34}$$

From these results, we see that neither the approximate 95% nor the 99% confidence intervals contain the true value of g. Note that these approximate confidence intervals are wider than the approximate intervals obtained in Section 6.9.2 ((6.29) and (6.30)).

6.11 Checking True Coverage of Bootstrap-Based Confidence Intervals without δ_i's

Since the confidence intervals computed in Sections Section 6.10.1 and 6.10.2 are approximations, we need to check their true coverage. We consider four cases as in Section 6.8.

- First simulation: $g = 8.77$ m/s^2 and $C = 0.11$ m^{-1}.
- Second simulation: $g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.
- Third simulation: $g = 8.44$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.

- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.392$ s.

We generate bootstrap data as explained in Section 6.10.1 We generate 10^6 bootstrap data samples. Then, we compute the approximate 95% confidence interval for g using (6.18), (6.19) and (6.20) for every 1000 bootstrap data samples. Finally, we check if these approximate 95% confidence intervals for g contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the true coverage of the approximate confidence intervals. The results of the simulation study are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	99.80 ± 0.01	0.414 ± 0.001	99.80 ± 0.01	0.423 ± 0.001
g, C, t_0	99.80 ± 0.01	1.55 ± 0.003	99.80 ± 0.01	1.58 ± 0.003

These results show that the approximate 95% confidence intervals computed in Sections 6.10.1 and 6.10.2 are conservative. If we compare these results to the results obtained in in Section 6.9.3, we see that the average length of the Bias-Corrected interval is slightly shorter than the average length of the approximate confidence interval obtained using (6.9).

CHAPTER 7

CONSTRUCTING A CONFIDENCE REGION BY INVERSION

In Section 5.1, we showed that the assumption

$$\frac{\epsilon_i}{\delta_i} \sim N(0, 0.2^2)$$

is more reasonable than that in ALA12. An alternative approach to approximate confidence intervals for the parameters of interest is to construct an approximate confidence region and then map back the set of values for the parameters that generate the approximate confidence region (e.g., [16]). Let

$$\varepsilon_i = \frac{\epsilon_i}{\sigma\delta_i} \qquad \text{for } i = 1, 2, \ldots, 31$$

where $\sigma = 0.2$. So that $\underline{\varepsilon} \sim N(0, I)$. Then,

$$||\underline{\varepsilon}||^2 = \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_{31}^2 = \sum_{i=1}^{31} \left(\frac{d_i - f_i(\mathbf{m})}{\sigma\delta_i} \right)^2 \sim \chi^2_{31}, \tag{7.1}$$

where d_i is the ith measurement, δ_i is the uncertainty in the ith measurement and $f_i(\mathbf{m})$ is the physical model (2.5) evaluated at time i. We define R_α as:

$$R_\alpha = \left\{ \mathbf{m} : \sum_{i=1}^{31} \left(\frac{d_i - f_i(\mathbf{m})}{\sigma\delta_i} \right)^2 \leq \chi^2_{31}(\alpha) \right\}, \tag{7.2}$$

where $\chi^2_{31}(\alpha)$ is the upper α percentile of the χ^2 distribution with 31 degrees of freedom. Then, we have

$$\mathbb{P}\left[\mathbf{m} \in R_\alpha\right] = 1 - \alpha, \tag{7.3}$$

which implies that R_α is an approximate $100(1-\alpha)\%$ confidence region for $\mathbf{m}$.

7.1 First Results

First, we let $\mathbf{m}$=(g, C) be the parameter in the model (2.5). We find an approximate 95% confidence region for $\mathbf{m}$ by finding the intersection of the surface $S(\mathbf{m})$, where $S(\mathbf{m})$ is given by:

$$S(\mathbf{m}) = \sum_{i=1}^{31} \left(\frac{d_i - f_i(\mathbf{m})}{\sigma \delta_i} \right)^2 \tag{7.4}$$

evaluated over the same grid of points that was used in Section 3.4., and the upper 95 percentile of the χ^2 distribution with 31 degrees of freedom as shown in Figure 7.1.

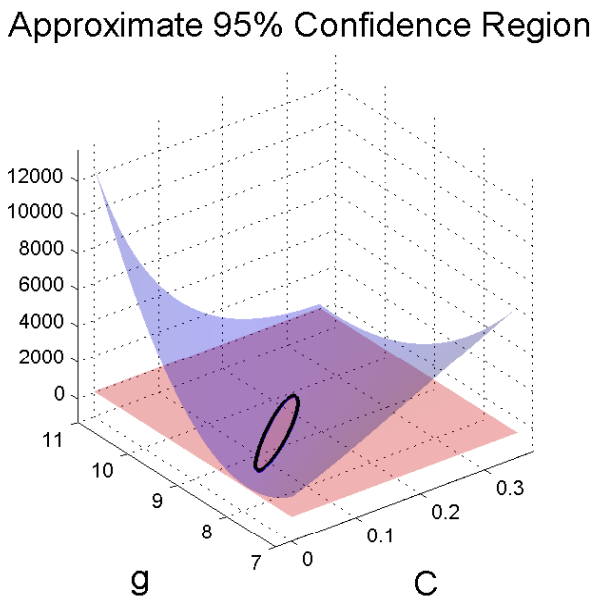

Figure 7.1: Approximate 95% confidence region.

In Figure 7.1, the light blue surface represents (7.4) evaluated over the same grid of points that was used in Section 3.4. The pink plane represents the upper 95 percentile of the χ^2 distribution with 31 degrees of freedom. The region enclosed by the black line represents an approximate 95% confidence region for $\mathbf{m}$. The projection of the approximate 95% confidence region of $\mathbf{m}$ onto the g and C plane is shown in Figure 7.2.

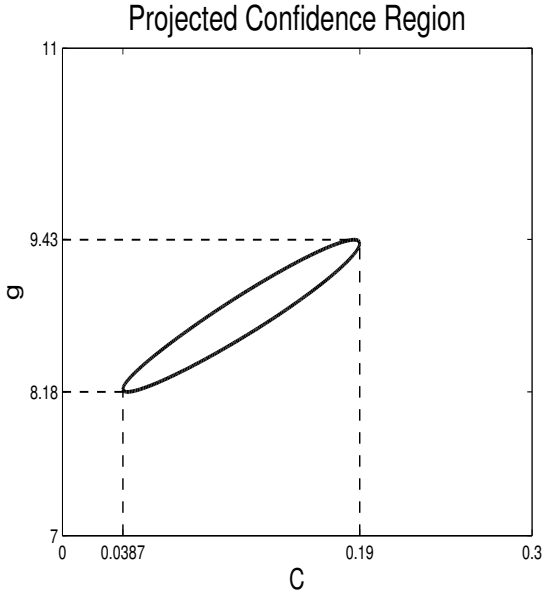

Figure 7.2: Projected 95% confidence region.

By projecting the approximate region of $\mathbf{m}$ onto the g and C axes, we obtain approximate 95% confidence intervals:

$$\begin{array}{ll} g: & [8.18 \text{ m/s}^2 , 9.43 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} , 0.19 \text{ m}^{-1}]. \end{array} \tag{7.5}$$

And the approximate 99% confidence intervals for g and C:

$$\begin{array}{ll} g: & [8.11 \text{ m/s}^2 , 9.51 \text{ m/s}^2] \\ C: & [0.03 \text{ m}^{-1} , 0.20 \text{ m}^{-1}]. \end{array} \tag{7.6}$$

From these results, we see that these approximate confidence intervals are wider than the approximate confidence intervals obtained in Sections 5.3 and 6.4.

7.2 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include the starting time, t_0, as another parameter in the model. Now, we let $\mathbf{m}=(g, C, t_0)$ be the parameter in the model (2.5). We find an approximate 95% confidence region for $\mathbf{m}$ by finding the intersection of the surface region $S(\mathbf{m})$ (7.4) evaluated over the same grid of points that was used in Section 3.5, and the upper 95 percentile of the χ^2 distribution with 31 degrees of freedom. Then, we project the

approximate 95% confidence region of $\mathbf{m}$ onto the g, C, and t_0 axes. The approximate 95% confidence intervals for g, C and t_0 are:

$$\begin{array}{ll} g: & [6.98 \text{ m/s}^2 , 10.66 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} , 0.24 \text{ m}^{-1}] \\ t_0: & [1.35 \text{ s} , 1.43 \text{ s}] \end{array} \tag{7.7}$$

And the approximate 99% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [6.93 \text{ m/s}^2 , 10.99 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} , 0.26 \text{ m}^{-1}] \\ t_0: & [1.35 \text{ s} , 1.44 \text{ s}] \end{array} \tag{7.8}$$

From these results, we see that these approximate confidence intervals are wider than the approximate confidence intervals obtained in Sections 5.3.1 and 6.5. Note that both the approximate 95% and 99% confidence intervals contain the true value of the gravity constant ($\approx 9.79 \text{ m/s}^2$).

7.3 Checking True Coverage

Since the confidence intervals computed in Sections 7.1 and 7.2 are approximations, we need to check their true coverage. We consider four cases as in Section 5.2.2.

- First simulation: $g = 8.77 \text{ m/s}^2$ and $C = 0.11 \text{ m}^{-1}$.
- Second simulation: $g = 9.79 \text{ m/s}^2$ and $C = 0.11 \text{ m}^{-1}$.
- Third simulation: $g = 8.44 \text{ m/s}^2$, $C = 0.09 \text{ m}^{-1}$, and $t_0 = 1.392 \text{ s}$.
- Fourth simulation: $g = 9.79 \text{ m/s}^2$, $C = 0.09 \text{ m}^{-1}$, and $t_0 = 1.392 \text{ s}$.

We use the same synthetic data sets used in Section 5.3.2. We then compute the approximate 95% confidence interval for g as described in Sections 7.1 (considering g and C as the only parameters in the model) and 7.2 (including t_0 as another parameter in the model). Finally, we check if these intervals contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the coverage of the true coverage of the approximate confidence intervals. The results of the simulation study are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	96.13 ± 0.012	0.942 ± 0.015	96.03 ± 0.01	0.942 ± 0.014
g, C, t_0	99.80 ± 0.01	3.354 ± 0.007	99.80 ± 0.01	4.01 ± 0.001

From these results, we see that the approximate 95% confidence intervals computed in Sections 7.1 and 7.2 are conservative. If we compare these results to the results obtained in Section 5.3.2, we see that these approximate 95% confidence intervals for g are wider than the approximate 95% confidence intervals for g obtained in Section 5.3.2.

7.4 Analysis without δ_i's

As we did in Section 5.4, we conduct the same analysis as in Sections 7.1 and 7.2. However, we now exclude the δ_i's.

7.4.1 First Results

First, we let $\mathbf{m}$=(g, C) be the parameter in the model (2.5). We find an approximate 95% confidence region for $\mathbf{m}$ by finding the intersection of the surface $S(\mathbf{m})$, where $S(\mathbf{m})$ is given by:

$$S(\mathbf{m}) = \sum_{i=1}^{31} \left(\frac{d_i - f_i(\mathbf{m})}{\sigma} \right)^2 \tag{7.9}$$

evaluated over the same grid of points that was used in Section 3.4., and the upper 95 percentile of the χ^2 distribution with 31 degrees of freedom. Note that in this case $\sigma = 1.25$ as explained in Section 5.4. By projecting the approximate region onto the g and C axes, we obtain approximate 95% confidence intervals:

$$\begin{array}{ll} g: & [8.24 \text{ m/s}^2 , 9.37 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} , 0.182 \text{ m}^{-1}]. \end{array} \tag{7.10}$$

And the approximate 99% confidence intervals for g and C:

$$\begin{array}{ll} g: & [8.18 \text{ m/s}^2 , 9.45 \text{ m/s}^2] \\ C: & [0.04 \text{ m}^{-1} , 0.19 \text{ m}^{-1}]. \end{array} \tag{7.11}$$

From these results, we see that these approximate confidence intervals are narrower than the approximate confidence intervals obtained in Section 7.1. We next include t_0 as another parameter in the model (2.5).

7.4.2 Adding a Third Parameter: Starting Time

As explained in Section 3.5, we now include the starting time, t_0, as another parameter in the model. We find an approximate 95% confidence region for $\mathbf{m}$ by finding the intersection of the surface $S(\mathbf{m})$ (7.9) evaluated over the same grid of points that was used in Section 3.5., and the upper 95 percentile of the χ^2 distribution with 31 degrees of freedom. Note that in this case $\sigma = 1.25$ as explained in Section 5.4. Then, we project the approximate 95% confidence region of $\mathbf{m}$ onto the g, C, and t_0 axes. The approximate 95% confidence intervals for g, C and t_0 are:

$$\begin{array}{ll} g: & [6.94 \text{ m/s}^2 , 11.03 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} , 0.25 \text{ m}^{-1}] \\ t_0: & [1.35 \text{ s} , 1.44 \text{ s}] \end{array} \tag{7.12}$$

And the approximate 99% confidence intervals for g, C, and t_0:

$$\begin{array}{ll} g: & [6.91 \text{ m/s}^2 , 11.43 \text{ m/s}^2] \\ C: & [0 \text{ m}^{-1} , 0.27 \text{ m}^{-1}] \\ t_0: & [1.35 \text{ s} , 1.45 \text{ s}] \end{array} \tag{7.13}$$

From these results, we see that these approximate confidence intervals are wider than the approximate confidence intervals obtained in Section 7.2 ((7.6) and (7.7)). Note that both the approximate 95% and 99% confidence intervals contain the true value of the gravity constant (≈ 9.79 m/s^2).

7.5 Checking True Coverage

Since the confidence intervals computed in Sections 7.4.1 and 7.4.2 are approximations, we need to check their true coverage. We consider four cases as in Section 7.3.

- First simulation: $g = 8.77$ m/s^2 and $C = 0.11$ m^{-1}.
- Second simulation: $g = 9.79$ m/s^2 and $C = 0.11$ m^{-1}.

- Third simulation: $g = 8.41$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.391$ s.

- Fourth simulation: $g = 9.79$ m/s^2, $C = 0.09$ m^{-1}, and $t_0 = 1.391$ s.

We use the same synthetic data sets used in Section 5.4.6. We then compute the approximate 95% confidence interval for g as described in Sections 7.4.1 (considering g and C as the only parameters in the model) and 7.4.2 (including t_0 as another parameter in the model). Finally, we check if these intervals contain their corresponding true value. We use the Agresti-Coull interval (3.19) to compute the coverage of the true coverage of the approximate confidence intervals. The results of the simulation study are:

	$\widehat{g}$		$g = 9.79$ m/s^2	
	coverage	average length	coverage	average length
g, C	96.13 ± 0.012	0.869 ± 0.012	96.22 ± 0.012	0.892 ± 0.015
g, C, t_0	97.20 ± 0.01	3.184 ± 0.053	97.02 ± 0.01	3.14 ± 0.048

From these results, we see that the approximate 95% confidence intervals computed in Sections 7.4.1 and 7.4.2 are conservative. If we compare these results to the results obtained in Section 5.4.6, we see that these approximate 95% confidence intervals for g are wider than the approximate 95% confidence intervals for g obtained in Section 5.4.6. Note that the average length of the approximate 95% confidence intervals for g (excluding δ_i) are narrower than the approximate 95% confidence intervals for g obtained in Section 7.3.

CHAPTER 8
CONCLUSIONS AND FURTHER RESEARCH

We have presented a comparison between Bayesian and frequentist techniques for parameter estimation. After a careful validation of the assumptions made by ALA12, we proposed a new likelihood function that gives us very similar estimates for the parameters of interest either using Bayesian or non-Bayesian techniques. Then, we focused on the comparison of the frequentist coverage of the shortest 95% Bayesian credible interval of g with the coverage of the 95% confidence interval of g using maximum likelihood. In both cases, we obtained coverages close to the target coverage.

We also presented other approaches to approximate confidence intervals for the parameters of interest such as bootstrap-based confidence intervals, nonlinear regression confidence intervals, and confidence intervals obtained by projecting a confidence region. We studied the coverage of each confidence interval and compared them with each other. The purpose of doing all this tedious process is to show how these confidence intervals perform on real life experiments.

All the methods we used seem to underestimate true value of the gravity constant. This may reflect a problem with the experiment. Now, the question is how we could obtain an estimate of g closer to the true value. Optimal experimental design will tell us how we should measure things to obtain the best estimates of the parameters of interest given the available resources and experimental constraints.

This thesis by no means tries to show that Bayesian methods are better than frequentist methods or vice versa. What is more important is to be able to validate whatever approach

we choose. One approach may be better than the other, depending on the problem. However, both schools of inference have something to offer. We prefer to keep an open mind. At the end of the day, we are statisticians and should be able to use all available tools to solve a problem.

REFERENCES CITED

[1] M. Allmaras, W. Bangerth, J. Linhart, J. Polanco, F. Wang, K. Wang, J. Webster & S. Zedler. *Estimating Parameters in Physical Models Through Bayesian Inversion: A Complete Example*. SIAM Journal of Applied Mathematics, 2012.

[2] L. D. Brown, T. T. Cai & A. DasGupta. *Interval estimation for a binomial proportion*. Statistical Science, pp. 101-117, 2001.

[3] G. Casella & R. L. Berger. *Statistical Inference (2nd Edition)*. Duxbury, 2002.

[4] A.C. Davison & D.V. Hinkley. *Bootstrap Methods and their Applications*. Cambridge University Press, 1997.

[5] B. Efron & D.V. Hinkley. *Assesing the Accuracy of the Maximum Likelihood Estimator: Observed Versus Expected Fisher Information*. Biometrika, Vol. 65, No. 3, pp. 457-482, 1978.

[6] B. Efron & R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, 1993.

[7] R. A. Fisher. *Theory of statistical estimation*. Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 22, No. 05. Cambridge University Press, 1925.

[8] R. A. Fisher. *Two new properties of mathematical likelihood*. Proceedings of the Royal Society of London. Series A 144, No. 852, pp. 285-307, 1934.

[9] A. R. Gallant. *Nonlinear Statistical Models*. Wiley, 1987.

[10] J. S. Galpin & D. M. Hawkins. *The Use of Recursive Residuals in Checking Model Fit in Linear Regression*. The American Statistician, Vol. 38, No. 2, pp. 94-105, 1984.

[11] Y. Pawitan. *In All Likelihood: Statistical Modeling and Inference Using Likelihood*. Oxford University Press, 2001.

[12] J. O. Rawlings, S. G. Pantula & D. A. Dickey. *Applied Regression Analysis A Research Tool (2nd Edition)*. Springer, 1998.

[13] C. P. Robert & G. Casella. *Monte Carlo Statistical Methods (2nd Edition)*. Springer, 2004.

[14] G. A. F. Seber & C. J. Wild. *Nonlinear Regression*. Wiley, 2003.

[15] J.A. Scales & L. Tenorio. *Prior information and uncertainty in inverse problems*. Geophysics 66.2 pp. 389-397, 2001.

[16] P. B. Stark (1992). *Inference in Infinite-Dimensional Inverse Problems: Discretization and Duality*. Journal of Geophysical Reseach, Vol. 97, No. B10, 14055-14082, 1992.

[17] P. B. Stark & L. Tenorio. *A primer of frequentist and Bayesian inference in inverse problems*. In Computational Methods for Large-Scale Inverse Problems and Quantification of Uncertainty, pp. 9-32, Wiley, 2011.

[18] A. C. Tamhane & D. D. Dunlop. *Statistics and Data Analysis from Elementary to Intermediate*. Prentice-Hall, 2000.

[19] C. R. Vogel. *Computational Methods for Inverse Problems*. SIAM, 2002.