

The Method of Finite Difference Regression

Arjun Banerjee

Computer Science Department, Purdue University, West Lafayette, IN, USA

Email: banerjearjun73@gmail.com

How to cite this paper: Banerjee, A. (2018)

The Method of Finite Difference Regression. *Open Journal of Statistics*, 8, 49-68.

<https://doi.org/10.4236/ojs.2018.81005>

Received: December 23, 2017

Accepted: January 30, 2018

Published: February 2, 2018

Copyright © 2018 by author and
Scientific Research Publishing Inc.

This work is licensed under the Creative
Commons Attribution International
License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper¹ I present a novel polynomial regression method called Finite Difference Regression for a uniformly sampled sequence of noisy data points that determines the order of the best fitting polynomial and provides estimates of its coefficients. Unlike classical least-squares polynomial regression methods in the case where the order of the best fitting polynomial is unknown and must be determined from the R^2 value of the fit, I show how the t -test from statistics can be combined with the method of finite differences to yield a more sensitive and objective measure of the order of the best fitting polynomial. Furthermore, it is shown how these finite differences used in the determination of the order, can be reemployed to produce excellent estimates of the coefficients of the best fitting polynomial. I show that not only are these coefficients unbiased and consistent, but also that the asymptotic properties of the fit get better with increasing degrees of the fitting polynomial.

Keywords

Polynomial Regression, t -Test, Finite Differences

1. Introduction

In this paper I present a novel polynomial regression method for a uniformly sampled sequence of noisy data points that I call the method of Finite Difference Regression. I show how the statistical t -test² can be combined with the method of finite differences to provide a more sensitive and objective measure of the order of the best fitting polynomial and produce unbiased and consistent estimates of its coefficients.

¹Portions of this independent research were submitted to the 2013 Siemens Math, Science, and Technology Competition, and to the 2014 Intel Science Talent Search Competition for scholarship purposes. It received the Scientific Research Report Badge from the latter competition.

²The t -test often called Student's t -test was introduced in statistics by William S. Gosset in 1908 and published under the author's pseudonym "Student" (see [10]).

Regression methods, also known as fitting models to data, have had a long history. Starting from Gauss (see [1] pp. 249-273) who used it to fit astronomical data, it is one of the most useful techniques in the arsenal of science and technology. While historically it has been used in the “hard” sciences, both natural and biological, and in engineering, it is finding increasing use in all human endeavors, even in the “soft” sciences like social sciences (see for example [2]), where data is analyzed in order to find an underlying model that may explain or describe it.

Polynomial regression methods are methods that determine the best polynomial that describes a sequence of noisy data samples. In classical polynomial regression methods *when the order of the underlying model is known*, the coefficients of the terms are estimated by minimizing the sum of squares of the residual errors to the fit. For example, in classical *linear* regression (see [2] pp. 675-730) one assumes that the order of the fitting polynomial is 1 and then proceeds to find the coefficients of the best fitting line as those that minimize the sum of the squares of its deviations from the noisy samples. The same method can be extended to polynomial regression. On the other hand, *when the model order is unknown*, the order of the best fitting polynomial must be determined *before* finding the coefficients of the polynomial terms. In classical regression methods this order is determined from the R^2 value of its fit. More on these two different aspects of polynomial regression are provided below.

1) *Determining the order of the best fitting polynomial (Model order unknown)*

In classical regression, the goodness of the order of the fitting polynomial is given by the R^2 value of the best fitting polynomial—the greater the R^2 value the better the fit. However, this means that one can go on fitting with higher degree polynomials and thereby increasing R^2 values. Given the insensitivity of these R^2 values (see for example Section 2.4) it is indeed hard to find a stopping criterion. To get around this problem, heuristic methods like AIC, BIC, etc. (see [3] for example) that penalize too many fitting parameters have been proposed. In contrast, in the method of Finite Difference Regression described below, the order of the best fitting polynomial is determined by using a t -test, namely determining the index of a finite difference row (or column) that has a high probability (beyond a certain significance level) of being zero. The sensitivity of this method is high compared to classical regression methods as I shall demonstrate in Section 2.5.

2) *Determining the coefficients of the best fitting polynomial (Model order known/determined)*

In classical polynomial regression, the polynomial coefficients of the fit that minimizes the sum of squares of the residual errors are found by a computation that is equivalent to inverting a square matrix. In contrast in the method I describe below, I once again employ the finite differences I used to find the order, to iteratively determine the polynomial coefficients. In Section 2.7, I show that the results of this method are remarkable—not only are the coefficients unbiased

and consistent, but also that the asymptotic properties of the fit get better with increasing degrees of the fitting polynomial.

Before describing the Finite Difference Regression method in Section 2, I briefly recapitulate the general method of finite differences whose formulation can be traced all the way back to Sir Isaac Newton (see [4]). The finite differences method (see [5]) is used for curve fitting with polynomial models. On its 2nd row, **Table 1** shows the y -value outputs for the sequence of regularly spaced x -value inputs at points 1, 2, ..., 10 shown on the 1st row. A plot of this data is shown in **Figure 1**. The problem is to determine the order of the polynomial that can generate these outputs and to find its coefficients.

The remaining rows in **Table 1** show the successive differences of the y -values. The fact that the 3rd difference row in **Table 1** is composed of all zeros implies that a quadratic polynomial best describes this data. In fact this data was generated by the quadratic polynomial: $y = 2x^2 + x + 3$.

I assume throughout this paper that the data are sampled at regular intervals. Without loss of generality these intervals can be assumed to be of unit length. Then the following general observations can be made:

1) If the $(m+1)^{\text{st}}$ difference row is all zero, then the data is best described by an m^{th} order polynomial: $a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$

2) If the $(m+1)^{\text{st}}$ difference row is all zero, then $m!a_m$ is the constant in the m^{th} row. From this observation the coefficient of the highest term can be found. For example in **Table 1** since the 3rd difference row is all zero then $2!a_2$ is the constant in the 2nd difference row. Since this constant is 4, it immediately follows that the highest term should be $4x^2/2! = 2x^2$

3) The lower order terms can be found successively by back substituting and subtracting the highest term. In this method, the data y in **Table 1** is replaced by y' where $y' = y - a_m x^m$, and then reapplying the method of finite differences on the y' data. Using this approach for the data in **Table 1**, one obtains the next term as follows:

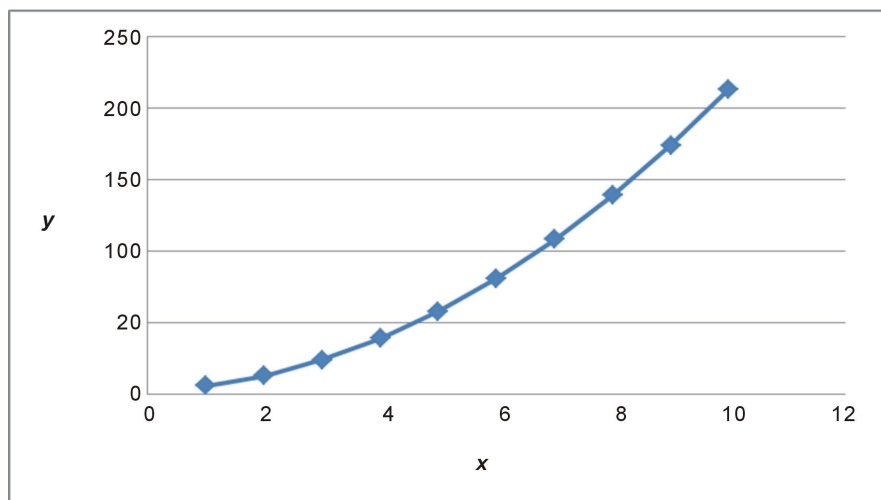


Figure 1. Graph of data in **Table 1**.

In **Table 2** the 2nd difference row is all zero. Hence $1!a_1$ is the constant in the 1st difference row. Since this constant is 1, the next highest term should be $1x/1! = x$.

Again, the final term can be found by replacing y' in **Table 2** by y'' where $y'' = y' - a_{m-1}x^{m-1}$:

The computation in **Table 3** yields 3 as the constant term for the polynomial fit.

Hence, by successive applications of the finite difference method in this manner, the best fit polynomial to the y data in **Table 1** is found to be $2x^2 + x + 3$.

2. The Method of Finite Difference Regression

In this section I describe the Finite Difference Regression method, compare the results obtained from this method to those from classical regression, and show unbiasedness and consistency properties of the Finite Difference Regression coefficients.

The intuitive idea behind the method of Finite Difference Regression is simple. Suppose one were to add normally distributed noise with mean 0 independently to each of the y -values for example in **Table 1** to make them noisy, then unlike the non-noisy situation, one cannot expect all entries in a difference row to be all

Table 1. Finite Differences on y -value outputs for the sequence of regularly spaced x -value inputs.

x	1	2	3	4	5	6	7	8	9	10
y	6	13	24	39	58	81	108	139	174	213
1st Difference		7	11	15	19	23	27	31	35	39
2nd Difference			4	4	4	4	4	4	4	
3rd Difference				0	0	0	0	0	0	
4th Difference					0	0	0	0		

Table 2. Second iteration of finding terms to fit the model.

x	1	2	3	4	5	6	7	8	9	10
y	6	13	24	39	58	81	108	139	174	213
$y' = y - 2x^2$	4	5	6	7	8	9	10	11	12	13
1st Difference		1	1	1	1	1	1	1	1	1
2nd Difference			0	0	0	0	0	0	0	

Table 3. Final iteration of finding terms to fit the model.

x	1	2	3	4	5	6	7	8	9	10
y'	4	5	6	7	8	9	10	11	12	13
$y'' = y' - x$	3	3	3	3	3	3	3	3	3	3
1st Difference		0	0	0	0	0	0	0	0	0

identically equal to zero. However, since differences between normally distributed random variables are also normally distributed (albeit with higher variance), each of the samples in a difference row (as in the rows of **Table 1**) would also be normally distributed with *different means* until one hits a difference row where the mean is zero with high probability. It is this highly probable zero-mean row that should be the all zero difference row in the absence or elimination of noise. The condition that the mean is zero with high probability can be objectively tested on each difference row using a t -test on the samples in that row, where the null hypothesis H_0 tests for zero population mean and the alternative hypothesis H_a tests for non-zero population mean. This is the main idea behind the method of Finite Difference Regression—the replacement of the notion of an all zero difference row in the non-noisy case by the notion of a *highly probable* row of all zeros in the noisy case and usage of the statistical t -test to test for the zero population mean hypothesis.

2.1. Statistical Independence and Its Relation to the Degrees of Freedom

While it is true that successive differences will produce normally distributed samples, not all the samples so produced are statistically independent of one another. For example as shown in **Figure 2**, the noisy data samples y_1, y_2, y_3, y_4 generate the next difference sequence y'_1, y'_2, y'_3 —not all of these values are independent. Since y'_1 and y'_2 both depend upon the value of y_2 they are not independent. Similarly, y'_2 and y'_3 are not independent because both depend upon y_3 . However, y'_1 and y'_3 having no point in common, are independent!

This observation can be generalized by the following process. Suppose I had K observations to begin with. If I were to produce successive differences by skipping over every other term, in the manner of a binary tree as shown in **Figure 3**, I would obtain independent normally distributed samples. Each level of the tree would have independent normally distributed finite differences denoted in **Figure 3** (shown with $K = 8$) by y', y'', \dots and so on. This method would retain all the properties of finite differences, in particular the highly probable zero mean row; plus it would have the added benefit of not having any dependent samples. If this method were implemented, and a t -test invoked at every level of the binary tree, then the number of samples in the test would be halved at each level, and consequently the degrees of freedom would be affected. In other words, if k denotes the number of samples³ at a particular level of the tree, then in the next level one would get $k/2$ samples. Then according to the theory of t -tests, the degrees of freedom for the first level would be $k - 1$, in the next level it would be $k/2 - 1$, and so on. However, discarding good data will have negative consequences for coefficient estimation (see Section 2.6). While some of the data at every level of a successive finite difference table (**Table 1** for example) are not independent, they could still be used for t -tests. *The only thing to remember is*

³Assume that K and k are powers of 2 for the purposes of this discussion.

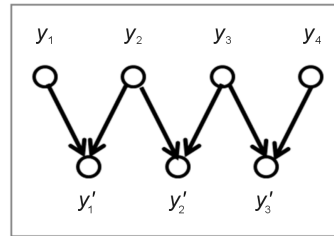


Figure 2. Successive finite differences.

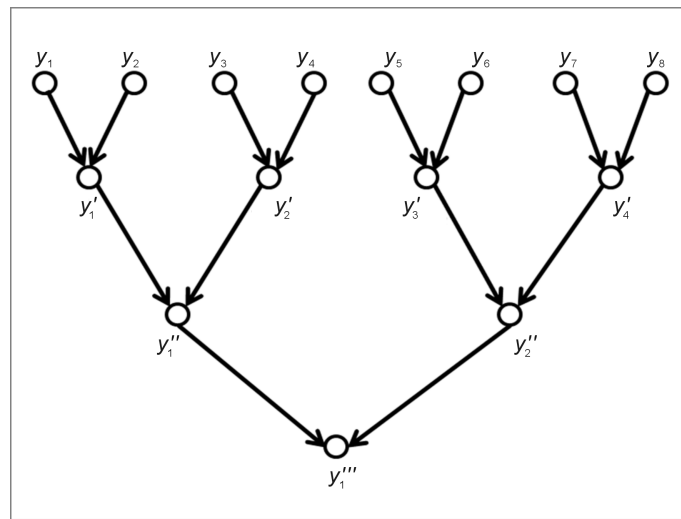


Figure 3. A binary tree of successive differences.

that at every stage the number of independent samples is halved and therefore the degrees of freedom is one less than this number.

2.2. An Example of Order Finding

Table A1 in **Appendix 2** gives an example of how the Finite Difference Regression method works to find the order of the best fitting polynomial. The 1st two columns of **Table A1** are the *pivoted rows* of **Table 1**. Independent zero-mean normally distributed random noise with standard deviation of **30.0** is generated in the 3rd column and added to the values of y in the 4th column. Thus, the 4th column represents noisy values of the data y . A graph of this noisy data is shown in **Figure 4**. Finally, columns 5 through 9 represent successive finite differences from 1st to 5th, computed in the same way as in **Table 1**. The last 6 rows of **Table 9** highlighted in yellow show the statistical computations associated with the method. These are described below in Section 2.3.

2.3. The Statistical Computations

The last 6 rows of **Table A1** use the one-sample t -test for a population mean as described in [2] pp. 547-548. Note that the **Null Hypothesis** is $H_0 : \mu = 0$ because the *hypothesized value* is 0. The **Alternative Hypothesis** is $H_a : \mu \neq 0$. The descriptions of these rows are:

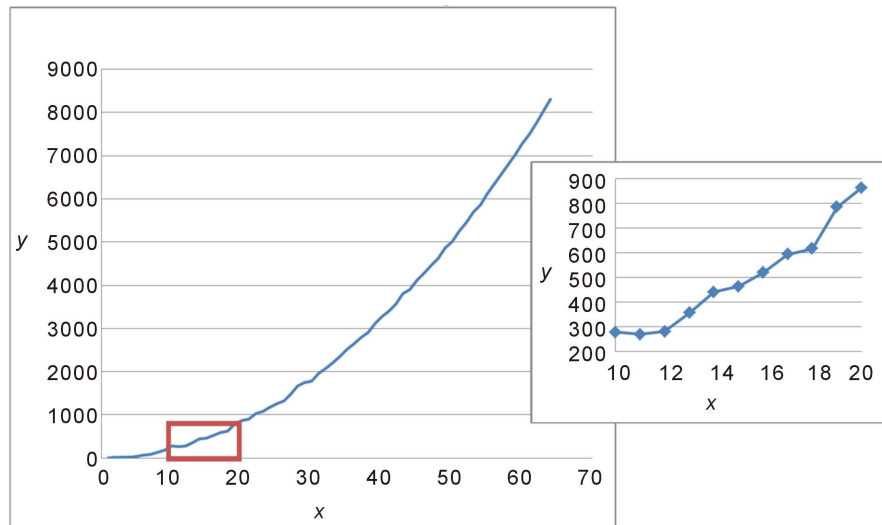


Figure 4. Graph of the noisy data in **Table 5** with insert showing magnified view for $10 \leq x \leq 20$.

Header titled n = Number of samples

Header titled \bar{x} = Sample mean

Header titled s = Sample standard deviation

Header titled t = Test statistic = $\frac{x - \mu}{s/\sqrt{n}} = \frac{x}{s/\sqrt{n}}$ (substituting $\mu = 0$)

Header titled df = Degrees of freedom computed as detailed in Section 2.1

Header titled P -value = $\begin{cases} 2 \times \text{Area to right of } t \text{ with given } df & \text{for } t > 0 \\ 2 \times \text{Area to left of } t \text{ with given } df & \text{for } t < 0 \end{cases}$

2.4. Conclusions from the t -Test

One of the assumptions of the t -test is that the sample size n is large. Typically, $n \geq 30$ is required. This assumption is violated in some of the columns of **Table A1**. However the table is merely for illustrative purposes. Even so, it shows the power of the method to correctly ascertain the order of the best fitting polynomial because one can hardly fail to notice the precipitous drop in P -value in the 3rd difference column (titled **Diff 3**). The plot in **Figure 5** showing the P -values of the 5 difference columns graphically displays this drop.

The insignificant P -value in the 3rd difference column implies that with high probability the order of the best fitting polynomial should be $m = 2$, *i.e.* a quadratic polynomial is the best fit polynomial to the data in **Table A1**.

2.5. Comparison to Classical Regression

In classical polynomial regression, R^2 values instead of P -values represent the goodness of fit. Using Microsoft Excel's built-in program to generate R^2 values to polynomial fits, I generated **Table 4** by restricting the highest degree m of the best fitting polynomial $a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$ for the noisy data in **Table A1**.

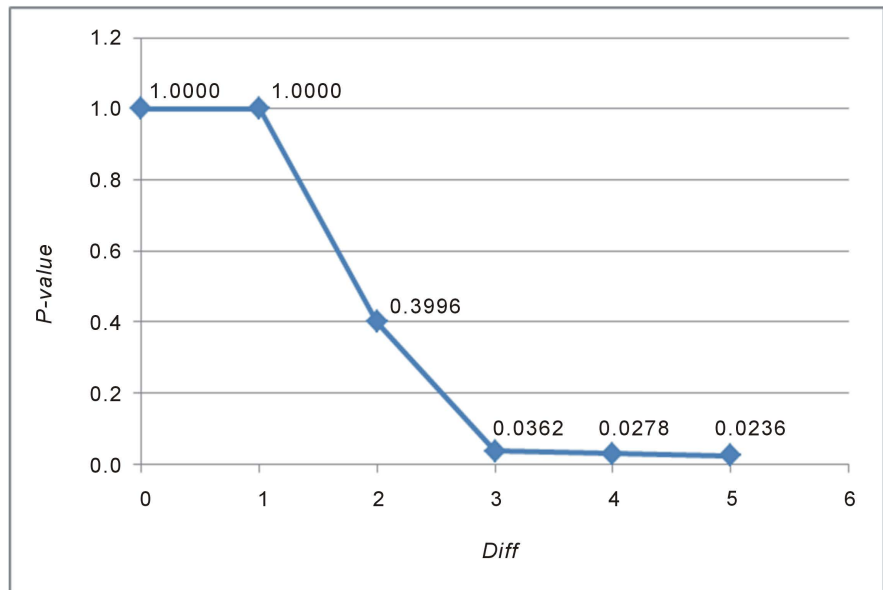


Figure 5. Graph of P -value vs. $Diff$.

Table 4. R^2 values in classical regression as a function of degree m .

m	1 (Linear)	2 (Quadratic)	3 (Cubic)	4	5
R^2 value	0.9403	0.9999	0.9999	0.9999	0.9999

Table 4 shows that the R^2 values used in classical regression do not have the sensitivity that the P -values with t -tests possess. The clear drop that occurs in the P -values as shown in Figure 5 does not have a corresponding sharp rise in R^2 values in Table 4. Also, unlike classical regression which requires R^2 values to be fitted with every polynomial the Finite Difference Regression is a one-shot test that gives a clear estimate of the order of the best fitting polynomial by testing the significance of the probabilities (P -values). The first successive difference column whose P -value is less than the *significance level* (a belief threshold for the viable model) comes out to be the best fitting polynomial model.

Unlike classical regression, however, this method requires a large sample size n , preferably one that is a power of 2. Also, because the degrees of freedom get halved at each successive difference and the degrees of freedom have to remain positive, the highest order polynomial that can be fit by this method is of the order $\log_2 n$. This is generally acceptable because the order of the best fitting polynomials should by definition be small constants. Other classical model order selection heuristics like AIC or BIC (see [3]) would also heavily penalize high order models.

2.6. Determining the Coefficients

Once the order of the best fitting polynomial has been found using the t -test in combination with finite differences, I will evaluate the coefficients of the best fitting polynomial using the method of successive finite differences as detailed

in Section 1.

In the non-noisy situation described in Section 1, the matter was simple: there was a difference row of all zeros and therefore the previous difference row was a constant that could be used to determine the coefficient of the highest degree. In the noisy case, the issue is complicated because there is no such row that consists of all zeros and therefore the previous difference row is not a constant. However, the intuitive idea behind the method of Finite Difference Regression (see the 2nd paragraph of Section 2), led me to *assume* that the sample mean \bar{x} would be an excellent choice for the previous difference row constant, since it is the unique number that minimizes the sum of squares of the residuals. From **Table A1**, the all zero column is the 3rd difference column, and therefore $\bar{x} = 4.29$ (the sample mean of the 2nd difference column in **Table A1**) corresponds to the constant difference. Thus as described in Section 1, the highest term should be $\bar{x}x^2/2! = 2.145x^2$; the next term determined from the sample mean of the “constant” column after subtracting the values of the highest term yields $-7.464x$; and finally the constant term is 78.69. Thus $\hat{y} = 2.145x^2 - 7.464x + 78.69$ is the estimated polynomial model.

Table A2 in **Appendix 2** compares the noisy data to the best fit polynomial estimate using the method of successive finite differences. Note that the residuals are not very small; however, their average turns out to be -0.00078 indicating a good fit. The polynomial fit is overlaid on the noisy y data in **Figure 6**, in which the blue line shows the noisy data, and the dashed red line depicts the values of the estimated \hat{y} .

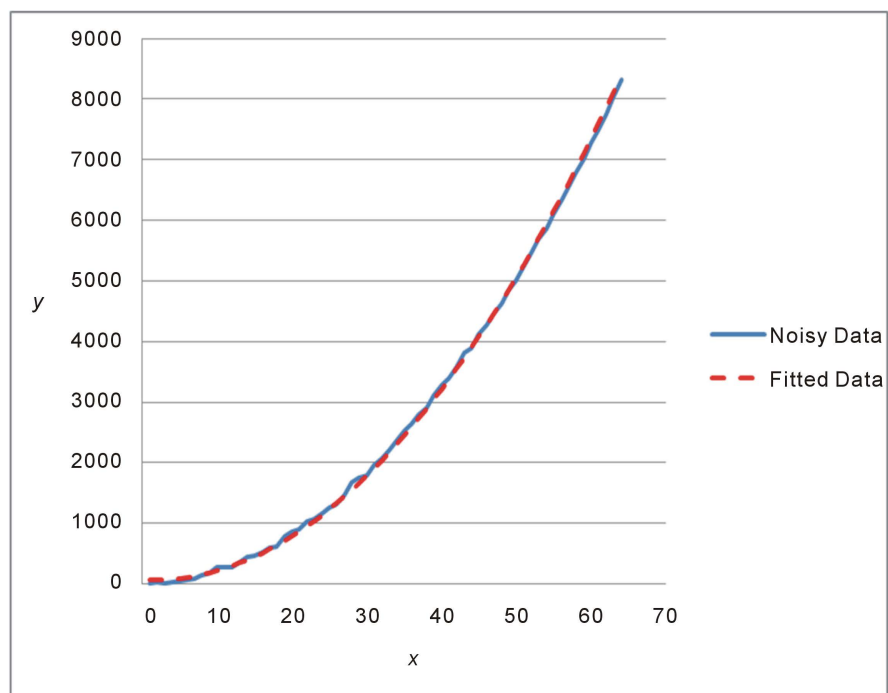


Figure 6. Graph of best fitting polynomial $\hat{y} = 2.145x^2 - 7.464x + 78.69$ overlaid on the noisy observations of **Table A1**.

2.7. Unbiasedness and Consistency of the Coefficient Estimates

In this section I prove that the coefficients estimated by the Finite Difference Regression method as detailed in Section 2.6 are unbiased and consistent. The properties of unbiasedness and consistency are very important for estimators (see [6] p. 393). The unbiasedness property shows that there is no systematic offset between the estimate and the actual value of a parameter. It is proved by showing that the mean of the difference between the estimate and the actual value is zero. The consistency property shows that the estimates get better with more observations. It is proved by showing that a non-zero difference between the estimate and the actual value becomes highly improbable with a large number of observations. The analysis focuses on the estimate of the leading coefficient *i.e.* that associated with highest degree because the asymptotic properties of the polynomial fit will be most sensitive to that value.

Consider the following table that shows the symbolic formulas for the successive differences for the n observations y_1, y_2, \dots, y_n :

For subsequent ease of notation, I introduce a change of variables $z_i \stackrel{\text{def}}{=} y_{i+1}$ for $0 \leq i \leq n-1$. With this renaming/re-indexing, **Table 5** can be rewritten as (note that the row indices have also changed):

For any m , $0 \leq m \leq n-1$, the m^{th} Diff column consists of expressions in rows indexed from m through $n-1$. Using mathematical induction, the m^{th} Diff column can be shown to be:

If the underlying polynomial is an m^{th} order polynomial $f(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$, then in the case of *noiseless* observations, every row of **Table 7** evaluates to the constant $m!a_m$. In particular, for the k^{th} row in the m^{th} Diff column shown in **Table 7** one gets:

Table 5. Successive differences for n observations y_1, y_2, \dots, y_n .

	<i>Diff0</i>	<i>Diff1</i>	<i>Diff2</i>	...
1st Row	y_1			
2nd Row	y_2	$y_2 - y_1$		
3rd Row	y_3	$y_3 - y_2$	$y_3 - 2y_2 + y_1$	
...	
n^{th} Row	y_n	$y_n - y_{n-1}$	$y_n - 2y_{n-1} + y_{n-2}$	

Table 6. Changing y_{i+1} from **Table 5** to z_i for $0 \leq i \leq n-1$.

	<i>Diff0</i>	<i>Diff1</i>	<i>Diff2</i>	...
0th Row	z_0			
1st Row	z_1	$z_1 - z_0$		
2nd Row	z_2	$z_2 - z_1$	$z_2 - 2z_1 + z_0$	
...	
$(n-1)^{\text{th}}$ Row	z_{n-1}	$z_{n-1} - z_{n-2}$	$z_{n-1} - 2z_{n-2} + z_{n-3}$	

Table 7. The m^{th} diff column.

	<i>Diff m</i>
m^{th} Row	$\binom{m}{0}z_m - \binom{m}{1}z_{m-1} + \binom{m}{2}z_{m-2} + \dots + (-1)^m \binom{m}{m}z_0$
$(m+1)^{\text{th}}$ Row	$\binom{m}{0}z_{m+1} - \binom{m}{1}z_m + \binom{m}{2}z_{m-1} + \dots + (-1)^m \binom{m}{m}z_1$
...	...
k^{th} Row	$\binom{m}{0}z_k - \binom{m}{1}z_{k-1} + \binom{m}{2}z_{k-2} + \dots + (-1)^m \binom{m}{m}z_{k-m}$
...	...
$(n-1)^{\text{th}}$ Row	$\binom{m}{0}z_{n-1} - \binom{m}{1}z_{n-2} + \binom{m}{2}z_{n-3} + \dots + (-1)^m \binom{m}{m}z_{n-1-m}$

$$\begin{aligned}
 & \binom{m}{0}z_k - \binom{m}{1}z_{k-1} + \binom{m}{2}z_{k-2} + \dots + (-1)^m \binom{m}{m}z_{k-m} \\
 &= \binom{m}{0}y_{k+1} - \binom{m}{1}y_k + \binom{m}{2}y_{k-1} + \dots + (-1)^m \binom{m}{m}y_{k+1-m} \tag{1} \\
 &= \binom{m}{0}f(k+1) - \binom{m}{1}f(k) + \binom{m}{2}f(k-1) + \dots + (-1)^m \binom{m}{m}f(k+1-m) \\
 &= m!a_m
 \end{aligned}$$

In case of noisy observations, let each re-indexed observation z now be corrupted by a sequence of *additive zero-mean independent identically distributed* random variables N_0, N_1, \dots, N_{n-1} , each with variance σ^2 . In other words:

$$\begin{aligned}
 \mathbb{E}[N_i] &= 0 \text{ and } \mathbb{E}[N_i^2] = \sigma^2 \text{ for } 0 \leq i < n \\
 \mathbb{E}[N_i N_j] &= \mathbb{E}[N_i] \mathbb{E}[N_j] = 0 \text{ for } 0 \leq i < j < n
 \end{aligned}$$

where $\mathbb{E}[\]$ denotes the expectation operator. (See for example [7] pp. 220-223.)

This addition of random variables makes the z 's also random variables denoted now as Z , and so $Z_i = f(i+1) + N_i$ for $0 \leq i < n$. In this case, the k^{th} row in the m^{th} Diff column shown in **Table 7** is the random variable R_k where:

$$\begin{aligned}
 R_k &= \binom{m}{0}Z_k - \binom{m}{1}Z_{k-1} + \binom{m}{2}Z_{k-2} + \dots + (-1)^m \binom{m}{m}Z_{k-m} \\
 &= \binom{m}{0}N_k - \binom{m}{1}N_{k-1} + \binom{m}{2}N_{k-2} + \dots + (-1)^m \binom{m}{m}N_{k-m} \\
 &\quad + \binom{m}{0}f(k+1) - \binom{m}{1}f(k) + \binom{m}{2}f(k-1) + \dots + (-1)^m \binom{m}{m}f(k+1-m) \tag{2} \\
 &= \binom{m}{0}N_k - \binom{m}{1}N_{k-1} + \binom{m}{2}N_{k-2} + \dots + (-1)^m \binom{m}{m}N_{k-m} + m!a_m
 \end{aligned}$$

where, the last step of Equation (2) follows from the identity in Equation (1).

Let the random variable $\hat{A}_m(n)$ denote the estimate of a_m with the n observations. Then as described in Section 2.6, since there are $(n-m)$ rows in the m^{th} Diff column, the sample mean of the \mathbf{R} 's divided by $m!$ provides the required estimate. In other words,

$$\hat{A}_m(n) = \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} \mathbf{R}_k \quad (3)$$

Now, define the *reduced* row random variable \mathbf{Q}_k for the m^{th} Diff column shown in **Table 7** as:

$$\mathbf{Q}_k \stackrel{\text{def}}{=} \mathbf{R}_k - m!a_m \text{ for } m \leq k \leq n-1 \quad (4)$$

Then, substituting back into Equation (3):

$$\begin{aligned} \hat{A}_m(n) &= \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} \mathbf{Q}_k + \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} m!a_m \\ &= \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} \mathbf{Q}_k + \frac{m!a_m(n-m)}{m!(n-m)} \\ &= \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} \mathbf{Q}_k + a_m \end{aligned}$$

Which, upon rearranging yields:

$$\hat{A}_m(n) - a_m = \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} \mathbf{Q}_k \quad (5)$$

Note that it is the statistics of the expression $\hat{A}_m(n) - a_m$ on the left-hand side of Equation (5) that will prove the properties of unbiasedness and consistency of the estimate $\hat{A}_m(n)$. Also note that by definition (4) and from Equation (2), the reduced row random variable \mathbf{Q}_k for the m^{th} Diff column can be explicitly written as:

$$\mathbf{Q}_k = \binom{m}{0} N_k - \binom{m}{1} N_{k-1} + \binom{m}{2} N_{k-2} + \cdots + (-1)^m \binom{m}{m} N_{k-m} \quad (6)$$

2.7.1. Result 1: $\hat{A}_m(n)$ Is an Unbiased Estimate of a_m

Taking the expectation operator $\mathbb{E}[\]$ on both sides of Equation (6):

$$\begin{aligned} \mathbb{E}[\mathbf{Q}_k] &= \mathbb{E} \left[\binom{m}{0} N_k - \binom{m}{1} N_{k-1} + \binom{m}{2} N_{k-2} + \cdots + (-1)^m \binom{m}{m} N_{k-m} \right] \\ &= \binom{m}{0} \mathbb{E}[N_k] - \binom{m}{1} \mathbb{E}[N_{k-1}] + \cdots + (-1)^m \binom{m}{m} \mathbb{E}[N_{k-m}] = 0 \end{aligned} \quad (7)$$

The last step in Equation (7) follows from the fact that the random variables N_i are all zero-mean. Then, taking the expectation operator on both sides of Equation (5), and using the result from Equation (7):

$$\mathbb{E}[\hat{A}_m(n) - a_m] = \frac{1}{m!(n-m)} \sum_{k=m}^{n-1} \mathbb{E}[\mathbf{Q}_k] = 0 \quad (8)$$

Hence, $\mathbb{E}[\hat{A}_m(n) - a_m] = 0$ for all n . This proves the unbiasedness property of $\hat{A}_m(n)$. ■

2.7.2. Result 2: $\hat{A}_m(n)$ Is a Consistent Estimate of a_m

Since the random variable $\hat{A}_m(n) - a_m$ has been shown to be of zero-mean by virtue of its unbiasedness property, the variance of this random variable is simply $\mathbb{E}\left[\left(\hat{A}_m(n) - a_m\right)^2\right]$. I show the consistency property of the estimate

$\hat{A}_m(n)$ by way of deriving an upper-bound for $\mathbb{E}\left[\left(\hat{A}_m(n) - a_m\right)^2\right]$ and showing that this upper-bound vanishes to 0 in the limit $n \rightarrow \infty$.

Writing out the reduced row random variables Q_k for each row k for the m^{th} Diff column from Equation (6) one gets the following series of $(n - m)$ equations:

$$\begin{aligned} Q_m &= \binom{m}{0} N_m - \binom{m}{1} N_{m-1} + \binom{m}{2} N_{m-2} + \dots + (-1)^m \binom{m}{m} N_0 \\ Q_{m+1} &= \binom{m}{0} N_{m+1} - \binom{m}{1} N_m + \binom{m}{2} N_{m-1} + \dots + (-1)^m \binom{m}{m} N_1 \\ &\dots \\ Q_k &= \binom{m}{0} N_k - \binom{m}{1} N_{k-1} + \binom{m}{2} N_{k-2} + \dots + (-1)^m \binom{m}{m} N_{k-m} \\ &\dots \\ Q_{n-1} &= \binom{m}{0} N_{n-1} - \binom{m}{1} N_{n-2} + \binom{m}{2} N_{n-3} + \dots + (-1)^m \binom{m}{m} N_{n-1-m} \end{aligned}$$

Summing up these $(n - m)$ equations column by column:

$$\sum_{k=m}^{n-1} Q_k = \binom{m}{0} S_m^{n-1} - \binom{m}{1} S_{m-1}^{n-2} + \binom{m}{2} S_{m-2}^{n-3} + \dots + (-1)^m \binom{m}{m} S_0^{n-1-m} \tag{9}$$

where for $(n - 1 - m) \leq i \leq (n - 1)$ and $0 \leq j \leq m$ define:

$$S_j^i \stackrel{\text{def}}{=} \sum_{k=j}^i N_k \tag{10}$$

Squaring both sides of Equation (10), and then taking expectations:

$$\begin{aligned} \mathbb{E}\left[\left(S_j^i\right)^2\right] &= \mathbb{E}\left[\left(\sum_{k=j}^i N_k\right)^2\right] = \mathbb{E}\left[\sum_{k=j}^i N_k^2 + \mathbf{SCT}\right] \\ &= \sum_{k=j}^i \left(\mathbb{E}\left[N_k^2\right] + \mathbb{E}[\mathbf{SCT}]\right) = \sum_{k=j}^i \sigma^2 = (i - j + 1) \sigma^2 \end{aligned} \tag{11}$$

In Equation (11) I have invoked the zero-mean, independent nature of the N 's to cause the "sum of cross-terms" term \mathbf{SCT} to vanish.

Define a new sequence of $(m + 1)$ random variables T_0, T_1, \dots, T_m as follows:

$$T_i \stackrel{\text{def}}{=} (-1)^i \binom{m}{i} S_{m-i}^{n-1-i} \text{ for } 0 \leq i \leq m \tag{12}$$

By squaring both sides of definition (12), and then taking expectations:

$$\mathbb{E}[T_i^2] = \mathbb{E}\left[(-1)^{2i} \binom{m}{i}^2 (S_{m-i}^{n-1-i})^2\right] = \binom{m}{i}^2 \mathbb{E}\left[(S_{m-i}^{n-1-i})^2\right] = \binom{m}{i}^2 (n-m)\sigma^2 \quad (13)$$

where, the final step in Equation (13) follows from Equation (11).

Using definition (12), Equation (9) can be rewritten as:

$$\sum_{k=m}^{n-1} Q_k = T_0 + T_1 + \dots + T_m \quad (14)$$

Then as $\mathbb{E}\left[(T_0 + T_1 + \dots + T_m)^2\right] \leq (m+1)(\mathbb{E}[T_0^2] + \mathbb{E}[T_1^2] + \dots + \mathbb{E}[T_m^2])$ (see

Appendix 1, squaring both sides of Equation (14), and then taking expectations:

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{k=m}^{n-1} Q_k\right)^2\right] &= \mathbb{E}\left[(T_0 + T_1 + \dots + T_m)^2\right] \leq (m+1)(\mathbb{E}[T_0^2] + \mathbb{E}[T_1^2] + \dots + \mathbb{E}[T_m^2]) \\ &= (m+1)(n-m)\sigma^2 \left(\binom{m}{0}^2 + \binom{m}{1}^2 + \dots + \binom{m}{m}^2\right) \quad [\text{From Equation (13)}] \\ &= (m+1)(n-m)\sigma^2 \binom{2m}{m} \quad [\text{From binomial identity (see [8] pg.64)}] \end{aligned} \quad (15)$$

Finally, by squaring both sides of Equation (5), and then taking expectations I obtain the following inequality from inequality (15):

$$\mathbb{E}\left[(\hat{A}_m(n) - a_m)^2\right] = \frac{1}{(m!)^2 (n-m)^2} \mathbb{E}\left[\left(\sum_{k=m}^{n-1} Q_k\right)^2\right] \leq \frac{(m+1)(2m)! \sigma^2}{(m!)^4 (n-m)} \quad (16)$$

where, the last step in (16) is a rewriting of the binomial coefficient $\binom{2m}{m}$ in terms of factorials and subsequent simplification.

It can be seen that the factor $(m+1)(2m)!/(m!)^4$ on the right hand side of inequality (16) is a super-exponentially decreasing function of m (above a constant) due to presence of the large $(m!)^4$ factor in its denominator. This factor is evaluated for a few values of m in **Table 8**.

Table 8 shows that:

$$\mathbb{E}\left[(\hat{A}_m(n) - a_m)^2\right] \leq \frac{4.5 \times \sigma^2}{(n-m)} \quad (17)$$

For the Finite Difference Regression method, $m \leq \log_2 n$ (see Section 0 and the comments in Section 0). Thus, the right-hand-side denominator $n-m \geq n - \log_2 n$ in inequality (17) and therefore,

$$\mathbb{E}\left[(\hat{A}_m(n) - a_m)^2\right] \leq \frac{4.5 \times \sigma^2}{(n - \log_2 n)} \quad (18)$$

Taking limits on both sides of inequality (18),

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[(\hat{A}_m(n) - a_m)^2\right] \leq \lim_{n \rightarrow \infty} \frac{4.5 \times \sigma^2}{(n - \log_2 n)} = 0 \quad (19)$$

Table 8. The factor $(m+1)(2m)!/(m!)^4$ evaluated for a few values of m .

m	0	1	2	3	4	5	6
$(m+1)(2m)!/(m!)^4$	1.000	4.00	4.500	2.222	0.608	0.105	0.012
m	7	8	9	10	11	12	13
$(m+1)(2m)!/(m!)^4$	1.1×10^{-3}	7.1×10^{-5}	3.7×10^{-6}	1.5×10^{-7}	5.3×10^{-9}	1.5×10^{-10}	3.8×10^{-12}

However, as $\mathbb{E}\left[\left(\hat{A}_m(n) - a_m\right)^2\right] \geq 0$, inequality (19) implies that:

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[\left(\hat{A}_m(n) - a_m\right)^2\right] = 0 \quad (20)$$

Then invoking Chebyshev's inequality (see [7] p. 233, [9] p. 151) yields:

$$\lim_{n \rightarrow \infty} \text{Prob}\left[\left|\hat{A}_m(n) - a_m\right| > \varepsilon\right] \leq \frac{\lim_{n \rightarrow \infty} \mathbb{E}\left[\left(\hat{A}_m(n) - a_m\right)^2\right]}{\varepsilon^2} = 0, \text{ for any } \varepsilon > 0$$

This proves the consistency property of $\hat{A}_m(n)$. ■

2.7.3. A Note on the Asymptotic Properties of the Fit

Inequality (16) and the values in **Table 8** show that the higher the degree m of the polynomial, the lower is the variance of the estimate of its leading coefficient. This excellent property implies that the asymptotic properties of the fit (that are most sensitive to this leading coefficient) are stable, in the sense that the value of the asymptote becomes less varying with increasing degrees of the fitting polynomial.

3. Conclusions and Further Work

In this paper I have presented a novel polynomial regression method for uniformly sampled data points called the method of Finite Difference Regression. Unlike classical regression methods in which the order of the best fitting polynomial model is unknown and is estimated from the R^2 values, I have shown how the t -test can be combined with the method of finite differences to provide a more sensitive and objective measure of the order of the best fitting polynomial. Furthermore, I have carried forward the method of Finite Difference Regression, reemploying the finite differences obtained during order determination, to provide estimates of the coefficients of the best fitting polynomial. I have shown that not only are these coefficients unbiased and consistent, but also that the asymptotic properties of the fit get better with increasing degrees of the fitting polynomial.

At least three other avenues of further research remain to be explored:

- 1) The polynomial coefficients obtained by this method do not minimize the sum of the squared residuals. Is there an objective function that they *do* minimize?

2) The classical regression methods work on non-uniformly sampled data sets. Extending this method to non-uniformly sampled data should be possible.

3) Automatically finding a good t -test significance level *i.e.* the precipitously low P -value that sets the order of the best fitting polynomial (as described in Section 2.4) remains an open problem. For this, heuristics in the spirit of AIC or BIC methods (see [3]) in the classical case may perhaps be required.

Acknowledgements

I thank Prof. Brad Lucier for his interest and for the helpful discussions and comments that made this a better paper. I am also grateful to Dr. Jan Frydendahl for encouraging me to pursue this research activity.

References

- [1] Gauss, C.F. (1857) Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections. Davis, C.H., Trans., Little, Brown & Co., Boston, MA.
- [2] Peck, R., Olsen, C. and Devore, J. (2005) Introduction to Statistics and Data Analysis. 2nd Edition, Thomson Brooks/Cole, Belmont, CA.
- [3] Burnham, K.P. and Anderson, D.R. (2004) Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods and Research*, **33**, 261-304. <https://doi.org/10.1177/0049124104268644>
- [4] Newton, I. (1848) Newton's Principia: The Mathematical Principles of Natural Philosophy. Motte, A., Trans., Daniel Adee, New York.
- [5] Hämmerlin, G. and Hoffmann, K.-H. (1991) Numerical Mathematics. Schumaker, L.L., Trans., Springer, New York. <https://doi.org/10.1007/978-1-4612-4442-4>
- [6] Cover, T.M. and Thomas, J.A. (2006) Elements of Information Theory. 2nd Edition, John Wiley & Sons, Inc., New York.
- [7] Feller, W. (1968) An Introduction to Probability Theory and Its Applications. 3rd Edition, Vol. I, John Wiley & Sons, Inc., New York.
- [8] Tucker, A. (1980) Applied Combinatorics. John Wiley & Sons, Inc., New York.
- [9] Feller, W. (1971) An Introduction to Probability Theory and Its Applications. 2nd Edition, Vol. II, John Wiley & Sons, Inc., New York.
- [10] Student (1908) The Probable Error of a Mean. *Biometrika*, **6**, 1-25. <https://doi.org/10.1093/biomet/6.1.1>

Appendix 1

To show that for any sequence of $(m+1)$ random variables T_0, T_1, \dots, T_m :

$$\mathbb{E}\left[(T_0 + T_1 + \dots + T_m)^2\right] \leq (m+1)\left(\mathbb{E}[T_0^2] + \mathbb{E}[T_1^2] + \dots + \mathbb{E}[T_m^2]\right)$$

The proof follows from a variance-like consideration. Given the random variables T_0, T_1, \dots, T_m and any random variable \bar{T} , the sum of squares *random variable* $\sum_{i=0}^m (T_i - \bar{T})^2$ is always non-negative. Its expected value $\mathbb{E}\left[\sum_{i=0}^m (T_i - \bar{T})^2\right]$ is therefore always non-negative. Choosing the random variable $\bar{T} \stackrel{\text{def}}{=} \sum_{i=0}^m T_i / (m+1)$ i.e. the “mean” of T_0, T_1, \dots, T_m proves the inequality as follows:

$$\begin{aligned} 0 &\leq \mathbb{E}\left[\sum_{i=0}^m (T_i - \bar{T})^2\right] \\ &= \mathbb{E}\left[\sum_{i=0}^m (T_i^2 - 2\bar{T}T_i + \bar{T}^2)\right] \\ &= \mathbb{E}\left[\sum_{i=0}^m T_i^2\right] - 2\mathbb{E}\left[\sum_{i=0}^m \bar{T}T_i\right] + \mathbb{E}\left[\sum_{i=0}^m \bar{T}^2\right] \\ &= \sum_{i=0}^m \mathbb{E}[T_i^2] - 2\mathbb{E}\left[\bar{T}\sum_{i=0}^m T_i\right] + \mathbb{E}\left[\bar{T}^2\sum_{i=0}^m 1\right] \\ &= \sum_{i=0}^m \mathbb{E}[T_i^2] - 2\mathbb{E}\left[\bar{T} \times (m+1)\bar{T}\right] + \mathbb{E}\left[\bar{T}^2 \times (m+1)\right] \quad (\text{from defn. of } \bar{T}) \\ &= \sum_{i=0}^m \mathbb{E}[T_i^2] - 2(m+1)\mathbb{E}[\bar{T}^2] + (m+1)\mathbb{E}[\bar{T}^2] \\ &= \sum_{i=0}^m \mathbb{E}[T_i^2] - (m+1)\mathbb{E}[\bar{T}^2] \\ &= \sum_{i=0}^m \mathbb{E}[T_i^2] - (m+1)\mathbb{E}\left[\left(\frac{\sum_{i=0}^m T_i}{m+1}\right)^2\right] \quad (\text{from defn. of } \bar{T}) \\ &= \sum_{i=0}^m \mathbb{E}[T_i^2] - \frac{(m+1)}{(m+1)^2} \mathbb{E}\left[(T_0 + T_1 + \dots + T_m)^2\right] \end{aligned}$$

Multiplying both sides of the above inequality by $(m+1)$ and rearranging yields the required result. ■

Appendix 2

Table A1. Example showing how the finite difference regression method works.

x	y	<i>Noise</i>	$y + \text{Noise}$	<i>Diff1</i>	<i>Diff2</i>	<i>Diff3</i>	<i>Diff4</i>	<i>Diff5</i>
1	6	-3.46	2.54					
2	13	11.08	24.08	21.53				
3	24	-10.78	13.22	-10.86	-32.39			
4	39	-12.23	26.77	13.55	24.41	56.81		
5	58	-15.22	42.78	16.01	2.46	-21.95	-78.76	
6	81	-20.56	60.44	17.65	1.64	-0.82	21.13	99.89
7	108	-18.25	89.75	29.31	11.66	10.02	10.84	-10.29
8	139	0.81	139.81	50.06	20.75	9.09	-0.94	-11.78
9	174	2.70	176.70	36.88	-13.18	-33.92	-43.01	-42.07
10	213	65.28	278.28	101.58	64.70	77.87	111.80	154.81
11	256	14.08	270.08	-8.20	-109.78	-174.47	-252.34	-364.14
12	303	-21.65	281.35	11.27	19.47	129.25	303.72	556.06
13	354	2.26	356.26	74.91	63.64	44.17	-85.07	-388.79
14	409	31.42	440.42	84.16	9.25	-54.39	-98.57	-13.49
15	468	-4.20	463.80	23.37	-60.79	-70.04	-15.65	82.92
16	531	-11.58	519.42	55.62	32.25	93.04	163.08	178.72
17	598	-4.56	593.44	74.03	18.41	-13.84	-106.88	-269.95
18	669	-52.78	616.22	22.77	-51.26	-69.66	-55.82	51.06
19	744	40.58	784.58	168.37	145.59	196.85	266.51	322.33
20	823	40.27	863.27	78.69	-89.68	-235.27	-432.12	-698.63
21	906	-1.21	904.79	41.52	-37.17	52.51	287.78	719.90
22	993	41.62	1034.62	129.83	88.32	125.49	72.98	-214.80
23	1084	-4.49	1079.51	44.89	-84.94	-173.26	-298.75	-371.73
24	1179	-8.48	1170.52	91.01	46.12	131.07	304.32	603.07
25	1278	-11.62	1266.38	95.86	4.84	-41.28	-172.35	-476.67
26	1381	-55.64	1325.36	58.99	-36.87	-41.71	-0.43	171.92
27	1488	-13.86	1474.14	148.78	89.79	126.66	168.37	168.80
28	1599	72.21	1671.21	197.07	48.30	-41.49	-168.15	-336.52
29	1714	35.99	1749.99	78.78	-118.29	-166.59	-125.10	43.05
30	1833	-46.35	1786.65	36.66	-42.12	76.18	242.77	367.87
31	1956	4.35	1960.35	173.70	137.04	179.15	102.97	-139.80
32	2083	-6.83	2076.17	115.82	-57.87	-194.91	-374.07	-477.04
33	2214	-5.17	2208.83	132.66	16.84	74.71	269.62	643.69
34	2349	10.94	2359.94	151.10	18.44	1.61	-73.10	-342.73
35	2488	32.31	2520.31	160.38	9.27	-9.17	-10.77	62.33

Continued

36	2631	15.10	2646.10	125.79	-34.59	-43.86	-34.70	-23.92
37	2778	10.35	2788.35	142.24	16.45	51.04	94.91	129.60
38	2929	-14.74	2914.26	125.91	-16.33	-32.79	-83.83	-178.74
39	3084	39.17	3123.17	208.91	83.00	99.33	132.12	215.95
40	3243	36.96	3279.96	156.80	-52.12	-135.12	-234.45	-366.57
41	3406	-4.73	3401.27	121.31	-35.49	16.63	151.74	386.19
42	3573	-6.60	3566.40	165.13	43.82	79.31	62.68	-89.06
43	3744	65.08	3809.08	242.68	77.55	33.74	-45.57	-108.25
44	3919	-22.02	3896.98	87.90	-154.77	-232.33	-266.06	-220.49
45	4098	13.68	4111.68	214.70	126.80	281.57	513.90	779.96
46	4281	-6.64	4274.36	162.67	-52.03	-178.82	-460.39	-974.29
47	4468	-5.14	4462.86	188.50	25.83	77.86	256.68	717.07
48	4659	-37.01	4621.99	159.13	-29.37	-55.20	-133.06	-389.74
49	4854	6.42	4860.42	238.43	79.30	108.67	163.88	296.94
50	5053	-46.83	5006.17	145.75	-92.68	-171.98	-280.66	-444.53
51	5256	7.28	5263.28	257.11	111.36	204.04	376.03	656.68
52	5463	-8.25	5454.75	191.47	-65.64	-177.00	-381.04	-757.07
53	5674	18.11	5692.11	237.36	45.89	111.53	288.53	669.57
54	5889	-20.01	5868.99	176.89	-60.48	-106.37	-217.90	-506.42
55	6108	5.34	6113.34	244.35	67.46	127.94	234.30	452.20
56	6331	1.67	6332.67	219.33	-25.02	-92.48	-220.42	-454.72
57	6558	8.07	6566.07	233.40	14.07	39.09	131.58	351.99
58	6789	7.01	6796.01	229.94	-3.46	-17.54	-56.63	-188.21
59	7024	-6.75	7017.25	221.24	-8.69	-5.23	12.31	68.94
60	7263	14.76	7277.76	260.51	39.27	47.96	53.19	40.88
61	7506	-18.95	7487.05	209.29	-51.22	-90.48	-138.44	-191.63
62	7753	-2.33	7750.67	263.61	54.32	105.54	196.03	334.47
63	8004	24.70	8028.70	278.03	14.42	-39.91	-145.45	-341.47
64	8259	57.11	8316.11	287.41	9.38	-5.04	34.87	180.32
<i>n</i>			64	63	62	61	60	59
\bar{x}			2833.75	131.96	4.29	0.68	-1.03	1.93
<i>s</i>			2517.12	83.47	63.10	113.72	211.20	399.59
<i>t</i>			9.01	12.55	0.54	0.05	-0.04	0.04
<i>df</i>			63	31	15	7	3	1
<i>P-value</i>			1.0000	1.0000	0.3996	0.0362	0.0278	0.0236

Table A2. The noisy data compared to the best fit polynomial estimate $\hat{y} = 2.145x^2 - 7.464x + 78.69$ determined from the successive finite differences method.

<i>x</i>	<i>y + Noise</i>	\hat{y}	<i>Residual</i>	<i>x</i>	<i>y + Noise</i>	\hat{y}	<i>Residual</i>
1	2.54	73.371	-70.831	33	2208.83	2168.283	40.547
2	24.08	72.342	-48.262	34	2359.94	2304.534	55.406
3	13.22	75.603	-62.383	35	2520.31	2445.075	75.235
4	26.77	83.154	-56.384	36	2646.10	2589.906	56.194
5	42.78	94.995	-52.215	37	2788.35	2739.027	49.323
6	60.44	111.126	-50.686	38	2914.26	2892.438	21.822
7	89.75	131.547	-41.797	39	3123.17	3050.139	73.031
8	139.81	156.258	-16.448	40	3279.96	3212.130	67.830
9	176.70	185.259	-8.559	41	3401.27	3378.411	22.859
10	278.28	218.550	59.730	42	3566.40	3548.982	17.418
11	270.08	256.131	13.949	43	3809.08	3723.843	85.237
12	281.35	298.002	-16.652	44	3896.98	3902.994	-6.014
13	356.26	344.163	12.097	45	4111.68	4086.435	25.245
14	440.42	394.614	45.806	46	4274.36	4274.166	0.194
15	463.80	449.355	14.445	47	4462.86	4466.187	-3.327
16	519.42	508.386	11.034	48	4621.99	4662.498	-40.508
17	593.44	571.707	21.733	49	4860.42	4863.099	-2.679
18	616.22	639.318	-23.098	50	5006.17	5067.990	-61.820
19	784.58	711.219	73.361	51	5263.28	5277.171	-13.891
20	863.27	787.410	75.860	52	5454.75	5490.642	-35.892
21	904.79	867.891	36.899	53	5692.11	5708.403	-16.293
22	1034.62	952.662	81.958	54	5868.99	5930.454	-61.464
23	1079.51	1041.723	37.787	55	6113.34	6156.795	-43.455
24	1170.52	1135.074	35.446	56	6332.67	6387.426	-54.756
25	1266.38	1232.715	33.665	57	6566.07	6622.347	-56.277
26	1325.36	1334.646	-9.286	58	6796.01	6861.558	-65.548
27	1474.14	1440.867	33.273	59	7017.25	7105.059	-87.809
28	1671.21	1551.378	119.832	60	7277.76	7352.850	-75.090
29	1749.99	1666.179	83.811	61	7487.05	7604.931	-117.881
30	1786.65	1785.270	1.380	62	7750.67	7861.302	-110.632
31	1960.35	1908.651	51.699	63	8028.70	8121.963	-93.263
32	2076.17	2036.322	39.848	64	8316.11	8386.914	-70.804