Name: Ashutosh Rabia Roll no.: 220238 (BS-MTH)

1 Dataset

```
t_data: [0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.11 0.12 0.13 0.14 0.15
 0.16 0.17 0.18 0.19 0.2  0.21 0.22 0.23 0.24 0.25 0.26 0.27 0.28 0.29
 0.31 0.32 0.33 0.34 0.35 0.36 0.37 0.38 0.39 0.4  0.41 0.42 0.43 0.44
 0.45 0.46 0.47 0.48 0.49 0.51 0.52 0.53 0.54 0.55 0.56 0.57 0.58 0.59
 0.6  0.61 0.62 0.63 0.64 0.65 0.66 0.67 0.68 0.69 0.71 0.72 0.73 0.74
 0.75 0.76 0.77 0.78 0.79 0.8  0.81 0.82 0.83 0.84 0.85 0.86 0.87 0.88
 0.89 0.91 0.92 0.93 0.94 0.95 0.96 0.97 0.98 0.99 1.  ]
y_data: [ 3.5145  3.6702  3.6325  3.7286  3.5693  3.7675  3.6442  3.7747  3.8626
  3.7776  3.9294  3.9816  3.9611  4.063   4.1039  4.2111  4.2615  4.141
  4.2011  4.3101  4.312   4.2744  4.3409  4.3845  4.6733  4.5197  4.6075
  4.5709  4.6902  4.7305  4.8956  4.8101  5.0101  4.9595  5.1347  4.9934
  5.1542  5.1346  5.1899  5.3145  5.5038  5.4142  5.4439  5.5385  5.6056
  5.7527  5.7775  5.8537  5.9821  5.9717  6.1009  6.3081  6.1539  6.3737
  6.4645  6.409   6.5207  6.8695  6.9168  7.0009  7.1238  6.9516  7.1404
  7.3612  7.387   7.6027  7.5852  7.7302  7.9688  7.9247  8.0991  8.1753
  8.5472  8.549   8.6368  8.6822  9.0126  9.0778  9.237   9.2429  9.4427
  9.6744  9.7115 10.141  10.1216 10.2393 10.3881 10.5981 10.9565 11.072
 11.1468 11.3425 11.7212 11.7782 11.9857]
```

2 Models

Model 1: $y(t) = \alpha_0 + \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + \epsilon(t).$

Model 2: $y(t) = \dfrac{\alpha_0 + \alpha_1 t}{\beta_0 + \beta_1 t} + \epsilon(t).$

Model 3: $y(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \epsilon(t).$

where $\{\epsilon(t)\}$ is a sequence of i.i.d. normal random variables with mean zero and variance σ^2.

3 Least Square Estimators

To find the least squares estimators for the parameters in each model, I used an iterative optimization approach, specifically gradient descent. This method minimizes the sum of squared residuals by iteratively updating the parameter estimates in the direction that reduces the loss. Given the structure of the models and the availability of residuals, I implemented gradient descent to find parameter values that minimize the difference between the observed data and the model predictions. At each step, the parameters were updated using the gradient of the loss function with respect to the parameters.

For model 1, I choose the initial guess as [1,1,1,1,1] .Similarly for model 2 : [1,1,1,1] and for model 3: [1,1,1,1,1] . This uniform starting point was selected due to the lack of specific prior knowledge about the scale or expected values of the parameters. While it is a simple approach, this choice allowed for unbiased initial conditions, ensuring that no parameter was favored at the start. This choice may also help avoid any potential biases in the estimation process, though it could lead to slower convergence if parameters vary widely in scale.

So we get the Least Square Estimators are as follows:

- **Model 1:**

  $\alpha_0 = 2.036306748855132 \times 10^{-6}, \quad \alpha_1 = 2.506707491556617, \quad \alpha_2 = 1.0420625834796016,$
  $\beta_1 = 0.22886790568881882, \quad \beta_2 = 2.13699684715203$

- **Model 2:**

  $\alpha_0 = 9.87502340904815, \quad \alpha_1 = 31.882273923342535,$
  $\beta_0 = 3.993504806206553, \quad \beta_1 = 7.977560114217491 \times 10^{-27}$

- **Model 3:**

  $\beta_0 = 3.5191462229110155, \quad \beta_1 = 3.2589724454523243, \quad \beta_2 = 0.8531231390959737,$
  $\beta_3 = 3.235467008035791, \quad \beta_4 = 1.0936700978146765$

4 Best Model

To find which model is the "Best fitted" , we use a parameter called MSE(mean squared error) and these are as follows for given models :

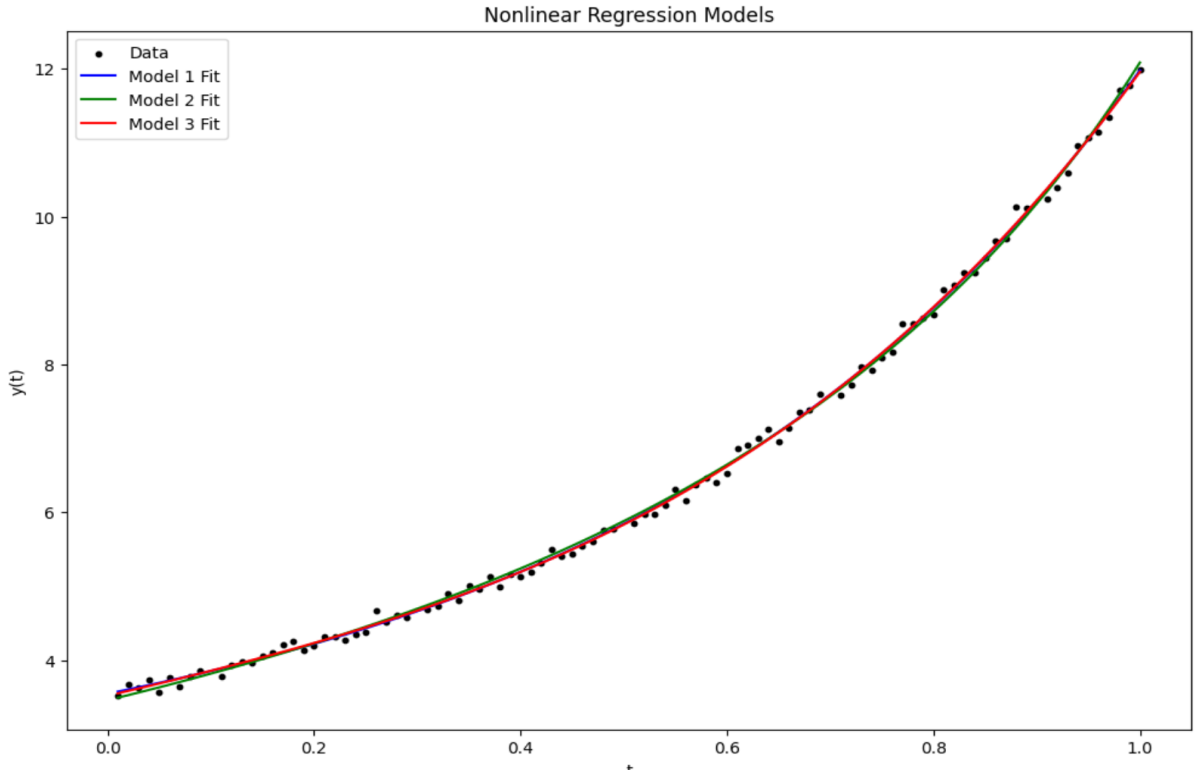
- **Model 1:** 0.0068191397062944475
- **Model 2:** 0.34725792558848967
- **Model 3:** 0.006749312276619548

As we get the minimum MSE for model 3, hence model 3 is the best fitted.

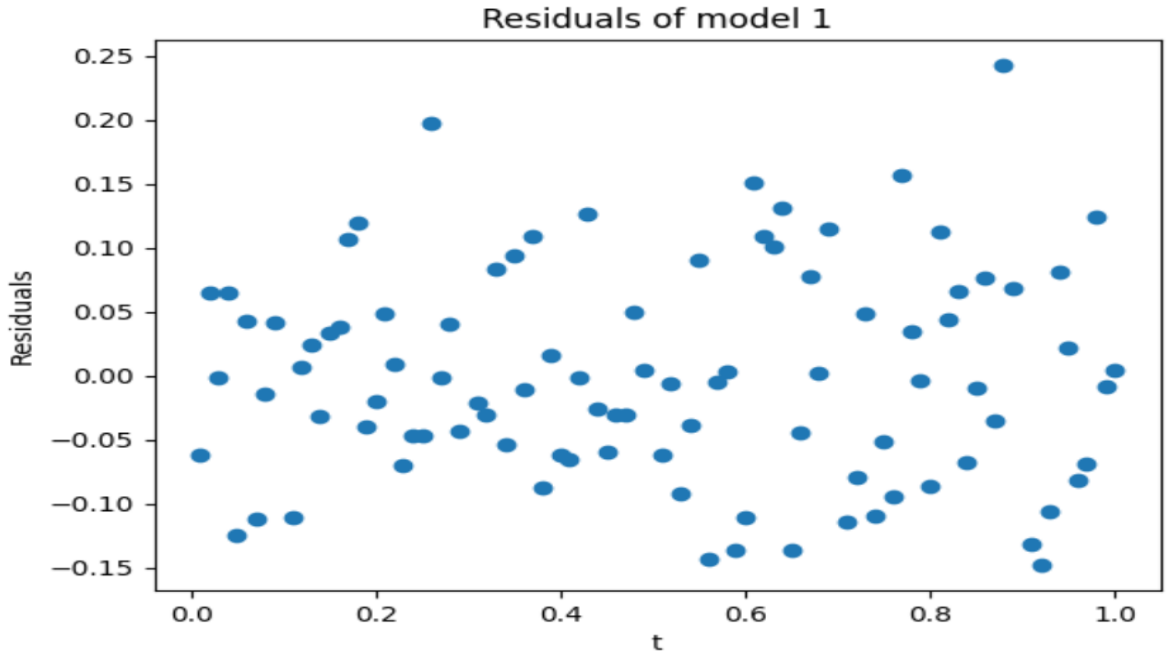
4.1 Estimate of σ^2

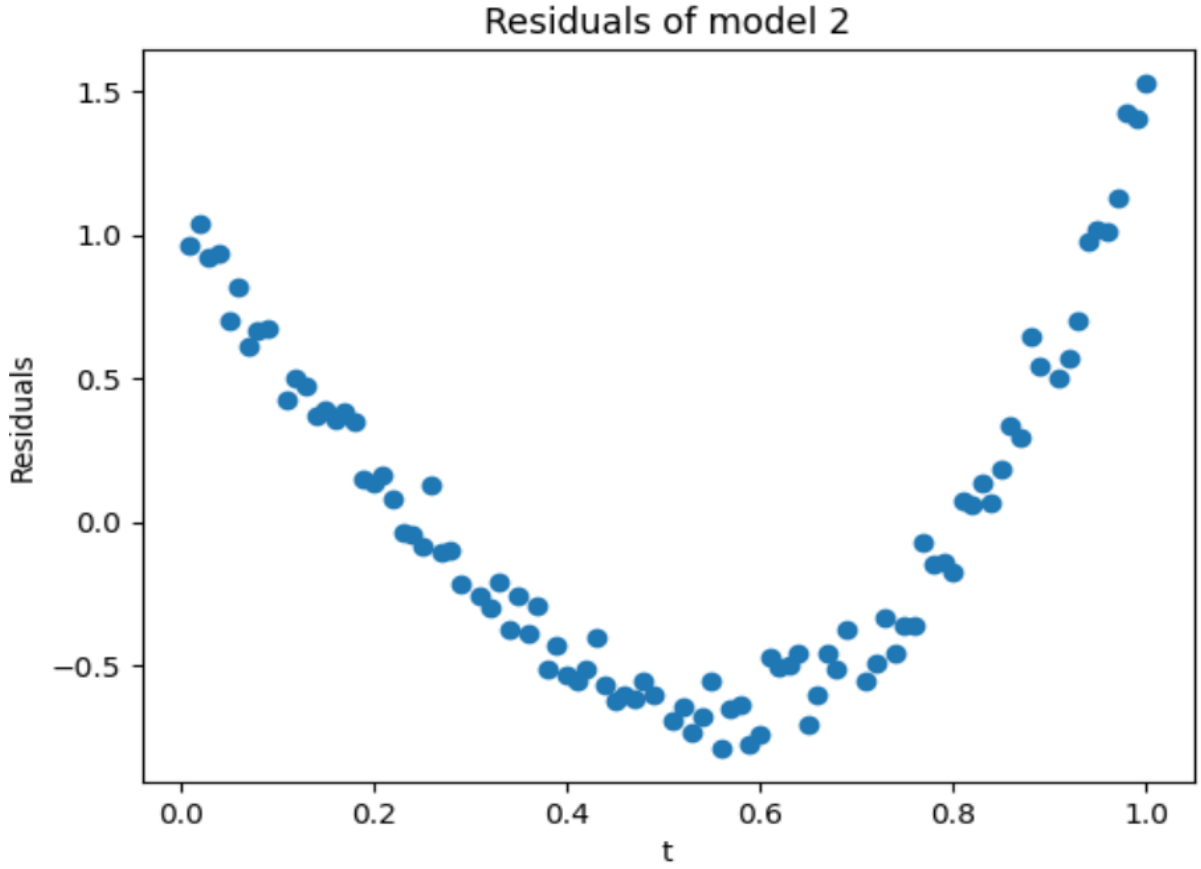
- **Model 1:** 0.007197980801088583
- **Model 2:** 0.36252201022974195
- **Model 3:** 0.007124274069765078

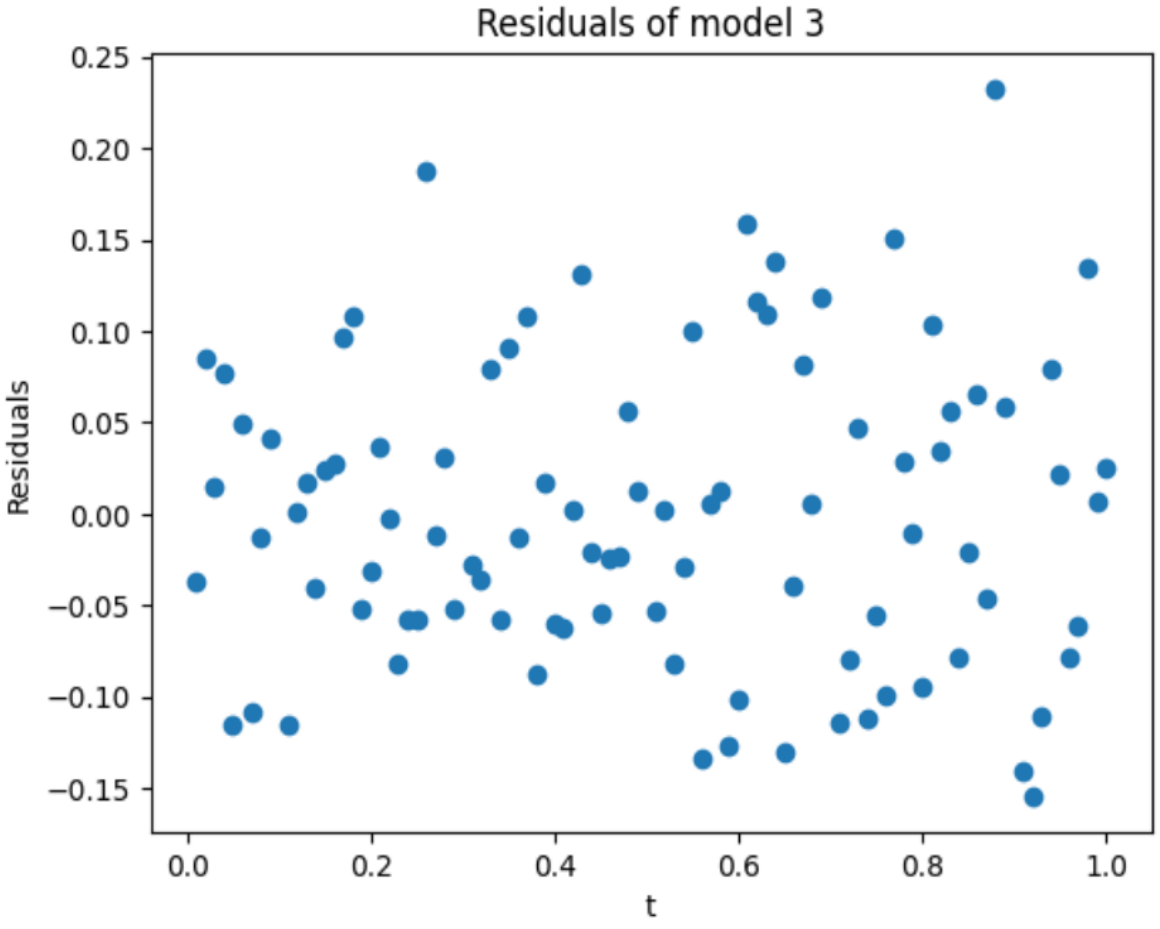
5 Fitted Curves



6 Residuals




Residuals of model 2
Residuals
t
1.5
1.0
0.5
0.0
−0.5
0.0
0.2
0.4
0.6
0.8
1.0


Residuals of model 3
Residuals
t
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
0.0
0.2
0.4
0.6
0.8
1.0

7 Confidence Interval

7.1 Model 1:

- α_0: Estimate = 0.0000, 95% Confidence Interval = [-4326.8505, 4326.8505]
- α_1: Estimate = 2.5061, 95% Confidence Interval = [-4236.3530, 4241.3651]
- α_2: Estimate = 1.0421, 95% Confidence Interval = [-86.8142, 88.8983]
- β_0: Estimate = 0.2289, 95% Confidence Interval = [-440.9498, 441.4075]
- β_1: Estimate = 2.1370, 95% Confidence Interval = [-38.2203, 42.4942]

7.2 Model 2:

- α_0: Estimate = 9.8750, 95% Confidence Interval = [-20580424.7477, 20580444.4978]
- α_1: Estimate = 31.8823, 95% Confidence Interval = [-66445487.1351, 66445550.8996]
- β_0: Estimate = 3.9935, 95% Confidence Interval = [-8322818.4225, 8322826.4096]
- β_1: Estimate = 0.0000, 95% Confidence Interval = [-1.3480, 1.3480]

7.3 Model 3:

- β_0: Estimate = 3.5191, 95% Confidence Interval = [2.4688, 4.5695]
- β_1: Estimate = 3.2590, 95% Confidence Interval = [-11.0197, 17.5376]
- β_2: Estimate = 0.8531, 95% Confidence Interval = [-56.3679, 58.0741]
- β_3: Estimate = 3.2355, 95% Confidence Interval = [-81.7660, 88.2369]
- β_4: Estimate = 1.0937, 95% Confidence Interval = [-40.6621, 42.8495]

8 Normality Distribution test

To assess whether the residuals of the model satisfy the normality assumption, we performed the Shapiro-Wilk test. The Shapiro-Wilk test is a widely used statistical test specifically designed to test the null hypothesis that a given sample comes from a normally distributed population. It is known for its power to detect deviations from normality, especially in small to moderately sized datasets.

The Shapiro-Wilk test is particularly suitable for testing normality because it takes into account both the skewness and kurtosis of the sample distribution. In cases where the residuals of a regression model are expected to follow a normal distribution, using this test helps in determining if the residuals significantly deviate from normality. A p-value lower than a predefined significance level (typically 0.05) would indicate that the residuals do not follow a normal distribution, which could suggest issues such as model mis-specification or the presence of outliers.

- **Model 1 :** Stat = 0.9812859738655647 , p-value = 0.19270903530098588
- **Model 2 :** 0.9235982445694861 , p-value = $3.4481699063757094 \times \mathbf{10^{-5}}$
- **Model 3 :** Stat = 0.9821745842046904 , p-value = 0.22352544157484422

Conclusion:

- Model 1 and Model 3 pass the normality distribution test (residuals are approximately normally distributed).
- Model 2 does not pass the normality distribution test (residuals significantly deviate from a normal distribution).