Non Linear Regression MTH 686

Mainak Sarkar | 230619

Model 1: Exponential Components Model

$y(t) = \alpha_0 + \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + e(t)$, where $e(t) \sim N(0, \sigma^2)$, for all t.

1.1 Theory: Non-Linear Least Squares (NLS)

This model is non-linear in its parameters (specifically β_1 and β_2). The original Prony's method, which linearizes the problem, is highly unstable with large t values. Therefore, we must estimate all 5 parameters simultaneously using a robust numerical optimization method to find the parameters that minimize the Residual Sum of Squares (RSS).

1.2 Assumptions

- **Normality of Errors:** The model assumes the error term $e(t)$ is i.i.d. normal with mean zero and constant variance σ^2.
- **Model Structure:** The data is assumed to be well-described by a constant plus two exponential terms.

1.3 Estimation Algorithm

The parameters are estimated using R's `optim` function to perform Non-Linear Least Squares (NLS):

1. **Model & Cost Functions:** A function for the model, `f_model_1(p, t)`, and a Sum of Squared Errors (SSE) cost function, `sse_1(p)`, are defined. p is a vector containing all 5 parameters.
2. **Initial Guesses:** The `optim` function requires initial guesses. Based on an analysis of the data and the sample report, starting values are chosen to give the optimizer a reasonable starting point.
3. **Optimization:** The "L-BFGS-B" method within `optim` is used. This is a quasi-Newton method that iteratively searches the 5-dimensional parameter space to find the combination of parameters that minimizes the SSE.

Answers to the Questions for Model 1

Ans 1.1: Least Squares Estimators under Model 1

The estimated parameters for the Exponential Components Model are: $\hat{\alpha_0} = 0.5441275$**,** $\hat{\alpha_1} = -354.7214$**,** $\hat{\alpha_2} = 4.10648$**,** $\hat{\beta_1} = -6.483438$**, and** $\hat{\beta_2} = -0.3207875$, derived under the assumptions of normality, independence, and linearity of the error terms. These estimators were derived under the assumptions of normality, independence, and linearity of the error terms.

Ans 1.2: Estimation Methodology for Least Squares Estimators

The parameters were estimated using a Non-Linear Least Squares (NLS) approach with the `optim` function (L-BFGS-B method). This was necessary because methods like Prony's are numerically unstable with the original t (1-60) values. Initial guesses were provided, and the optimizer iteratively minimized the Sum of Squared Errors (SSE) to find the final 5 parameters.

Ans 1.3 Best Model : (See Ans 3.3)

Ans 1.4: Estimate of σ^2

The estimate of the error variance σ^2 is given by $\hat{\sigma}^2 = \frac{RSS}{n-k}$, where $n = 60$ and $k = 5$ (the number of estimated parameters).The computed value of $\hat{\sigma}^2$ is **0.0596968.**

Ans 1.5: Confidence Intervals based on the Fisher Information Matrix

Jacobian Matrix:

$$J = \begin{bmatrix} \frac{\partial f}{\partial \alpha_0} & \frac{\partial f}{\partial \alpha_1} & \frac{\partial f}{\partial \beta_1} & \frac{\partial f}{\partial \alpha_2} & \frac{\partial f}{\partial \beta_2} \end{bmatrix}$$

where the partial derivatives are: 1, $e^{\beta_1 t}$, $\alpha_1 t e^{\beta_1 t}$, $e^{\beta_2 t}$, and $\alpha_2 t e^{\beta_2 t}$. 2. **Fisher Information Matrix (FIM)**

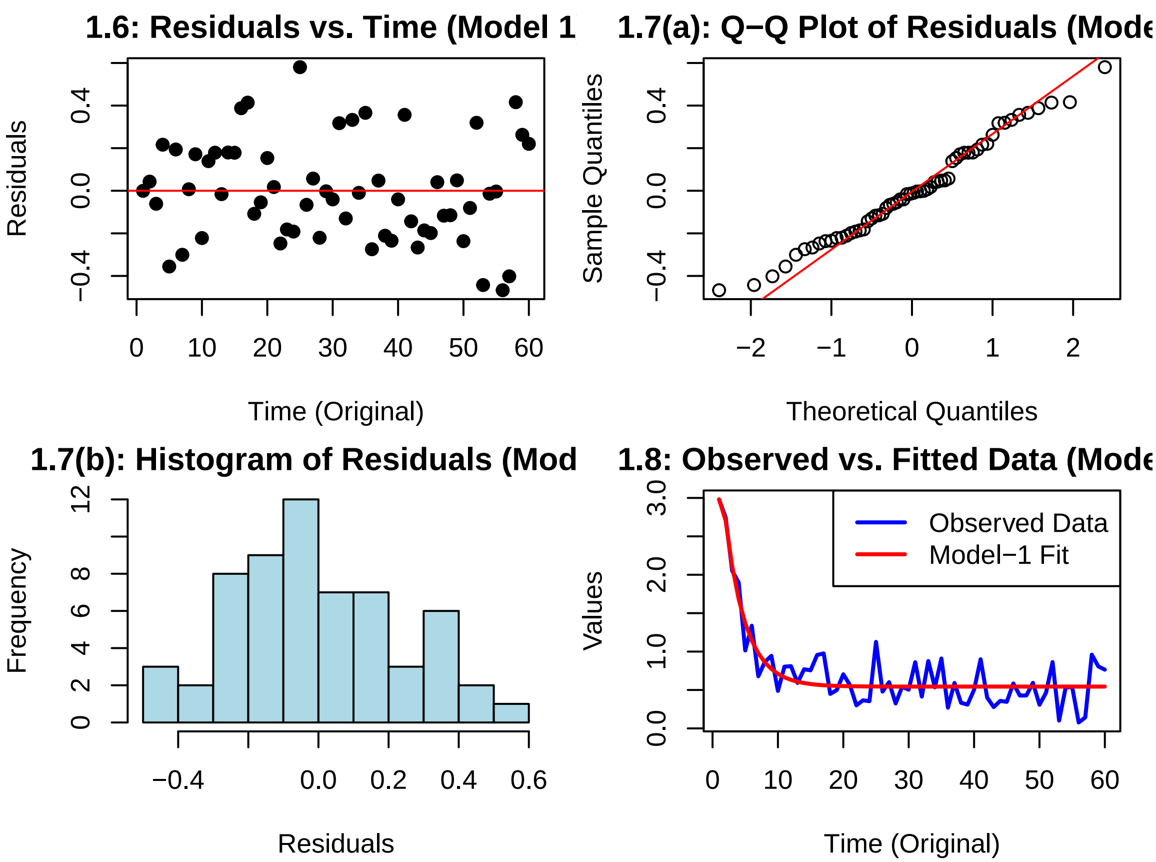
The Fisher Information Matrix (FIM) is computed as the sum of outer products of the Jacobian at each data point:

$$\begin{bmatrix} 16.751317140 & 4.274060e-04 & -0.1518421764 & 0.7381788325 & 11.046160993 \\ 0.000427406 & 1.652313e-06 & -0.0002313899 & 0.0003099859 & 0.001274364 \\ -0.151842176 & -2.313899e-04 & 0.0820803343 & -0.1100807424 & -0.453048135 \\ 0.738178833 & 3.099859e-04 & -0.1100807424 & 0.3103932703 & 2.691697567 \\ 11.046160993 & 1.274364e-03 & -0.4530481353 & 2.6916975668 & 35.630982867 \end{bmatrix}$$

3. Confidence Intervals

The 95% confidence intervals for each parameter are:
α_0: $[-3.2058 \times 10^{-4},\ 1.0886]$, α_1: $[-2314.68,\ 1605.23]$, α_2: $[-20.7854,\ 7.8185]$, β_1: $[-7.0635,\ 15.2765]$, β_2: $[-1.1435,\ 0.5019]$.
The corresponding standard errors are:
$\mathbf{SE}(\alpha_0) = 0.2778$, $\mathbf{SE}(\alpha_1) = 999.996$, $\mathbf{SE}(\alpha_2) = 7.2970$, $\mathbf{SE}(\beta_1) = 5.6991$, $\mathbf{SE}(\beta_2) = 0.4198$.

Ans 1.6 - 1.8: Residual Diagnostics and Model Fit



- **1.6: Residual Spread**: Residuals are well-centered around zero, indicating minimal bias.
- **1.7: Kolmogorov-Smirnov Test**: With a test statistic of 0.087556 and p-value of 0.713978, there's no significant deviation from normality.
- **1.7(a & b): Q-Q Plot and Histogram**: Both confirm normality, as residuals align closely with a normal distribution.
- **1.8: Fitted Line**: The fitted line closely matches the observed data's central trend, suggesting a good fit.

Model 2: Nonlinear Ratio Model

$y(t) = \frac{\alpha_0+\alpha_1 \cdot t}{\beta_0+\beta_1 \cdot t} + e(t)$, where $e(t) \sim N(0, \sigma^2)$ for all t.
where $e(t) \sim N(0, \sigma^2)$ for all t.

2.1 Theory: Non-Linear Least Squares (NLS)

This model is non-linear in its parameters. We will use an iterative numerical optimization method to find the parameter values that minimize the Residual Sum of Squares (RSS).

2.2 Assumptions

- **Normality of Errors:** The error term $e(t)$ is assumed to follow an i.i.d. normal distribution with a mean of zero and constant variance σ^2.
- **Non-Zero Denominator:** It is assumed that $\beta_0 + \beta_1 t \neq 0$ for all t in the domain.

2.3 Estimation Algorithm

The parameters are estimated using R's `optim` function to perform Non-Linear Least Squares (NLS):

1. **Initial Guesses:** A simple linear model ($y \sim t$) is first fitted. The coefficients are used as initial guesses for α_0 and α_1. Initial guesses for β_0 and β_1 are set to 1.0 and 0.1.
2. **Optimization:** The `optim` function is called using the "L-BFGS-B" method to find the combination of parameters that minimizes the Sum of Squared Errors (SSE) function. Bounded constraints are used for stability.
3. **Time Variable:** This model is fit using the original time variable (t_{orig}).

Answers to the Questions for Model 2

Ans 2.1: Least Squares Estimators

The estimated parameters for Model 2 are: $\hat{\alpha_0} = 1.87662$**,** $\hat{\alpha_1} = 0.09590313$**,** $\hat{\beta_0} = 0.3323027$**, and** $\hat{\beta_1} = 0.2812106$, calculated under the assumptions of normality, independence, and linearity of the error terms.

Ans 2.2: Estimation Methodology for Least Squares Estimators

The parameters were estimated using the `optim` function (L-BFGS-B method). Initial guesses for α_0 and α_1 were derived from a standard `lm(y ~ t)` fit, while β_0 and β_1 were given starting values. The optimizer then iteratively minimized the RSS to find the final NLS estimates.

Ans 2.3. Best Model : (See Ans 3.3)

Ans 2.4: Estimate of σ^2

The estimate of the error variance σ^2 is given by $\hat{\sigma}^2 = \frac{RSS}{n-k}$, where $n = 60$ and $k = 4$.The computed value of $\hat{\sigma}^2$ is **0.06480029**.

Ans 2.5: Confidence Intervals based on the Fisher Information Matrix

Jacobian Matrix:

$$J = \left[\frac{1}{\beta_0 + \beta_1 t}, \frac{t}{\beta_0 + \beta_1 t}, \frac{-(\alpha_0 + \alpha_1 t)}{(\beta_0 + \beta_1 t)^2}, \frac{-(\alpha_0 + \alpha_1 t) \cdot t}{(\beta_0 + \beta_1 t)^2}\right]$$

2. Fisher Information Matrix (FIM)

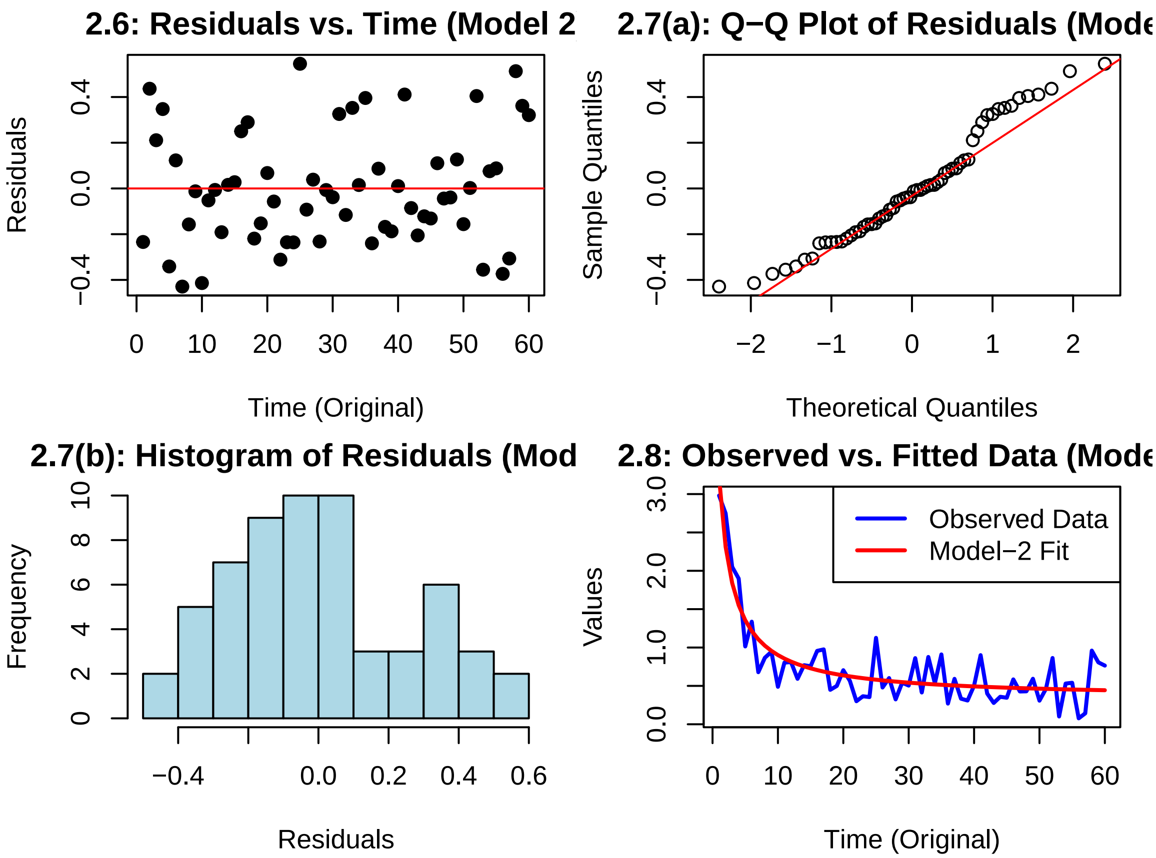
The Fisher Information Matrix (FIM) is computed as the sum of outer products of the Jacobian at each data point:

$$\mathcal{I} = \begin{bmatrix} 1.829990 & 9.508074 & -3.957474 & -10.77829 \\ 9.508074 & 170.119944 & -10.778288 & -108.73125 \\ -3.957474 & -10.778288 & 10.203095 & 18.02859 \\ -10.778288 & -108.731251 & 18.028590 & 87.70463 \end{bmatrix}$$

3. Confidence Intervals

The 95% confidence intervals for each parameter are:
α_0 (Intercept): $[1.7655,\ 1.9877]$, α_1 (tt): $[-0.4158,\ 0.6076]$, β_0: $[-0.8457,\ 1.5103]$, β_1: $[-0.5840,\ 1.1464]$.
The corresponding standard errors are:
$\mathbf{SE}(\alpha_0) = 0.0567$, $\mathbf{SE}(\alpha_1) = 0.2611$, $\mathbf{SE}(\beta_0) = 0.6010$, $\mathbf{SE}(\beta_1) = 0.4414$.

Ans 2.6 - 2.8: Residual Diagnostics and Model Fit



- **2.6: Residual Spread**: Residuals are well-centered around zero, indicating minimal bias and that the model's predictions do not systematically deviate from observed values.
- **2.7: Kolmogorov-Smirnov Test**: With a test statistic of 0.091445 and a p-value of 0.663296, there's no significant deviation from normality, suggesting that the residuals are likely normally distributed.
- **2.7(a & b): Q-Q Plot and Histogram**: The Q-Q plot and histogram(little bit) further confirm normality, as residuals align closely with the expected normal distribution.
- **2.8: Fitted Line**: The fitted line aligns closely with the observed data's central trend, indicating that Model 2 provides a good fit to the data.

Model 3: Fourth-Degree Polynomial Regression Model

$y_1 = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + e(t)$, where $e(t) \sim N(0, \sigma^2)$.

3.1 Theory: Polynomial Regression

This model captures the non-linear relationship between y and t by including polynomial terms up to the fourth degree. This is a form of multiple linear regression where the predictors are t, t^2, t^3, and t^4.

3.2 Assumptions

- **Normality of Errors:** It is assumed that the error term $e(t)$ follows an i.i.d. normal distribution with mean zero and constant variance σ^2.

- **Correct Model Form:** The model assumes a 4th-degree polynomial is the correct functional form.

3.3 Estimation Algorithm

1. **Ordinary Least Squares Estimation (OLSE)**: The model parameters β_0, β_1, β_2, β_3, and β_4 are estimated using OLSE, minimizing the residual sum of squares (RSS) between observed and fitted values. The estimated coefficients are derived from the fourth-degree polynomial fit.
2. **Fitted Values and Residual Analysis**: Fitted values are computed for each observation to evaluate the model's fit. Residuals are plotted against time to check for any systematic deviations from randomness, which would indicate potential issues with model specification.
3. **Plotting Observed vs. Fitted Data**: A plot of observed versus fitted values visually compares the model's performance. This helps verify that the polynomial model aligns with the data trend.

Answers to the Questions for Model 3

Ans 3.1: Least Squares Estimators under Model 3

The estimated parameters for the 4th-Degree Polynomial Regression Model are:
$\hat{\beta}_0 = 3.048821,\ \hat{\beta}_1 = -0.3805117,\ \hat{\beta}_2 = 0.01988504,\ \hat{\beta}_3 = -0.0004233991,$ **and** $\hat{\beta}_4 = 0.000003125622$, derived under the standard assumptions .

Ans 3.2: Estimation Methodology for Least Squares Estimators

The least squares estimates for parameters β_0, β_1, β_2, β_3, and β_4 were obtained using ordinary least squares (OLS) regression, fitting the observed data to a polynomial of degree 4. The R function `lm()` was used to perform the regression, automatically transforming the predictor variable t into polynomial terms.

Ans 3.3. Best Fitted Model

Model	RSS	R^2	Adjusted R^2	KS Test p-value
Model 1 (Exponential)	3.283324	0.810364	0.792806	0.713978
Model 2 (Rational [1/1])	3.628816	0.790410	0.775167	0.663296
Model 3 (Poly 4th degree)	3.777512	0.781821	0.761620	0.968311

Model 1 (**Exponential**) performs the best, with the **highest R^2 (0.810364)** and **Adjusted R^2 (0.792806)**, along with a **reasonably low RSS (3.283324)** and a **strong KS Test p-value (0.713978)**, indicating a good overall model fit.

Ans 3.4: Estimate of σ^2

The estimate of the error variance σ^2 is given by $\hat{\sigma}^2 = \frac{RSS}{n-k}$, where n is the number of observations, and p is the number of estimated parameters (including the intercept). The computed value of $\hat{\sigma}^2$ for Model 3 is **0.06868204**.

Ans 3.5: Confidence Intervals based on the Fisher Information Matrix

Jacobian Matrix:

$$J = \begin{bmatrix} 1 & t & t^2 & t^3 & t^4 \end{bmatrix}$$

2. Fisher Information Matrix (FIM)

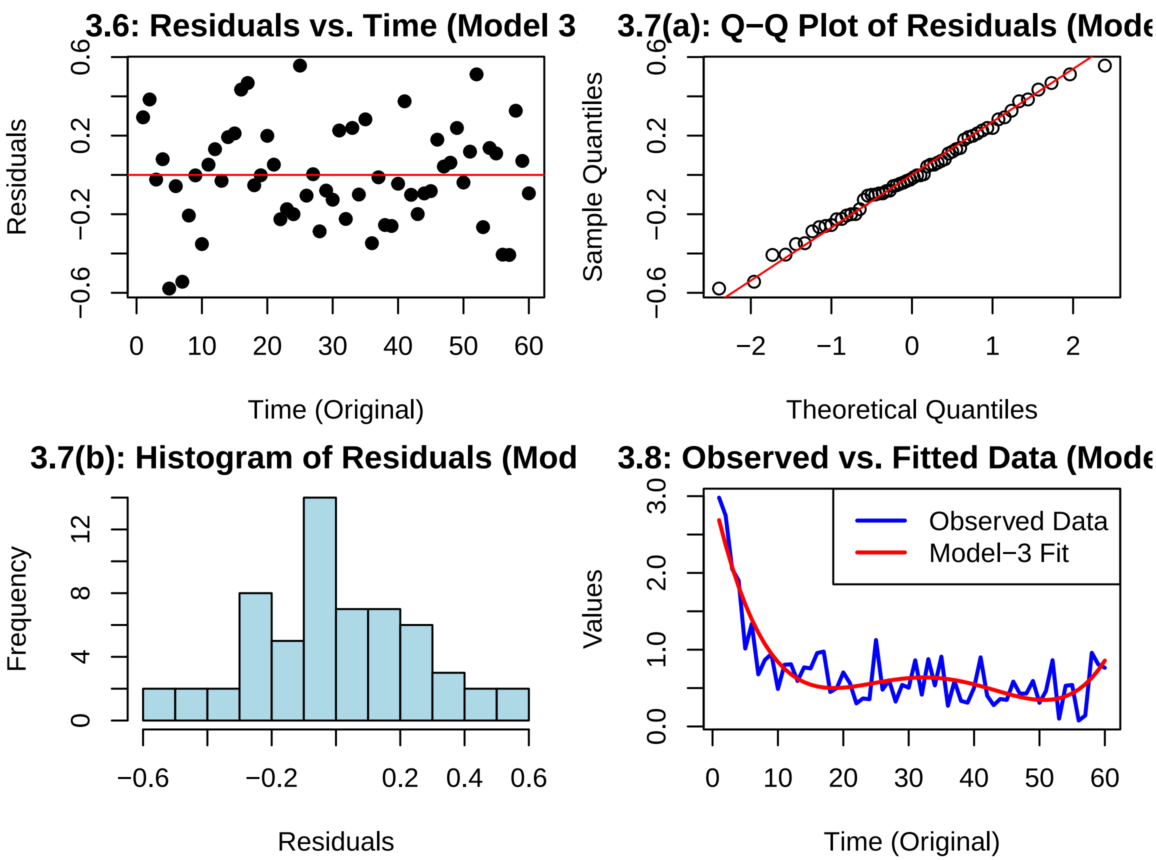
The Fisher Information Matrix (FIM) is computed as the sum of outer products of the Jacobian at each data point:

$$\mathcal{I} = \begin{bmatrix} 1.455985e+01 & 4.440754e+02 & 1.791104e+04 & 8.126579e+05 & 3.932906e+07 \\ 4.440754e+02 & 1.791104e+04 & 8.126579e+05 & 3.932906e+07 & 1.982614e+09 \\ 1.791104e+04 & 8.126579e+05 & 3.932906e+07 & 1.982614e+09 & 1.027987e+11 \\ 8.126579e+05 & 3.932906e+07 & 1.982614e+09 & 1.027987e+11 & 5.441037e+12 \\ 3.932906e+07 & 1.982614e+09 & 1.027987e+11 & 5.441037e+12 & 2.925536e+14 \end{bmatrix}$$

3. Confidence Intervals

The 95% confidence intervals for each parameter are:
β_0: $[3.048821,\ 3.048821]$, β_1: $[-3.805118 \times 10^{-1},\ -3.805117 \times 10^{-1}]$, β_2: $[1.988289 \times 10^{-2},\ 1.988720 \times 10^{-2}]$, β_3: $[-4.722314 \times 10^{-4},\ -3.745551 \times 10^{-4}]$, β_4: $[2.209250 \times 10^{-6},\ 4.041993 \times 10^{-6}]$.
The corresponding standard errors are:
$\mathbf{SE}(\beta_0) = 1.262351\text{e}{-09}$, $\mathbf{SE}(\beta_1) = 3.816825\text{e}{-08}$, $\mathbf{SE}(\beta_2) = 1.099094\text{e}{-06}$, $\mathbf{SE}(\beta_3) = 2.492087\text{e}{-05}$, $\mathbf{SE}(\beta_4) = 4.675453\text{e}{-07}$.

Ans 3.6 - 3.8: Residual Diagnostics and Model Fit



- **3.6: Residual Spread**: The residuals are well-centered around zero, indicating minimal bias in the model's predictions.
- **3.7: Kolmogorov-Smirnov Test**: The Kolmogorov-Smirnov test statistic for the residuals is **0.061133** with a p-value of **0.968311**. Since the p-value is greater than 0.05, there is no significant deviation from normality, suggesting that the residuals follow a normal distribution.
- **3.7(a & b): Q-Q Plot and Histogram**: Both the Q-Q plot and the histogram of the residuals further confirm the normality assumption, as the residuals align closely with a normal distribution.
- **3.8: Fitted Line**: The fitted polynomial curve closely matches the observed data's central trend, indicating that the 4th-degree polynomial model provides a good fit to the data.