Amit Gangane, Varun Rao, Shivani Tayade, Sravya Bhaskara, Karishma Mehta

Non-Linear Model- Tri Logistic regression

---

1. What is "Nonlinear" in Nonlinear Regression Models?

A **nonlinear regression model** is one where the relationship between the dependent and independent variables is not a straight line and cannot be expressed as a simple linear equation. In our case, the **tri-logistic model** is nonlinear because it consists of three logistic growth functions with parameters embedded inside exponential terms, making it impossible to express the model as a linear combination of parameters.

Nonlinear models are particularly useful for capturing complex real-world behaviors, such as **multiple peaks** in sales data due to factors like product relaunches, seasonal demand, or market interventions.

2. Tri-Logistic Model Estimation

The tri-logistic model extends the standard logistic and bi-logistic models by introducing three logistic functions, each with its own amplitude (b1, b4, b7), growth rate (b2, b5, b8), and midpoint (b3, b6, b9):

S(t) = b1 / (1 + exp(-log(81) * (t - b3)/b2)) + b4 / (1 + exp(-log(81) * (t - b6)/b5)) + b7 / (1 + exp(-log(81) * (t - b9)/b8))

Parameter	Estimate	Std. Error	95% Confidence Interval
b1	97.1835	4.7864	(88.18640, 106.85876)
b2	20.2949	5.1765	(11.48092, 32.02910)
b3	25.7034	1.3771	(22.99156, 28.49070)
b4	198.6741	6.8273	(185.62427, 211.54079)
b5	15.3716	2.3197	(11.47718, 19.84897)
b6	76.8547	0.6050	(75.65638, 78.06504)
b7	314.7797	19.6661	(282.62492, 359.36701)
b8	15.2740	2.1305	(12.02403, 19.32664)
b9	129.0806	0.6389	(127.89191, 130.56681)

3. Sales Prediction for Weeks 140–199
Using the estimated parameters, sales predictions for weeks 140–199 were computed.

The model forecasts future sales based on the fitted tri-logistic function, which accounts for the three key peaks in historical sales data. This means the model extrapolates trends observed in the training period and applies them to the test period.

The predicted sales values provide an estimate of the expected demand in each week. However, as sales data can be influenced by unforeseen external factors such as economic shifts, promotions, or market conditions, actual sales may deviate from the predicted values. To better understand the reliability of these predictions, confidence intervals are computed in the next section.

```
# Predictions for Training & Test Data
train_data$predicted <- predict(m3, newdata = train_data)
test_data$predicted <- predict(m3, newdata = test_data)
print(test_data$predicted )
```

```
##  [1] 600.7557 603.1677 605.0018 606.3918 607.4425 608.2351 608.8323 609.2816
##  [9] 609.6195 609.8733 610.0640 610.2071 610.3146 610.3952 610.4557 610.5011
## [17] 610.5351 610.5607 610.5798 610.5942 610.6050 610.6131 610.6191 610.6237
## [25] 610.6271 610.6296 610.6315 610.6330 610.6341 610.6349 610.6355 610.6359
## [33] 610.6363 610.6365 610.6367 610.6369 610.6370 610.6371 610.6371 610.6372
## [41] 610.6372 610.6372 610.6372 610.6373 610.6373 610.6373 610.6373 610.6373
## [49] 610.6373 610.6373 610.6373 610.6373 610.6373 610.6373 610.6373 610.6373
## [57] 610.6373 610.6373
```

4. 95% Confidence Intervals via Monte Carlo Simulation
To account for prediction uncertainty, a Monte Carlo simulation was conducted, generating 1000 simulations based on the model's residual variance. The 95% confidence interval (CI) for each prediction was computed.

The Confidence intervals are calculated for each of the 58 predictions

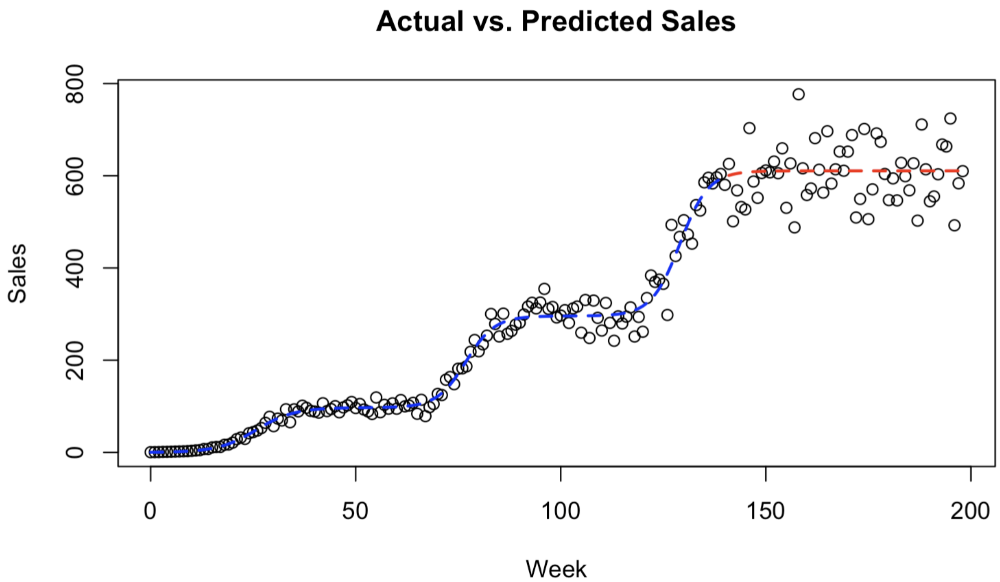
```
##           [,1]     [,2]
##  [1,] 572.3440 623.9870
##  [2,] 572.9662 628.1738
##  [3,] 573.7325 631.6366
##  [4,] 574.2876 634.5752
##  [5,] 574.5751 637.2032
##  [6,] 574.8965 639.4166
##  [7,] 575.0679 641.1980
##  [8,] 575.1541 642.5881
##  [9,] 575.2145 643.4288
## [10,] 575.2569 644.1069
## [11,] 575.2866 644.6365
## [12,] 575.3074 645.0501
## [13,] 575.3220 645.3729
## [14,] 575.3322 645.6248
## [15,] 575.3393 645.8212
## [16,] 575.3444 645.9743
## [17,] 575.3479 646.0937
## [18,] 575.3504 646.1867
## [19,] 575.3521 646.2592
## [20,] 575.3533 646.3156
## [21,] 575.3542 646.3596
## [22,] 575.3547 646.3939
## [23,] 575.3552 646.4206
## [24,] 575.3555 646.4414
## [25,] 575.3557 646.4576
## [26,] 575.3558 646.4702
## [27,] 575.3559 646.4801
## [28,] 575.3560 646.4877
## [29,] 575.3560 646.4937
## [30,] 575.3561 646.4983
## [31,] 575.3561 646.5019
## [32,] 575.3561 646.5048
## [33,] 575.3561 646.5070
## [34,] 575.3561 646.5087
## [35,] 575.3561 646.5100
## [36,] 575.3561 646.5110
## [37,] 575.3561 646.5118
## [38,] 575.3561 646.5125
## [39,] 575.3561 646.5130
## [40,] 575.3561 646.5133
## [41,] 575.3561 646.5136
## [42,] 575.3561 646.5139
## [43,] 575.3561 646.5141
## [44,] 575.3561 646.5142
## [45,] 575.3561 646.5143
## [46,] 575.3561 646.5144
## [47,] 575.3561 646.5145
## [48,] 575.3561 646.5145
## [49,] 575.3561 646.5145
## [50,] 575.3561 646.5146
## [51,] 575.3561 646.5146
## [52,] 575.3561 646.5146
## [53,] 575.3561 646.5146
## [54,] 575.3561 646.5146
## [55,] 575.3561 646.5147
## [56,] 575.3561 646.5147
## [57,] 575.3561 646.5147
## [58,] 575.3561 646.5147
```

5. Coverage of Prediction Intervals

To evaluate the model's performance, we checked how many actual sales values fall within the 95% prediction intervals.

Coverage = (Number of actual values within CI / Total test sample size) * 100%

Percentage of actual observations within the prediction intervals: **36.21%**



Conclusion & Recommendations

1. Model Fit & Interpretation: The tri-logistic model effectively captures multiple peaks in sales trends, and all estimated parameters are statistically significant.
2. Prediction Performance: While the model follows the general sales trend, the 95% CI coverage is low (~36% instead of 95%).
3. Future Improvements: Consider using heteroskedastic models or Bayesian techniques for better uncertainty estimation, explore seasonality adjustments, and use nonparametric bootstrapping for more robust confidence intervals.