
The GenIQ-enhanced Regression Model

Bruce Ratner PhD
DM STAT-1 CONSULTING
1 800 DM STAT-1 www.DMSTAT1.com

GenIQ[®]

The GenIQ Model

- The GenIQ Model is a machine learning alternative model to the statistical ordinary least squares, logistic regression models.
- GenIQ let's the data define the model – automatically
 1. **Data mines** for new variables
 2. Performs **variable selection**
 3. **Specifies the model** (equation) - so as to "**optimize the decile table**," to fill the upper deciles with as much profit/many responses as possible.
- GenIQs optimizing of the deciles **is equivalent to** predicting the best possible ranking (permitted by the data) of the target variable based on the GenIQ Model score **GenIQvar**.

Objective: The GenIQ-enhancer

- The GenIQ Model can be used on a final regression model to let GenIQs data mining prowess **enhance**, data permitting, the results of the final model.
- GenIQ let's the data define the model – automatically
 1. Data mines for new variables,
 2. Performs variable selection, and then
 3. Specifies the model equation - so as to "optimize the decile table," to fill the upper deciles with as much profit/many responses as possible.
- GenIQs Utility: **Enhance** a final regression model by running GenIQ with only one predictor: the final regression equation score.

Build a Logistic Regression Model

- **OBJECTIVE #1:** Build a binary response model with data in Table 1.
I build a logistic regression model (LRM) for predicting the target variable Response using predictors XX1 and XX2.
- The LRM equation is:
 - Logit of Response (=Yes) = $0.1978 - 0.0328 * XX1 + 0.0308 * XX2$

Table 1.

ID	XX1	XX2	Response
1	31	38	Yes
2	12	30	No
3	35	21	Yes
4	23	30	No
5	45	37	No
6	16	13	No
7	45	5	Yes
8	30	30	Yes
9	6	10	Yes
10	30	10	No

Build a Logistic Regression Model

LRM RESULTS: *The Response ranking is not perfect.*

The ranking of Response based on the LRM probability score

Prob_of_Response, which is obtained directly from the LRM logit score, is below in Table 2. Clearly, the Yes-No ordering is not perfect; amusingly, it seems random!

Table 2.

ID	XX1	XX2	Response	Prob_of_Response
7	45	5	Yes	0.75472
10	30	10	No	0.61728
3	35	21	Yes	0.57522
5	45	37	No	0.53452
6	16	13	No	0.48164
8	30	30	Yes	0.46556
9	6	10	Yes	0.42336
1	31	38	Yes	0.41299
4	23	30	No	0.40913
2	12	30	No	0.32557

GenIQ Model: Its Output

■ **Unsuspected Equation – GenIQs Output**

1. The visual display called a parse tree, depicts the GenIQ Model
2. The GenIQ Model "equation" is actually **computer program/code**

■ The regression modeler, anticipating an equation:

- ▶ $Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n$, is stupefied when s/he unsuspectedly sees the GenIQ computer code.

■ **Ungainly Interpretation – "*Bit much to grasp.*"**

- ▶ The GenIQ parse tree and computer code can be a "bit much to grasp." The visual display provides the modeler with an ocular sense, albeit mote, of comfort and confidence for understanding and using GenIQ.

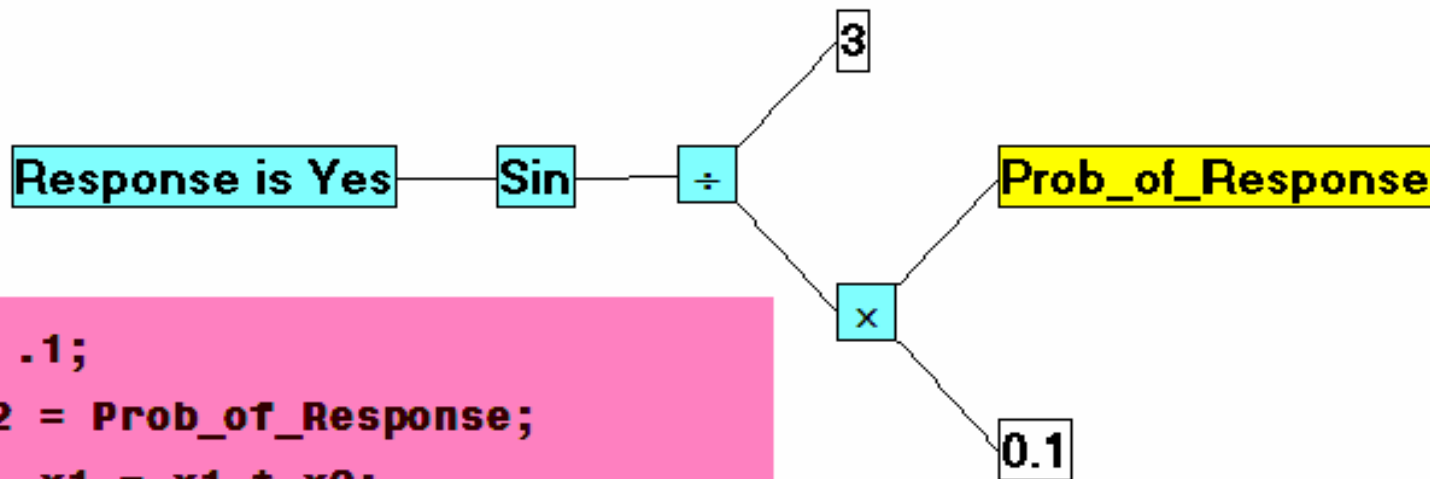
GenIQ Model: Its Output

- **Ungainly Interpretation** – *"Bit much to grasp."*
 - **The GenIQ tree**
 - ▶ although not a "black box" like most other machine learning methods
 - **gives the modeler a graphic, albeit Picasso-like, to make sense of.**
- **GenIQ for the everyday regression model produces a GenIQ tree defined by**
 - "branches" formed by **yellow-boxed** predictor variables attached to various numbers of **turquoise-boxed** functions, as displayed in next slide.

GenIQ Model: Its Output

■ Running GenIQ with the final LRM score:

- Logit of Response (=Yes) = $0.1978 - 0.0328 * XX1 + 0.0308 * XX2$



```
x1 = .1;  
x2 = Prob_of_Response;  
x1 = x1 * x2;  
x2 = 3;  
If x1 NE 0 Then x1 = x2 / x1; Else x1 = 1;  
x1 = Sin(x1);  
GenIQvar = x1;
```


GenIQ-enhanced LRM

The GenIQ-enhanced LRM Results: *The Response ranking is **perfect!***

The ranking of Response - based on the GenIQ Model with only the LRM score **Prob_of_Response** - produces a GenIQ Model score **GenIQvar**, in Table 3, yielding a perfect Yes-No ordering.

Table 3.

ID	XX1	XX2	Prob_of_Response	Response	GenIQvar
8	30	30	0.46556	Yes	0.99934
9	6	10	0.42336	Yes	0.98444
3	35	21	0.57522	Yes	0.95002
7	45	5	0.75472	Yes	0.88701
1	31	38	0.41299	Yes	-0.37436
5	45	37	0.53452	No	-0.41121
6	16	13	0.48164	No	-0.51791
2	12	30	0.32557	No	-0.86183
4	23	30	0.40913	No	-0.87665
10	30	10	0.61728	No	-0.99553

Thank you

Danke

Xie xie

Khawp khun

Yum botie

Mahalo

Salamat

Juspa

raña

Spacibo

Obrigada

Arigato