The Falling Child Project : Binary Logistic Regression

"Ari" Arivazhagan, PMP, PMI-SP,PMI-RMP,MBB

Founder-CEO at International Institute of Project Management (IIPM)

The Falling Child : This is part-2 of the series of articles on Borewell accidents in India. Part-1 of the series *predicted the probabilities of innocent children falling into open borewells,* using G-chart, sunk by farmers for agricultural and drinking water, while playing in the fields.

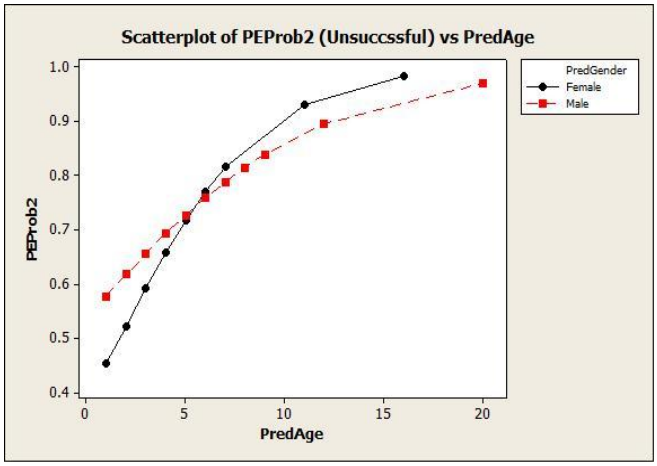
Binary Logistic Regression: This case study uses the power of predictive analytics to predict the probability of successfully rescuing a trapped child based on the inputs if it was a girl or a boy and the age of the child using Binary Logistic Regression.

The borewell accidents data collected and provided by The Falling Child Project (www.fallingchild.org), an NGO based out of USA, has been used for this predictive analysis. A part of the raw data provided by the NGO is shown Table-1. A total of 62 borewell accident cases in India have been documented from 2001 to January 2015.

	A	B	C	E	F	H	K	M	O
1	Year	State	District	Country	Date (mm/dd/yy	Victim_Age	Rescue_Outcon	Victim Gender	Borewell Depth,m
2	2015	Gujarat	Ahmedabad	India	1/15/2015	3	Successful	Female	45
3	2015	Telangana	Rangareddy	India	1/14/2015	5	Successful	Female	Unknown
4	2014	Maharashtra	Aurangabad	India	12/26/2014	2	Successful	Male	20
5	2014	Madhya Prades	Chhatarpur	India	12/18/2014	1	Unsuccessful	Male	100
6	2014	Gujarat	Ahmedabad	India	11/23/2014	3	Successful	Female	45
7	2014	Telangana	Ranga Reddy	India	10/12/2014	4	Unsuccessful	Female	300
8	2014	Karnataka	Bagalkot	India	8/4/2014	6	Unsuccessful	Male	300
9	2014	Andhra Prades	Nellore	India	8/4/2014	4	Unsuccessful	Female	35
10	2014	Madhya Prades	Shajapur	India	7/7/2014	8	Unsuccessful	Male	100
11	2014	Uttar Pradesh	Mainpuri	India	7/5/2014	60	Unsuccessful	Male	45
12	2014	Chhattisgarh	Rajnandgaon	India	6/28/2014	5	Unsuccessful	Female	40
13	2014	Gujarat	Jamnagar	India	6/21/2014	4	Unsuccessful	Male	350
14	2014	Karnataka	Bijapur	India	6/20/2014	4	Unsuccessful	Female	300
15	2014	Guwahati	Assam	India	5/4/2014	2	Unsuccessful	Male	30
16	2014	Gujarat	Bhavnagar	India	5/3/2014	3	Unsuccessful	Female	35
17	2014	Tamil Nadu	Viluppuram	India	4/6/2014	3	Unsuccessful	Female	30
18	2014	Tamil Nadu	Tiruvannamala	India	4/16/2014	1	Unsuccessful	Male	47
19	2014	Tamil Nadu	Tirunelveli	India	4/15/2014	3	Successful	Male	15
20	2014	Rajasthan	Churu	India	1/9/2014	2	Unsuccessful	Male	200
21	2013	Punjab	Chandigarh	India	9/4/2013	3	Successful	Male	Unknown
22	2013	Haryana	Palwal	India	5/30/2013	4	Successful	Female	200

Borewell Hole Accidents List

The predicted probabilities for unsuccessful events verses the Predicted Age and Predicted Gender are shown in Scatter Plot (Figure-1) below.



We can predict, with 70% confidence, that the probability of unsuccessful rescue is more by 15 % for a male child of age 2 than that for a female child of same age. However, it is surprising to note that , above age 5, girls have about 10 % more chances of being unsuccessfully rescued from the borewell pits than do the boys.

Binary Logistic Regression: Rescue versus Age, Gender

```
Link Function: Logit
Response Information
Variable  Value          Count
Rescue    Unsuccessful      43  (Event)
          Successful        19
          Total             62
Logistic Regression Table
                                                  Odds      95% CI
Predictor        Coef   SE Coef      Z      P  Ratio  Lower  Upper
Constant    -0.469872   1.20398  -0.39  0.696
Age          0.279562  0.310339   0.90  0.368   1.32   0.72   2.43
Gender
 Male        0.614843   1.37211   0.45  0.654   1.85   0.13  27.22
Gender*Age
 Male       -0.112769  0.334288  -0.34  0.736   0.89   0.46   1.72

Log-Likelihood = -36.336
Test that all slopes are zero: G = 3.741, DF = 3, P-Value = 0.291
```

It may be noted that one outlier data, a male of age 60, was replaced with a male of age 6 to reduce the unnecessary effect of outlier on the whole analysis / output.

Inferences : From the Binary Logistic Regression analysis above, one can predict that boys of age 5 and above have more chances of being successfully rescued than do the girls. Although the analysis indicates a P-value of 0.736,hinting that there is not much of interaction between the age of the child and its gender in predicted probabilities, the over all model's P-Value is reasonable at 0.291 hinting a moderate 70 % confidence level in the model.

However, the scatter plot of predicted probabilities, shown in Figure-1, paints a different picture. The age 5 seems to be critical age beyond which girl children have lesser chances of being rescued alive than the boys do.

The research and the findings are likely to be helpful to the rescue teams to plan their rescue efforts and accessories needed accordingly so that they can increase the chances of successfully rescuing every trapped child, boy or girl, alive always.

About the Author : The author, Ondiappan Arivazhagan "Ari", is an Honors graduate in Civil / Structural Engineering from University of Madras.He is a certified PMP, PMI-SP, PMI-RMP from PMI, USA. He is also a Master Black Belt in Lean Six Sigma and has done Business Analytics from IIM, Bangalore. He has 30 years of professional global project management experience in various countries around the World and has almost 14 years of teaching / training experience in Project management, Analytics, Risk Management and Lean Six Sigma .He is the Founder-CEO of International Institute of Project Management (IIPM), Chennai and can be reached at askari@iipmchennai.com