Volume-5, Issue-3, March - 2016 • ISSN No 2277 - 8160 IF : 3.62 | IC Value 70.36

Research Paper Engineering

Study of Causal Techniques and Evaluation Methods for Engineering Rsearch Models

Awari Mahesh Babu
Professor, Civil Engineering Department, Tirumala Engineering College, Hyderabad.TS. India.

ABSTRACT *In research problems and models causal techniques and appraisal of techniques are very important. Entity of the mainly considerable as well as numerous complexity concerning but here is statistical relationship along with a snap inconsistent Y and instructive inconsistent Xi. An alternative in the direction of reaction this predicament is just before obtain causal techniques in categorize to model its relationship. There are special kinds of causal techniques. The kind of the causal model depends planned the category of the allotment of Y. In modeling it is to predict the result Y place scheduled principles of a set of incoherent Xi. In this paper it is focused in causal and evaluation methods for validation of models. The intention of numerous proceedings techniques is in the track of wrest because untreated in progression the accurate evaluation.*

KEYWORDS : Causal, techniques, models, variance, covariance, evaluation

1.1 INTRODUCTION

In *statistical* modeling, causal techniques is a statistical method for approximation the relations among incoherent. It contains diverse methods on behalf of modeling furthermore evaluate a number of incoherent, Further explicitly, causal techniques ease one categorize how the characteristic appraisal of the contingent incompatible changes while numerous entity of the autonomous conflicting is speckled, whereas the other autonomous conflicting are held predetermined In causal techniques, it is further of implication in the direction of discern the deviation of the contingent conflicting about the causal role which know how to be described with a probability distribution. Causal techniques and evaluation methods investigational values of Delay and Carriage width (CW) are indentified.

2.0 CAUSAL TECHNIQUES

An easy causal techniques preserve demonstrate that the relations among an autonomous inconsistent X and a dependent inconsistent Y is linear, using the simple linear causal equation.

$Y = p_0 + p_1X_1$ ----------------- (1)

Where a and b are constants

Multiple causal will give an equation that calculate one inconsistent from two or more autonomous inconsistent,

$Y = p_0 + p_1X_1 + p_2X_2 \ldots\ldots$ (2)

Important steps concerned in performing causal techniques

Step-1: Construction of the causal model

The construction of an instructive model is a fundamental step in the causal techniques. It has to be described through position to the action theory of the intervention. It is expected that a number of category of inconsistent exist. An inconsistent might also correspond to an observable attribute or an unobservable one.

The model may imagine that a particular inconsistent evolves in a linear, logarithmic, exponential or other way. All the instructive models are constructed on the base of a model, such as the subsequent, for linear causal:

$Y = p_0 + p_1X_1 + p_2X_2 + \ldots + p_kX_k$, where -------- (3)

Y is the differ that the program is principally theoretical to make

$X_{1\text{-}k}$ are autonomous inconsistent likely to explain the change.$p_{0\text{-}k}$ are constants.

Step-2: Construction of a model

To apply multiple causal, a large model is generally essential.

Step-3: Records collection

Consistent information have to be collected, any from a monitoring scheme, from a questionnaire survey or from field survey.

Step-4: Computation of coefficients

Coefficients know how to be premeditated rather easily, using statistical software that is both reasonable and available to PC users.

Step-5: Valuation of the model

The model plan to elucidate as much of the variability of the experimental changes as possible. Assessment can be done from graphical and computational techniques and significance tests. The percentage of the variability in the y variable can be explained by the x variable.

2.1 LITERATURE REVIEW- EVALUATION OF METHODS

For techniques of the data it has to be fitted with one or more models. The techniques deals numerical fit techniques, graphical fit techniques and significance tests. The graphical measures deals with the residuals and prediction bounds and the numerical measures deals with the goodness of fit statistics as given below

1. Sum of squares due to errors (SSE)

The above statistic measures the total deviation of the response values from the fit to the response values. A value closer to zero indicates a better fit. Mathematically it is represented as

$$SSE = \sum_{i=1}^{n} w_i(yi - \hat{y}_i)^2 \text{ ------------- } (4)$$

Where is the response value and is the predicted value

2. R^2

These statistics dealings how successful the fit is in explaining the variation of the data. It is the square of the correlation between the response values and the predicted response values.

SSR is defined as

$$SSR = \sum_{i=1}^{n} w_i(\hat{y}_i - \bar{y})^2 \text{ --------------- } (5)$$

Where is the response vector

SST, which is furthermore called as squares about the mean, is defined as

$$SST = \sum_{i=1}^{n} w_i(y_i - \bar{y})^2 \text{ --------------- } (6)$$

Where SST = SSR+SSE

Therefore R^2 can be expressed as

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad \text{---------------} \ (7)$$

The coefficient of determination, R^2, is constructive since it gives the quantity of the discrepancy (fluctuation) of one inconsistent that is predictable from the other variable. It is a measure that allows us to determine how certain one can be in making predictions from a certain model/graph. The coefficient of determination is the ratio of the explained variation to the total variation.

The coefficient of determination is such that $0 < R^2 < 1$, and denotes the strength of the linear association between x and y.

3. *Degree of freedom (DFE)*

The residuals degree of freedom is distinct as the quantity of response values n less the quantity of fitted coefficients m predictable from the response values.

$$v = n - m \quad \text{----------------} \ (8)$$

v designate the quantity of independent pieces of information on which the approximation is based and is known as Degree of Freedom (DFE) of an estimate. A lesser amount of the degree of freedom less is the complication and degree of freedom also helps to evade the over fitting of the data.

4. Root Mean Square Error –RMSE. It is also known as Standard Error

This statistic is also identified as the standard error of the causal

$$RMSE = s = \sqrt{MSE} \quad \text{-----------} \ (9)$$

Where MSE is the mean square error or the residual mean square

$$MSE = \frac{SSE}{v} \quad \text{--------------} \ (10)$$

Standard Error (RMSE) value closer to 0 indicates a better fit.5

5. *Multiple correlation coefficients R*

It is the square root of R^2

Linear correlation coefficient,

r = Σ (xy) sqrt [(Σ x2) * (Σ y2)]-- -- -- -(11)

The quantity r, called the linear correlation coefficient, measures the potency and the direction of a linear correlation between two variables. The value of r is such that $-1 < r < +1$. The + and – signs are used for positive linear correlations and negative linear correlations, respectively. – ve sign indicates that one variable increases and other will decreases and vice versa.

3.0 CAUSAL TECHNIQUES AND EVALUATION METHODS
For case study transportation engineering problem is premeditated. The following table values are experimental values of Delay and stopping Carriage Width (CW)

Table 1

Links/ Variables	Delay	CW
1	17.00	8.21
2	11.23	9.40
3	8.72	1.80
4	9.56	5.61
5	19.73	1.23

Computation of good fit for causal models
i) Computation of good fit for risk parameters and equations
Computation of good fit for DELAY and CW

Linear model Poly3:

f(x) = +a2+a3x+a4--------(12)

Where x is normalized by mean 5.256 and std 3.678

Coefficients (with 95% confidence bounds):

a1 = -20.46 (-81.09, 40.17), a2 = 7.407 (-20.88, 35.69)

a3 = 21.35 (-44.62, 87.32), a4 = 6.637 (-19.64, 32.91)

The Strength and Significance coefficients:

MCC =0.97, SSE =4.18, R^2 = 0.95, SE= 2.04, DFE = 1

Linear model Poly4:

f(x) = a1 + +a3+a4x+a5---------(13)

Coefficients:
a1 = 0.03177, a2 = -1.117, a3 =12.37, a4 =-49.63, a5 = 64.29

The Strength and Significance coefficients:

MCC =1, SSE =1.89, R^2 = 1, SE = NaN, DFE =0

Table 6.6 Comparison of risk parameters

Polynomial degree	3rd degree polynomial				4th degree polynomial				Good fit polynomial
Parameter	SSE	R-square	RMSE	DFE	SSE	R-square	RMSE	DFE	
CW	4.18	0.95	2.04	1	1.89	1	NaN	0	3rd

iv) Comparison of risk parameters to get good fit graphical analysis
Comparison of parameters to get good fit graphical analysis for DELAY and CW

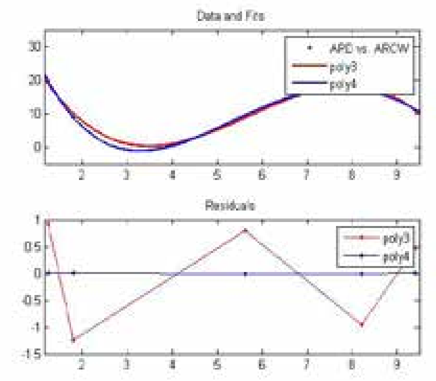

Fig. 1 Curves for good fit of DELAY and CW

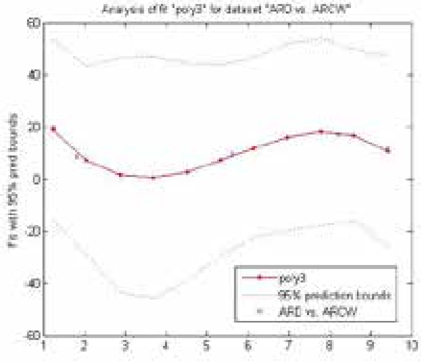

Fig.2 Prediction bounds for fit using poly3

Computation of good fit for DELAY and CW (Experimental values)

Linear model Poly3

f(x) = +n2+n3x+n4--------(12)

Where x is normalize by mean 45.15 and std 4.568.

Coefficients (with 95% confidence bounds):

n1 = 3.864 (-14.47, 22.19), n2 = -1.384 (-11.85, 9.083)

n3 = -2.596 (-23.56, 18.37), n4 = 32.42 (22.42, 42.42)

The Strength and Significance coefficients:

- Multiple correlation coefficient -MCC= 0.98, Sum of squares due to errors SSE = 0.76,coefficient of determination R^2= 0.96, Standard Error-SE = 0.87, Degree of freedom (DFE)=1

Linear model Poly4:

f(x)=n1 + +n3+n4x+n5---------(13)

Where x is normalized by mean 45.15 and std 4.568, Coefficients:

n1 = -2.039, n2 = 4.357, n3 = 1.555, n4 = -3.068 n5 = 31.81

The Strength and Significance coefficients:

MCC=1, SSE =2.52, R^2= 1, SE= NaN, DFE =0

3.1 CONLUSION

The 4th degree polynomial model gives R-square value equal to 1 but it over fits the data and increases the computational complexity. Fro graphical techniques it may be considered that 3rd degree polynomial fits the model and it is also observed that 3rd degree polynomial Sum of squares due to errors is less than 4th degree polynomial which indicates better fit. The causal techniques contrivance is advanced tools that be capable of categorize how unlike variables in a method are related. The causal tool will notify if one or multiple variables are interrelated among a method output. This in order can categorize where in the method direct is desired or what factors is the unsurpassed first point for a method upgrading task.

REFERENCES

1. J Shlens. 'A Tutorial on Principal Component Analysis'. Centre for Neural Science, New York University, New York City and Systems Neurobiology Laboratory, Salk Institute for Biological Studies La Jolla, CA 92037.
2. N. Ravishankar and D. K. Dey (2002), "A First Course in Linear Model Theory", Chapman and Hall/CRC, Boca Raton (p. 101)
3. Chiang, C.L, (2003) Statistical methods of techniques, World Scientific. ISBN 981-238-310-7 - page 274 section 9.7.4 "interpolation vs extrapolation"
4. Good, P. I.*; Hardin, J. W. (2009). Common Errors in Statistics (And How to Avoid Them) (3rd ed.). Hoboken, New Jersey: Wiley. p. 211*. ISBN 978-0-470-45798-6.
5. *Tofallis, C. (2009).* "Least Squares Percentage Causal". *Journal of Modern Applied Statistical Methods 7: 526–534*. doi:10.2139/ssrn.1406472.
6. M. H. Kutner, C. J. Nachtsheim, and J. Neter (2004), "Applied Linear Causal Models", 4th ed., McGraw-Hill/Irwin, Boston (p. 25)
7. *YangJing Long (2009)*. "Human age estimation by metric learning for causal problems" *(PDF). Proc. International Conference on Computer Techniques of Images and Patterns: 74–82*.
8. Altman DG. Practical Statistics for Medical Research. Chapman & Hall/CRC; 1991. .
9. Draper NR, Smith H. Applied Causal Techniques. Wiley Series in Probability and Statistics; 1998. Applied Causal Techniques.
10. Munro BH. Statistical Methods for Health Care Research. 5th ed. Lippincott Williams & Wilkins; 2005.
11. Steel, R.G.D, and Torrie, J. H., Principles and Procedures of Statistics with Special Reference to the Biological Sciences., McGraw Hill, 1960, page 288.
12. Rothman KJ, Greenland S. Modern Epidemiology. 2nd ed. Lippincot-Raven; 1998