AJMS

Asian Journal of Mathematical Sciences

RESEARCH ARTICLE

Big Data Investigation into the Causes and Treatment of Caries in Kindergarteners

Lam Kai Shun Carson

Department of Mathematics, University of Hong Kong, Pokfulam, Hong Kong

ABSTRACT

It has been thousands of years that tooth decay is a health problem among human beings (Chu, 2018). The disease is like our common daily influenza. The aim of this paper is to use the heated topic big data analysis and its related statistical mathematics to predict the possible behavior behind kindergarten children tooth care response – a predictive medicine for the prevention. Moreover, the paper also develops a thought experiment from the Bayes' decision tree. The aim is to determine some suitable strategies in the case of kindergarten tooth caries – for regenerative medicine. In the research, I have found that butterfly effect can form a predictive philosophy. I rationale it with Bayes theory and map each outcome with the corresponding domino effects (Heinrich theorem). While in the middle part, I insert random variables, respectively, as the connection. This event forms a completely new theory which can catch the chaos and dominos of the butterfly effects (philosophy) or the so-called Lorenz system. I propose the name should be the (HKLam's) Net-Seizing Theory. When my theory is expressed in terms of linear transformation, random matrix, and regression, one may use it in the prediction model (for the posterior distribution) of human behavior, etc. If we construct a reversed Bayesian tree with all necessary posterior predicted distributions (models), we may get the wanted forecasting prior (distribution) model and are a recursive formalism or the Bayes filter. We may establish the corresponding regression tree, etc. Furthermore, with the Savage theory, one can apply the machine learning technique to generate the necessary policy for handling the social problem which is just like the child's tooth care shown in this research paper – there is a need to subsidy our kindergarten as early as possible for the result of best social return.

Key words: Heinrich theorem, Influenza, Kindergarteners
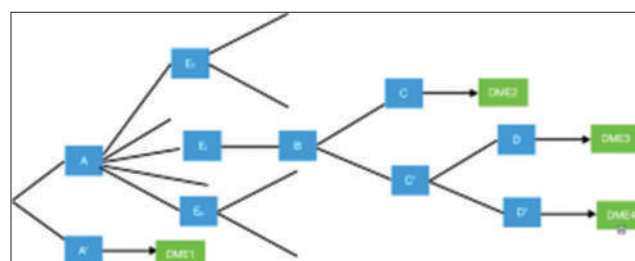
INTRODUCTION

Tooth decay is a common health issue as prevalent as influenza. Dental caries can often occur in children and young adults before the appearance of permanent teeth. Research has shown that fluoride can prevent tooth decay effectively. Recent studies have proven that the chemical compound, silver diamine fluoride when it is using in the tooth decay, it can treat dental carious lesions. Hence, the primary focus of this paper is to investigate the causes of cavities in kindergarteners (as they have milk teeth) in Hong Kong (3–6 years or students from nursery [K1] to upper kindergarten [K3]) and the efficky of fluoride treatment. Then, one may set up the wanted predictive model for cavity's

Address for correspondence:
Lam Kai Shun Carson
E-mail: h9361977@connect.hku.hk

development and hence develop a plan for the strategies of regenerative medicine in our teeth.

BACKGROUND – THE DECISION-MAKING THAT COMES FROM THE BUTTERFLY

The forward part of my decision making body (like the butterfly effect) is first illustrated from the divergent part of a Bayesian tree diagram that follows by a series of domino events using for particularly in the generation of a suitable policy [Figure 1].
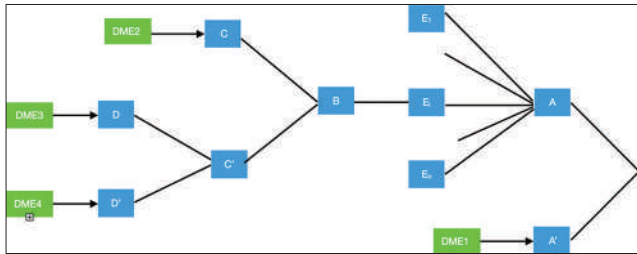
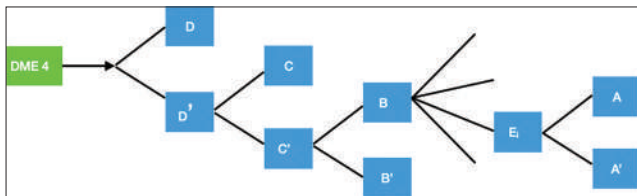**Figure 1:** Divergent part of my proposed decision-making body

Figure 2: The convergent part of my proposed decision-making body

E_1, E_2,… E_n are independent events with different probabilities Pr_1, Pr_2,…, Pr_n acts as the outcome probabilities. A and E_1, E_2,… E_n have their respective causal dependencies.

For the convergent part of Bayes theory plus Heinrich [Figure 2], one (like the second part of the mirrored image of the Bayes theorem) is mapped to those domino consequences (Dom_1, Dom_2, Dom_3,…Dom_n) corresponds to the different possibilities event E_1, E_2,…, E_n and the probabilities. Hence, one can use the Bayesian tree with the domino events (e.g. Dom_n,…,Dom_2, and Dom_1) to evaluate the efficiency of the hidden Markov Model (SEM).

Third, for the series of domino effect and the converse of the Bayes theorem, they constitute the backward part of my proposed decision-making body.

Indeed, DME 4 represents the corresponding domino consequence event 4 which is mapped to the inverse Bayes tree with the outcome event D and D'. One may then move upward the tree until event A and A'. I note that both the forward and the backward parts of the proposed philosophy come from the Bayes theory and the Heinrich theorem. Thus, I summarize my proposed one in catching the butterfly effects (maybe I name it as the (HKLam's) Net-Seizing theory) like the following:

1. A Bayes' probability tree diagram (Gorka and David, 2016)[1] which lists all the feasible outcomes of an event (forward and backward section from Bayes, Convergent, and Divergent Bayesian tree from decision-making – Savage theory)
2. A series of domino effect consequences or the Heinrich theory in risk management and engineering (Thierry, 2016)[2]
3. The immediate part (random variables) of linear mapping (Leung, 1974)[3] or transformation is added such that one may **rationale** (Lam, January 2019)[4] the philosophy. Hence, things can be approximated through the regression model.

My Original Discovery (Contribution) in the Linear Mapping of the Bayes Theorem

Consider the following imaginary statistical experiment that is often used when teaching statistics:

One should first toss three (or more) coins together in front of the participants such that heads are on the front and tails are on back. The participants are then required to list all possible outcomes after tossing the coins.

In general, the set of all feasible results (outcomes) would be:

S = {HHH, HHT, THH, HTT, THT, TTH, TTT} (suppose one will obtain such expected result). There are also other combinations and permutations of heads and tails when each of them is assigned with a number. Hence, the outcome space is not just unique.

Finally, the participants must count the number of tails for each possible outcome and then list them on a table, as seen below:

Possible outcomes (s_i)	HHH	HHT	HTH	THH	HTT	HTT	TTH	TTT
No. of tails obtained t_i	0	1	1	1	2	2	2	3

(Clearly, the above table does NOT list all possible outcomes for tails.)

A function f (which should be called the random variable) must also be considered. Intuitively, f maps all the possible outcomes (s_i) to the number of tails T_i

f_i: s_i | s t_i (i.e. no. of tails obtained as it takes values $t = 0,1,2,3$ … (Leung, 1974)

or strictly speaking

f: Sri T (i.e. **R – Real Number**)

Actually, through the linear mapping, the non-linear chaos can be converted into the linear one through a suitable choice of random variables. One of the cases is the mapping between Bayes

outcome space, random variables, and the domino. Hence, one can find the matching linear regression model approximations for different research areas. Obviously, in such sense, we can "catch the chaos and the dominos." This event is known as capturing the "Butterfly Effects (or the philosophy)" behind. Thus, I propose such kind of methodology should be termed as the HKLam's Net-Seizing theory.

When my theory is expressed in terms of both mathematics and statistics, it becomes

[Random Matrix] [LT] = Linear regression of the causal domino effects

And

[Linear regression of the causal domino effects] [LT] = [Random Matrix]

For the first part of Bayes plus Heinrich, one may further get a diagram [Figure 3] that rationales it and gets the structure of my Net-Seizing theory. Actually, one should insert the probability that multiplied with weight Pri*Wgi between the random variables RVi and the last branch of event Ei. For the converse part of my Net-Seizing theory [Figure 4], each domino causal effect is corresponding to a suitable linear transformation
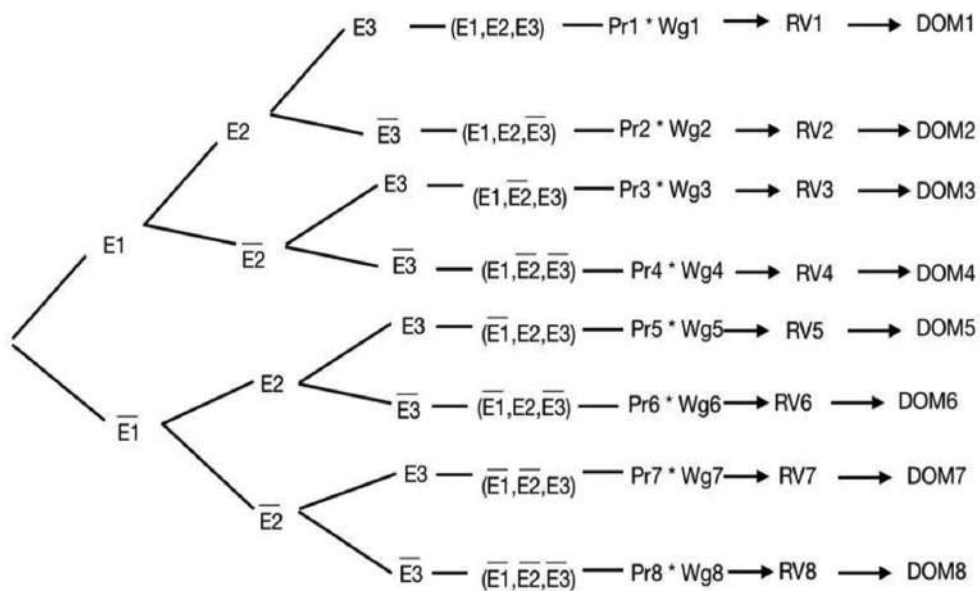

Figure 3: The rationalization of the forward part of my proposed sed decisionral Capital and Parental Mediation for Digital Inequity. Thousand Oaks, California: SA
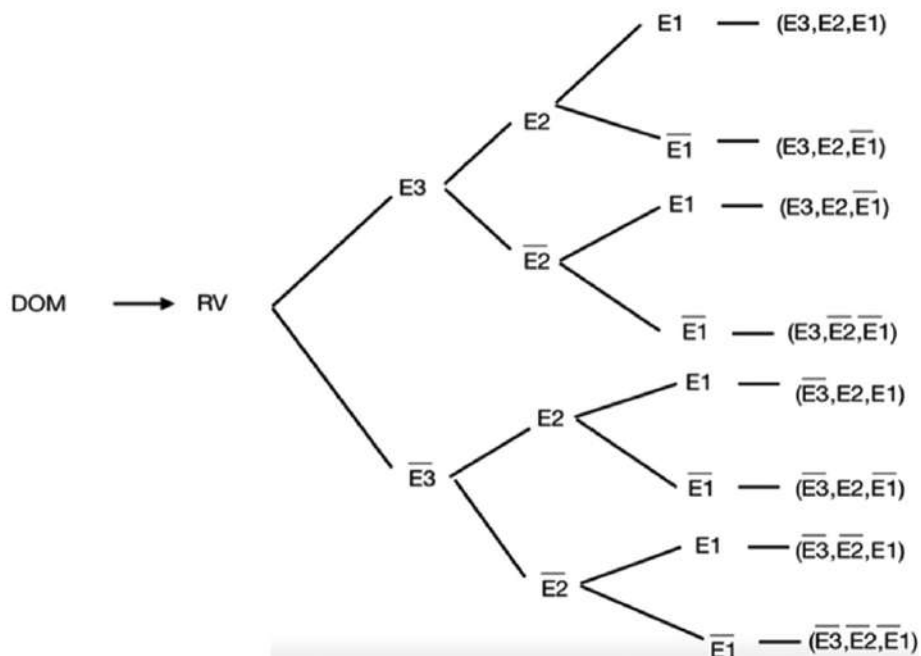

Figure 4: The rationalization of the converse part of my proposed "Net-Seizing" theory with random variables as the immediate connection between events and dominos

(random variable) and further mapped to a Bayesian tree. Hence, one can then express my Net-Seizing theory in terms of mathematics and statistics and get the following diagram [Figure 5] in matrix notation. The convergent (or the backward but NOT the convergent – mirror part) of the matrix notation diagram [Figure 6] is also true.

The study of the students reading behavior in Lam, November 2018[5] can be acted as my proposed theory's example [Figures 7-11]. The following diagrams outline the corresponding inverse

Bayesian tree diagram and domino consequences such that one may find the respective conditional probabilities for the convergent part of my proposed theory in the following page.

Practically, we can calculate the corresponding conditional probabilities of the convergent part of Bayes plus Heinrich from the inverse of the Bayesian tree diagram like the following:

At the same time, with referencer to the results in Lam, December 2019[6] and Lam, March 2020,[7] this author discovers that one may always
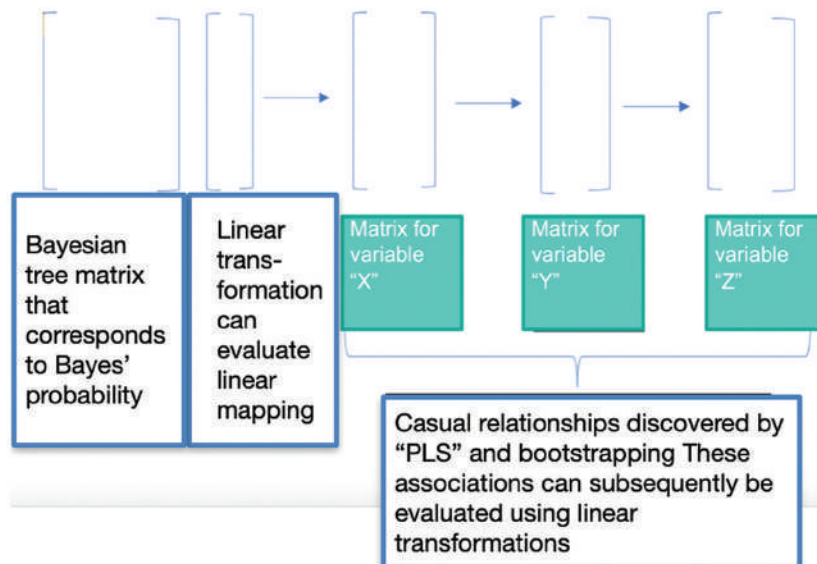

Figure 5: One can express the random probability matrix of my proposed philosophy in terms of a linear transformation to those causal relations that found by partial least square method like structural equation modeling
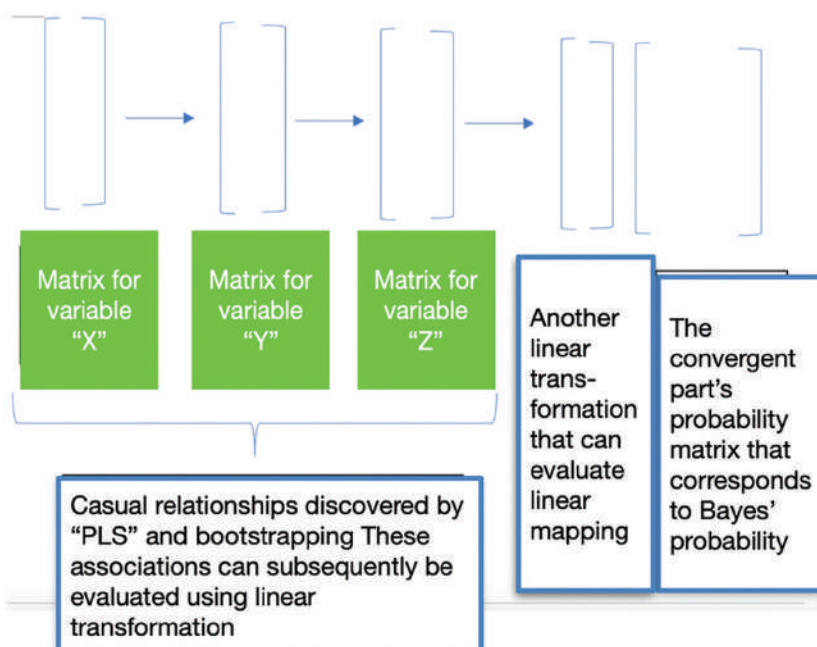

Figure 6: One can approximate the causal relationships that found by partial least square method from a linear transformation followed by the random probability matrix of my proposed philosophy, in matrix notation

express the Bayesian probability tree in terms of a matrix. The outcome matrix can finally be approximated by applying the Bayesian linear regression. Or we may have the following results (– assume the square matrix operation while the non-squared one can be converted to the squared by MatLab coding but is out of the scope of the present paper):

$$
\begin{pmatrix} P(A_1,\mu_1,\sigma_1^2) \\ P(A_2,\mu_2,\sigma_2^2) \\ . \\ . \\ . \\ P(A_n,\mu_n,\sigma_n^2) \end{pmatrix} = \begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ \beta'_n \end{pmatrix} +
$$

$$
\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}
$$

However, the probability matrix A can be expressed as:

$$
\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & . & . & . & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & . & . & . & P(A_2|B_n) \\ . & . & . & . & . \\ . & . & . & . & . \\ P(A_n|B_1) & P(A_n|B_2) & . & . & . & P(A_n|B_n) \end{pmatrix} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}
$$

$$
= \begin{pmatrix} P(A_1) \\ P(A_2) \\ . \\ . \\ . \\ P(A_n) \end{pmatrix}
$$

Therefore, accordingly, we have:

$$
\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & . & . & . & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & . & . & . & P(A_2|B_n) \\ . & . & . & . & . \\ . & . & . & . & . \\ P(A_n|B_1) & P(A_n|B_2) & . & . & . & P(A_n|B_n) \end{pmatrix} =
$$

$$
\begin{pmatrix} P(A_1) \\ P(A_2) \\ . \\ . \\ . \\ P(A_n) \end{pmatrix} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1}
$$

Substitute back the right hand side of the first equality into the above equation, we have:

$$
\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & . & . & . & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & . & . & . & P(A_2|B_n) \\ . & . & . & . & . \\ . & . & . & . & . \\ P(A_n|B_1) & P(A_n|B_2) & . & . & . & P(A_n|B_n) \end{pmatrix} = \{
$$

$$
\begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} \}_{} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1}
$$

However, from the result of Lam, March 2020,[19] the conditional probability matrix A (given B) can also be expressed in terms of linear regression through a suitable linear transformation (LT), thus we have:

$$
\begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} \}_{} \begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1} =
$$

$$
\{X + (XB+B_0) + [(XB+B_0)D + D_0]\} * (LT)^{-1}
$$

The above outcome implies that when the Bayesian linear regression is multiplied by the inverse of the event B's probability matrix, it can be well approximated through an ordinary linear regression equation multiplies with the inverse of the linear transform of conditional probability matrix A (given B). Or one can find out the inverse

of the event B's probability matrix if we know the inverse of Bayesian regression.

Hence, one can apply both linear regression, and the Bayesian regression (for predictive model) methods for the approximation to a random matrix and more about the inverse of probability matrix B:

$$
\begin{pmatrix} P(B_1) \\ P(B_2) \\ . \\ . \\ . \\ P(B_n) \end{pmatrix}^{-1}_{=} \left\{ \begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} \right\}^{-1}
$$

* {**X** + (**XB**+**B**$_0$) + [(**XB**+**B**$_0$)**D** + **D**$_0$]}*(**LT**)$^{-1}$

Substitute back into the equation:

$$\mathbf{A} = \mathbf{X} \, \boldsymbol{\beta'} \, \boldsymbol{\beta}^{-1} + \boldsymbol{\varepsilon'}$$

We have:

$$
A = X\beta' * \left[\left\{ \begin{pmatrix} x_{11} & . & . & . & x_{1n} \\ x_{21} & . & . & . & x_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & . & . & . & x_{nn} \end{pmatrix} \begin{pmatrix} \beta'_1 \\ \beta'_2 \\ . \\ . \\ . \\ \beta'_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix} \right\}^{-1} \right.
$$

* {**X** + (**XB**+**B**$_0$) + [(**XB**+**B**$_0$)**D** + **D**$_0$]}*(**LT**)$^{-1}$] + ε"——————————————-(*)

To sum up, one can always express the Bayesian linear regression of the probability matrix of event A in terms of the probability matrix B, an ordinary linear regression, and the inverse of the linear transform to conditional probability matrix A (given B). At the same time, the conditional probability matrix A (given B) can be expressed as the substitution back of the product among (the inverse of Bayesian linear regression, an ordinary linear regression, and the inverse of the linear transform) into the Bayesian linear regression equation of conditional probability matrix A. One of the applications of the above matrix equalities (or equations) is in the field of quantum computing. One can always use the causal relationships (quantum switch) and a random matrix, etc., to model different quantum systems in an approximated manner. The final result is shown in the above matrix expression (*) for our conditional probability A (and B) – a random matrix modeled to a quantum system.

In much a similar manner, one may also use the aforementioned matrix expression (*) to model our living organisms' mutation together with its cause and effect. The result can then be applied in tissue engineering or regeneration. In practice, the method to find out the inverse of a matrix is an interesting topic for our university's advanced mathematics study (mathematics major students) or the course such as Matrix theory and its application.

MATERIALS AND METHODS

This research aims to answer the following questions:

1. What is the level of tooth decay observed in K1 kindergarteners at the time of the first examination?
2. What is level of tooth decay observed in older (lower kindergarten [K2], K3) kindergarteners after fluoride treatment?
3. Why we need fluoride treatment and what are the implications of it?
4. What the predicted outcomes will see from K2 and K3 children fluoride treatment during the prevention of tooth decay?
5. How should the authorities develop corresponding strategies for regenerative medicine of teeth?

The investigation will comprise the correlation and survey methods to answer these questions. First, one will use quantitative statistical research methods to determine the relationships between the different variables. Second, one will also use a qualitative survey to improve the understanding of the implications of fluoride treatment in kindergarteners.

Both, the correlation studies and survey, will find out the implications of fluoride treatment before analyzing the causes. Finally, the correlation design will help in predicting the incidence of tooth decay after fluoride treatment.

The study will begin by making records of local Hong Kong kindergartens belonging to K1 and K2 (Chu, 2000).[8] Subsequently, one will use a tailor-made software to pull the necessary data into MS Excel. One will record down the variables such as name, examination and birth dates, the number of decayed and infected teeth, and treatment required. From, the collection of these data, one may apply them to Bayesian inference and Bayesian regression from the correlation relationship through simulation using software such as R. Hence, one can set up the needed predictive model for caries, etc. While

it is critical that sample size, the value of precise measurements, and the use of unbiased samples during the quantitative research. This study will use both explanatory and predictive designs. We propose a few variables for the former: Level of tooth decay observed at the first examination, the status of cavities after fluoride treatment, and the number of missing teeth. Subsequently, the researcher will use a model comprising these variables and the number of decayed teeth. The researcher will compare one group with decayed teeth after fluoride treatment with another group with decayed teeth without fluoride treatment, that is, the control group. The total number of participants will be 180,000. Indeed, there is a prediction research design. The investigator can identify those specific variables to help predict the behavioral outcomes from the prediction model setting up before we have mentioned. This study concerns with fluoride treatment, the level of tooth decay, and how it can aid in the prediction of healthy teeth. Tooth decay is a time-related variable. The research will correlate the predictor and outcome variables. We will use multiple regression to determine the predictor's effects on the variables. Hence, one can find out the desired causal relationships with a structural equation model or partial least square method from the collected data. Thus, one can predict the outcomes of after the fluoride treatment.

A longitudinal study (Creswell, 2012)[9] involves repeated observations of the same variables (e.g. population) over a certain period. Thus, we will also conduct qualitative yearly interviews (until K3) with the previous (K1) kindergarten students. The researcher will ask questions from an interview guide, listen to the answers and/ or see behavior, and record the responses. The interviewer asks open-ended questions, without options for responses, and records the answers provided in a qualitative interview. We usually use a longitudinal study to identify the reasons for the development of outcome implications (e.g. the implications of fluoride treatment). This information can guide the plan to develop better strategies for regenerative medicine for teeth.

There are being some ethical considerations while handling the mankind type of big data investigation. The concerns may include privacy, informed consent, data ownership and transparency, etc. This applicant believes that the researchers can solve those public worries if they can eliminate those misunderstandings among the investigated respondents.

RESULTS AND DISCUSSION

This applicant tries to reference the wanted research results (as suggested taking place in Hong Kong Kindergarten school) through using those data obtained from the United State Centre for Disease Control. The major working steps are:

1. Developing a suitable predictive philosophy such as the butterfly one (as shown in the previous section) and rationalizing it by my Net-Seizing theory

2. Using the structural equation model method to find out the hypothesis's causal relationships, so as the one indicated by Australia's influenza case in Lam June 2019[10]

3. Convert the data (collected from U.S.A CDC Department) into a time series. Perform the Granger Causal Connectivity Analysis (Seth, 2009)[11] by the software EViews. Fit those data into the maximum likelihood estimation (or using Hayes' mediated-moderated model) to approximate the causality behavior with multilevel modeling whenever necessary. Then, one can continue the Bayesian estimation

4. Applying Bayesian inference to find the prior and likelihood distribution of the collected data. Hence, one can find the deserved Posterior distribution from the United States' tooth cavity's data. SPSS software can perform all the mentioned work

5. After finding the parameters (mean and standard derivation) for the posterior distribution, one may then use the Bayesian regression for the following step of posterior's estimation. This applicant suggests one should use the software JASP for the calculation of the wanted Bayesian regression model (or equation), hence, one may predict the behavior of a child with tooth cavities

6. From the predicted posterior distributions (models) together with the reversed Bayesian tree, one may further construct the forecasting prior distribution (model) for our child tooth cavities. This is indeed a kind of recursive formalism

7. Establish the corresponding (probability) decision tree (or random forest) or even Bayesian adaptive tree for Bayesian causal inference from our tooth statistics; the aim is to determine how well is the tooth situation in the countries like the U.S.A. and thus create some suitable strategies in the regenerative medicine for the case of teeth

8. Implement the strategies or policies found in the previous step.

The following is the outcomes for Bayesian inference, where we can obtain the outputted tables from using the software SPSS:

Group statistical value					
Indicator_Value		Counts	Average value	Standard error	Standard average error
Data_Value	= 1	72	50.110	15.5512	1.8327
	= 2	46	45.896	12.8593	1.8960
	=3	72	20.844	8.8841	1.0470

Posterior distribution of independent sample average values					
	Mode	End-stage average	Variation	95% confidence interval	
				Upper limit	Lower limit
Data_Value= 1	−4.214	−4.214	7.218	−9.490	1.062
= 2	−29.265	−29.265	4.584	−33.470	−25.061
= 3	−25.051	−25.051	4.890	−29.398	−20.704
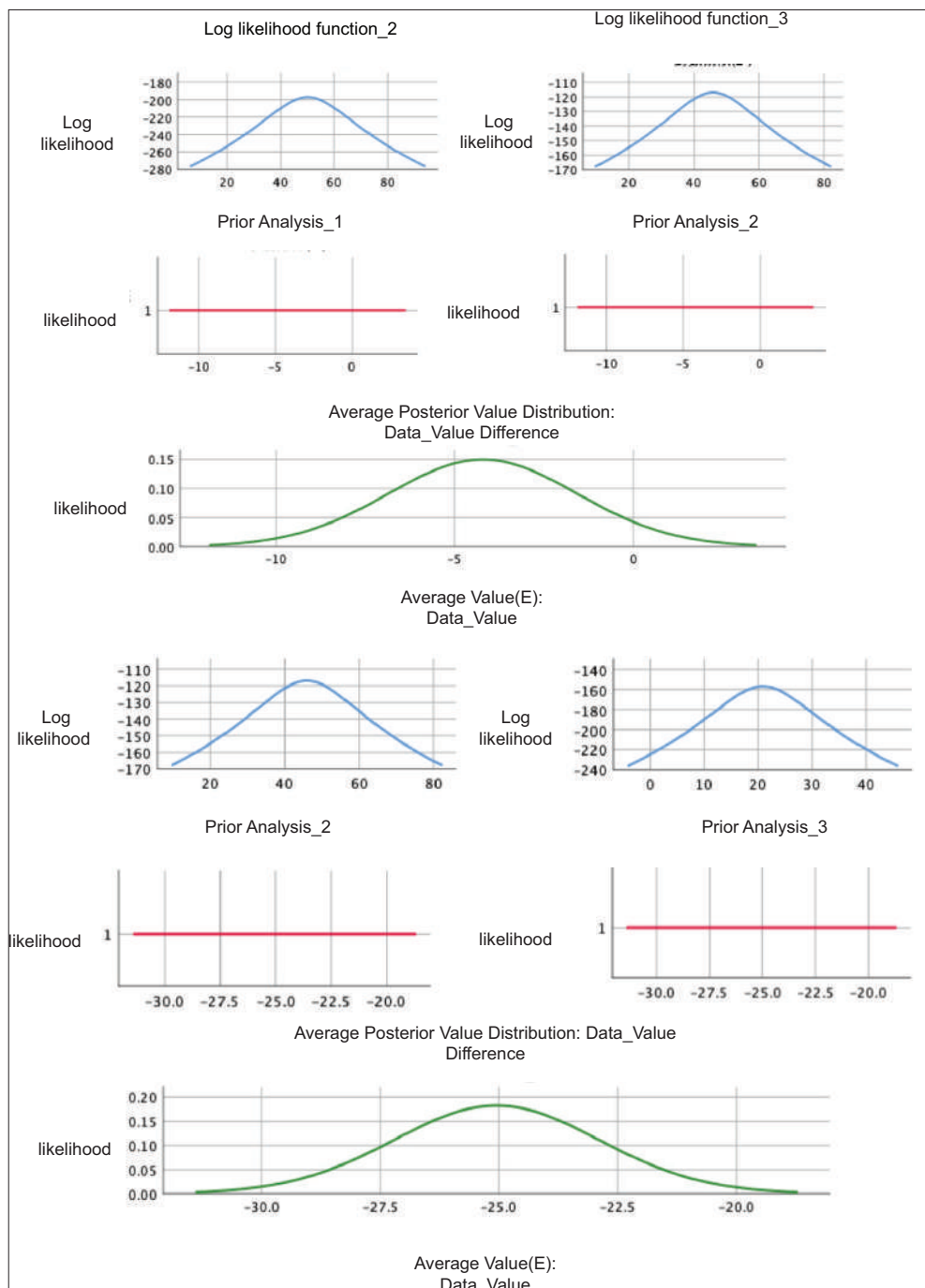
The regression equation (or model) is: Data_Value = 0.135*(Response_Rate) + 0.196*(Percent_eligible_for_the_NSLP_SampleStudents) + 24.881.

Coefficient	Mean
Model 1	
Intercept	24.881
Response_Rate	0.135
Percent_eligible_for_the_NSLP_SampleStudents	0.196
Bayesian Regression Equation (1)	Data_Value = 0.135*(Response_Rate) + 0.196*(Percent_eligible_for_the_NSLP_SampleStudents) + 24.881
Model 2	
Intercept	24.881
Sample_Size	0.001
Response_Rate	0.060
Bayesian Regression Equation (2)	Data_Value = 0.001*(Sample_Size) + 0.06*(Response_Rate) + 24.881
Model 3	
Intercept	24.881
Sample_Size	0.001
Percent_eligible_for_the_NSLP_SampleStudents	0.073
Bayesian Regression Equation (3)	Data_Value = 0.001*(Sample_Size) + 0.073*(Percent_eligible_for_the_NSLP_SampleStudents) + 24.881

The regression equation (or model) is: Data_Value = 0.001*(Sample_Size) + 0.06*(Response_Rate) + 24.881

The regression equation (or model) is: Data_Value = 0.001*(Sample_Size) + 0.073*(Percent_eligible_for_the_NSLP_SampleStudents) + 24.881.

This applicant notes that one can use the similar techniques as listed above in modeling the Data_Value corresponding to three different indicators of the tooth treatment data. Hence, one can predict the tendency of tooth decay that relates to the children daily behavior. One of the tooth indicators – untreated tooth decay Bayesian linear regression table (from statistical software JASP) is listed below in the next page.

With reference to the above result, one may obtain the Bayesian (predictive) regression equation (model) as: Data_Value = 6.302e-5*(Sample_Size) - 0.143*(Response_Rate) - 0.079*(Percent_eligible_for_the_NSLP_SampleStudents)

Coefficient	Mean
Model	
Intercept	23.410
Sample_Size	6.302e-5
Response_Rate	-0.143
Percent_eligible_for_the_NSLP_SampleStudents	-0.079
(Predictive) Bayesian regression equation	Data_Value = 23.410 + 6.302e-5*(Sample_Size) -0.143*(Response_Rate) – 0.079*(Percent_eligible_for_the_NSLP_SampleStudents)

Hence, it is obviously that the percentage of untreated tooth decay is indirectly proportional to the response rate. This event suggests that the investigated population may be among the low socioeconomic position (Ekholm, 2010).[12] My predictive models (or equations) tell us how
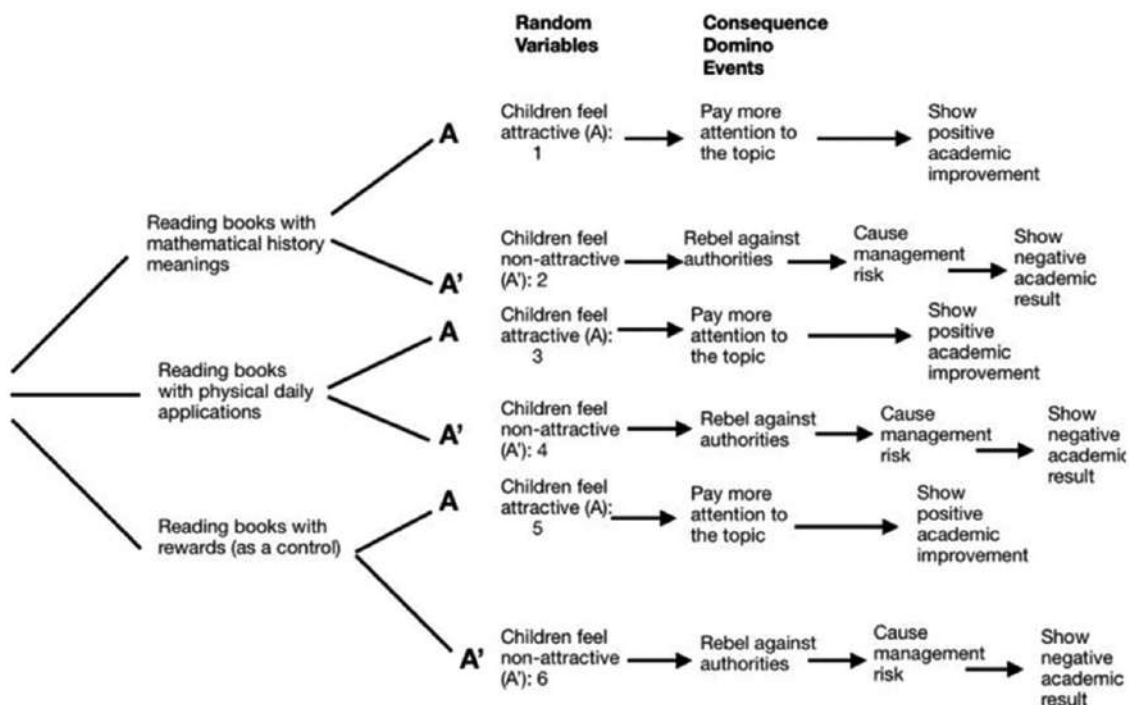

Figure 7: The forward pat of the students' reading behavior, random variables, and the domino consequences (similar to Savage theory in decision-making)
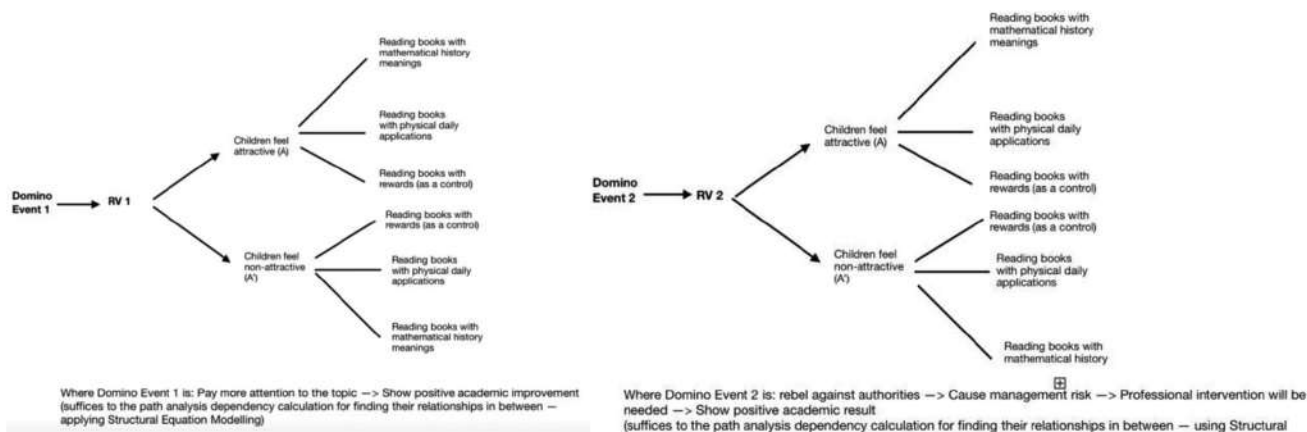
Figure 8: The backward parts of the students' reading behavior and the domino consequences

these low socioeconomic people will take less care (behavior) to their tooth decay. This event is because they may have more tooth decay without any medical treatment. Hence, one may perform a similar technique for the prediction of behavior to another two types of tooth indicators.

A Decision tree with expected values for finding a suitable strategy in teeth, regenerative medicine – a thought experiment

From my butterfly effect philosophy, one can further develop the Bayesian decision theory for finding the best strategy during decision making. Thus, another thing that we need to discuss is finding the most suitable strategy for our regenerative medicine. Suppose there is a thought experiment including five grades of children with different cost for making decisions such as whether one should join the tooth care scheme (JTC). This event (the thought experiment) is shown in the following diagram [Figure 12] in the next page by employing the software – SpiceLogic Decision Tree analysis.

The software finds that locationID > 8 is the evaluated policy for stochastic dominance (to determine the expected utility maximizer). One may also get all the feasible paths with their corresponding payoff:

With reference to both of the results (paths' payoff and the decision tree), the total payoff for the suggested path LocationID >8 is only 1140, which nearly attains payoff's highest value (1200 for the JTC with Tooth_Care_4). The minimum value of the payoff is 270 and the medium value is 560 corresponding to the path with Data_Value smaller than 21.3 and also there are some other paths that lead to the same payoff. If one selects the proposed path, this event shows that when there is a tooth care scheme, the society can always obtain the best return. Moreover, the thought experiment implies that when the researcher collects enough daily data, one can find the most suitable strategy. Hence, one can get the best path under the condition that there is a need to maximize the payoff. I remark here that there is a public policy selection theory (Ebenezer, 2016)[13] which is named as rational choice theory. Its main context is explaining social phenomena as the outcomes of a personal action that is considered to be rational. The advantage of rational theory is that it can guarantee the decision-making (selection) of a suitable policy (Emeka *et al.*, 2008)[14] can result in a maximum
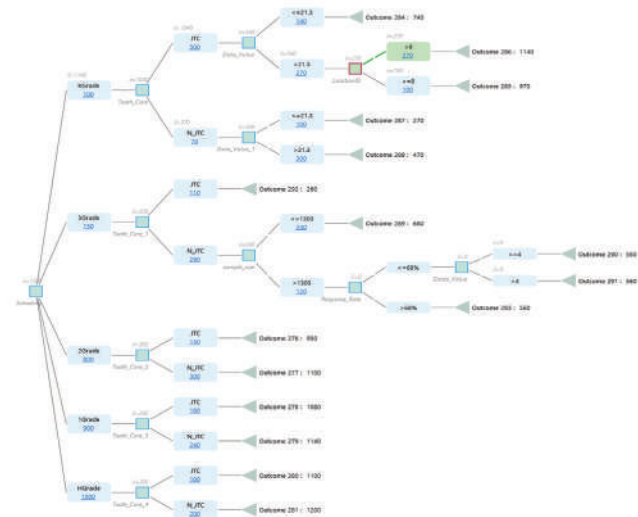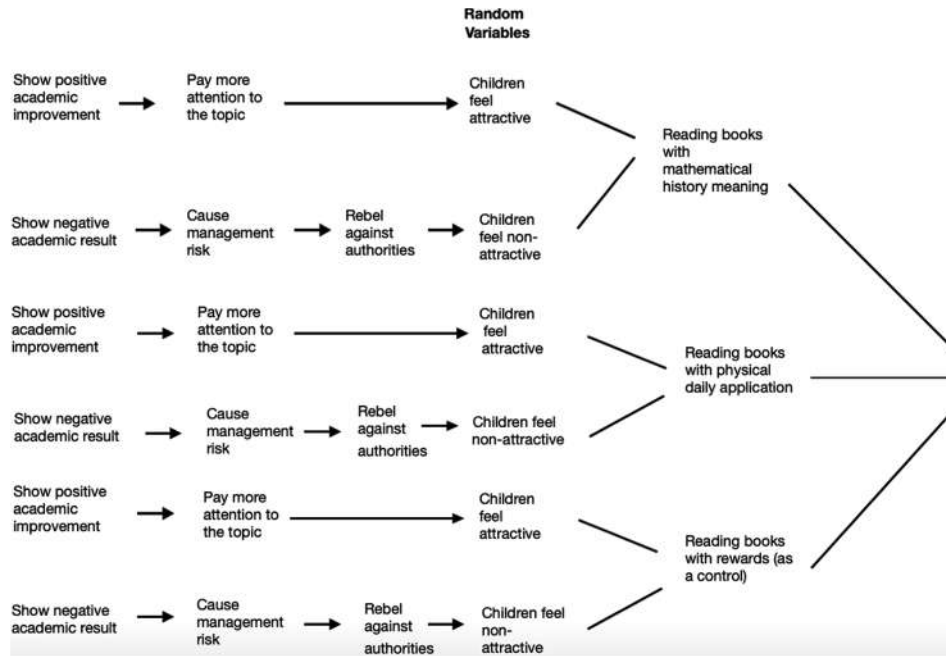
Figure 9: The convergent parts (mirrored image of Bayes' theory) of the students' reading behavior and the domino consequences.
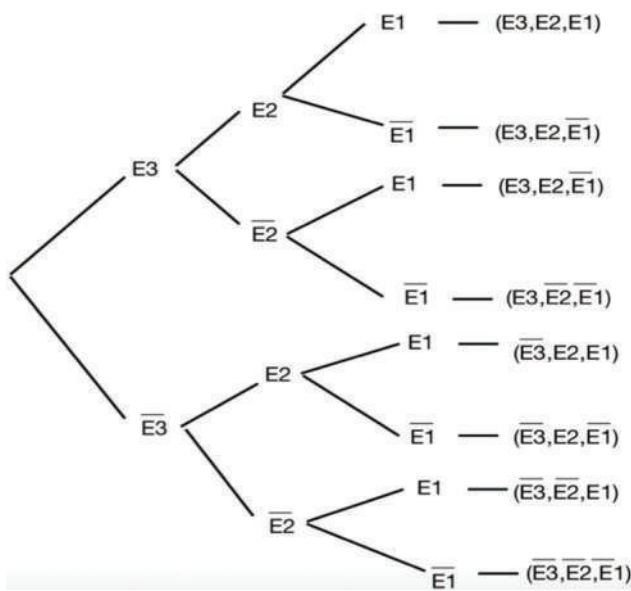

Figure 10: The inverse part of the Bayes theorem for computing the required conditional probabilities to the convergent part of my proposed philosophy
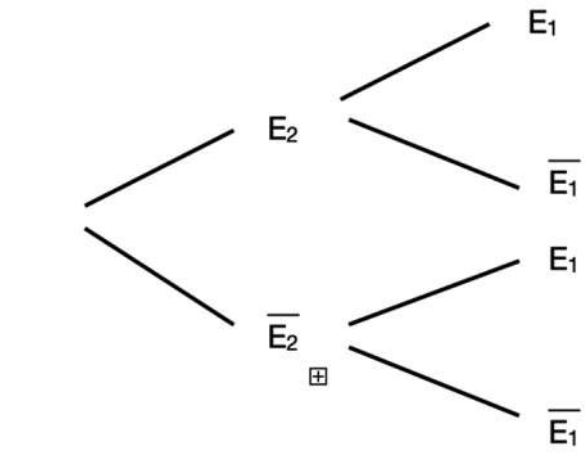

Figure 11: The second level of the inverse part of the Bayes theorem for computing the required conditional probabilities to the convergent part of my proposed philosophy

social gain or benefit. Obviously, my thought experiment fits well with the rational theory.

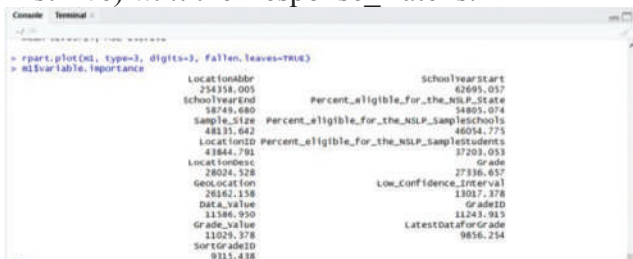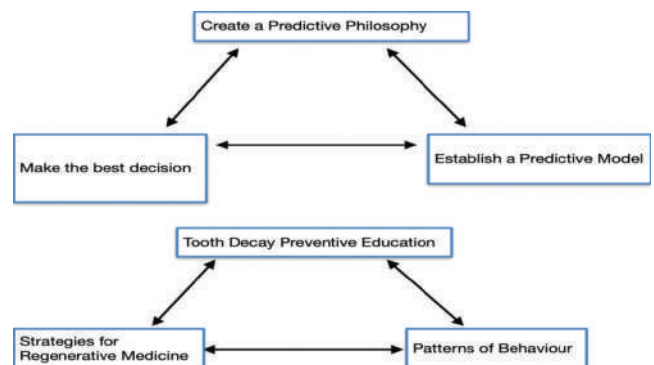## Regression tree for predicting the trend

Figure 12: The proposed of events for the prevention of tooth decay

One can also use a regression tree for the foreseeing the trend features of collected dental data from the

U.S. CDC. In the following section, this author applies the "rpart" and the "rpart.plot" function of the R program to find out the predicted Response_Rate percentage and hence determines the mean absolute error from the difference between actual and predicted values. In the next page, we can show a regression tree in terms of the Response_Rate.

The highest predicted Response_Rate is 85.9% with the probability of only 0.02. This event corresponds to the "locationabbr" of OR and WV together with the SchoolEndYear greater than year 2018. Another highest predicted Response_Rate is 78% with the probability of nearly 0.26. The event respects to the Percent_eligible_for_the_NSLP_SampleStudents of less than 19% and "locationabbr" other than AR, HI, IA, KY, NH, and SC. This event is also the most probably expected one that occurred. It seems that areas with "locationabbr" not included in AR, HI, IA, KY, NH, and SC have greater willingness of response. Thus, the complementary areas may be richer or have some kind of funding in dental care also. However, as the probability is not too low which tells us that the population of this group of people is about one in a fourth. The lowest predicted Response_Rate is zero with the probability of only 0.018. This event also corresponds to SchoolYearStart less than year 2000 with the Percent_eligible_for_the_NSLP_SampleStudents of less than 19%. The likelihood of the rest predicted Response_Rate is around 0.02 to 0.08, which are quite evenly distributed.

Indeed, the importance of difference variables (the first five) w.r.t the Response_Rate is:

The above result tells us that the five importance of variables which include the location (i.e. different states), the start and end of the school years, the percent of eligible for NSLP state, and the sample size do determine the response_rate. In other words, where the respondents live, when the research starts and how large is the sample size

with the percent of eligible are the key factors (or the social-economic background behind the population) that related to the respondent's population (or who will be investigated). It seems that area with code AR, HI, IA, KY, NH, and SC is much richer or having some kind of funding scheme in fighting tooth decay that starts from 2018. This author also remarks that the corresponding R code for the regression tree is shown in below:

```
> library(readxl)
> imputation_1 <- read_excel("C:/Users/User/Desktop/imputation_1.xlsx")
> View(imputation_1)
> library("rpart")
> library("rpart.plot")
> m1 <- rpart(Response_Rate ~ ., data = imputation_1, method="anova")
> m1
> rpart.plot(m1, type=3, digits=3, fallen.leaves=TRUE)
> m1$variable.importance
```

This author remarks that one may compare different Bayesian linear models from their Bayes factors with respect to their BF_M and BF_{01}. The smaller the value of BF_{01}, the higher the chance for the model hypothesis H_1 than H_0 to be likely occurred. This event is because BF_{01} is just the reverse of BF_{10}. In addition, one may find out the expectation values from the different BF_Ms. Hence, we can select the highest BF_{10} with well-balanced BF_Ms in finding the best Bayesian prior model. (Notes: Both the P(M) and P(M|Data) may need to be considered for selecting the best model also). Next, from the best prior model selected, we can have calculated the corresponding posterior model through Bayesian estimation with the calculated mean as the expectation values. Finally, one may discover the policy that relevant to that posterior model. Actually, the symmetrical theory is applied whenever one needs to public relation problems.

The following is the practice of what the above mentioned in theory:

The above data show that the best model found is:
Response_rate ~ Sample_Size + Percent_eligible_for_the_NSLP_SampleSchools + Grade_Value

The corresponding BF_{10} is 19.982 which is highly expected to occur than the null hypothesis. Moreover, the BF_M is 32.254 which is the expected value for that particular model. In the coming paragraph, this author will perform the Bayesian estimation for the posterior model.

When one is performing Bayesian estimation for the posterior model using statistical software SPSS, the result is listed below:

Models comparison	P(M)	P(M|Data)	BF$_M$	BF$_{10}$	R$_2$
Indicator_value + Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value + Data_Value + LocationID	0.143	0.189	1.394	1.000	0.065
Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value	0.007	0.188	32.254	19.982	0.058
Sample_size + Percent_eligible_for_the_NSLP_SampleStudents	0.010	0.115	13.531	9.161	0.047
Indicator_value +Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value + Data_Value	0.024	0.097	4.428	3.103	0.064
Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value + Data_Value	0.0100	0.082	9.235	6.489	0.060
Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value + LocationID	0.010	0.063	6.972	4.999	0.059
Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value + LocationID	0.024	0.058	2.533	1.852	0.062
Indicator_value +Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value	0.010	0.051	5.644	4.096	0.058
Indicator_value +Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + Grade_Value + LocationID	0.024	0.036	1.514	1.148	0.059
Sample_size + Percent_eligible_for_the_NSLP_SampleStudents + LocationID	0.007	0.019	2.674	2.003	0.047

Bayesian estimates of coefficients [a,b,c]

Posterior				95% confidence Interval	
Parameter	Mode	Mean	Variance	Lower Bound	Upper Bound
Sample_Size	39.535	39.535	15.547	31.804	47.266
Percent_eligible_for_the_NSLP_SampleStudents	0.001	0.001	0.000	0.000	0.001
NSLP_SampleSchools	0.223	0.223	0.002	0.136	0.310
Grade_Value	1.762	1.762	0.868	0.065	3.589

[a]Dependent Variable: Response_Rate, [b]Model: Sample_Size, Percent_eligible_for_the_NSLP_SampleStudents, NSLP_SampleSchools, Grade_Value, [c]Assume standard reference priors

From the above data, it is observed that Grade_Value has the highest mean value. This event implies that when the tooth healthy scheme is implemented in K1 or the earliest, the expected outcome and return to prevent tooth decay in U.S. kindergarten students will be better. Since for those students in the higher grade like K3 will soon have their milk tooth replacement in the coming one or two years. If the youngest kindergarten students can protect their tooth during the earliest stages, there will be a less expenditure that spent before the permanent tooth. Thus, this author suggests that there is a need to have tooth healthy checking scheme – a public policy for our kindergarten students in Hong Kong. My proposed public policy also agrees well with the aforementioned rational theory in the previous paragraph.

Predictive models and their implications

It should be note that from the above regression tree model, one can further perform the prediction with the R program code "predict" to obtain the wanted forecast values. In an example, if one's initial assumption believes that a particular coefficient (β) is positive but the fitting result gives a negative one, a contradiction appears. One of the case study is the fitting of physical experiment data that are used in the nuclear databases – inconsistent data that cause system error (Georg, 2018). This event shows the philosophy, logicism behind the model (Siu, 2009).[15] Georg in 2018 notes that one can finally celebrate these experiment data back to become best fitted and analyzed through the Bayesian methods (Georg, 2018).[16] In such case, those prescribed data without precise fitting celebrations may be acted as the deserved working case for our intelligent machines. A practical logicism model case study is the chromatic induction model for the integrated visual system. Simultaneously, when we continue to feedback the predictive values into the Bayesian tree, (say for 5 times) according to the procedural rules, then the outcome implies the existence of philosophy, formalism can be used to model (Lorenzo et al., 2011) the nested causal relationships behind. A generalized case study is the predictive loop

model that works under procedural laws. This event is used for the prediction of the earthquake when it cooperates together with the database management system. Finally, when one develops the Hidden Markov Model w.r.t. the data (such as in this thesis one) given, then this event provides us with the philosophy, intuitionism behind. Once we can compare the RMSE values calculated from the above three models mentioned and thus decide the philosophy indicated. This event is because the RMSE values usually tell us the level of model fitting. Obviously, the best fitted model (with the above three cases) will then demonstrate which kind of the philosophy laying backward among the data collected. In this thesis, if the Hidden Markov Model computed has the lowest RMSE value when one compares it with the other two, the Markov model will be the best fitted one, the tooth care behavior of our investigated people is thus only motivated according to their intuition.

To sum up, one can always express implicitly the three types of philosophy (can be extended to various ones) in terms of three kinds of mathematical model during their Prediction, respectively. In other words, one may always relate the models to philosophies. This event encourages us to model our body, mind, and spirit. It is because in Lam, July 2016, I connect the body, mind, and spirit with formalism, intuitionism, and logicism. If it is true that I can also mathematically model the above three philosophies, the implication is one can also indirectly model our body, mind, and spirit. Another side effect is we can also teach our computer what philosophy is Lau and Yuen, 2009.[17] Thus, the final result that associated with the above two implications may be the development of human-like artificial intelligent machine.

In particular, our planets can be viewed as the body since they are governed up by nature rules. While for our universe, the observations give us the clue of the natural constants. This event is the source of intuitionism or our mind. Finally, whenever there is a contradiction to these constants during computations, the universe's spirit – consciousness or the logicism shows itself.

Actually, a person's lifestyle has a close relationship with philosophy. One of the cases is a healthy philosophy (somehow a person's belief) of dieticians which may lead to a healthy lifestyle. This event will then give a satisfaction in his/her life and is linked with a group of person or a social class's consciousness. I note that one may

apply the cluster analysis (Yuen *et al.*, 2016)[18] to the qualitative data for the classification of them with similar characteristics or patterns. This implies the possibility of modeling different social classes' consciousness. If we model our human's conscious, the result is one may have an in-depth understanding of the integrated information theory or even the generalized one. The outcome is an advancement in the field of quantum physics, computer science, and mathematics. One of the interesting application is the feasibility to obtain a communication based model of consciousness or even to go ahead a step – the rationalization to the philosophical model of time consciousness (Mölder, 2014).[19] It is true that there are various types of consciousness model. It includes the functionalism or the philosophy of mind. The result of these consciousness models such as the architectural one may finally lead to the conclusive integrative theory of consciousness. One of the possible cases to model these human's consciousness can be referenced with the article Lam, March 2020 and will be left to the next cycles of investigation for those researchers interested.

CONCLUSION

The most significant aspect of this research is to investigate caries of prevalence in local kindergarten students and determine the effectiveness of fluoride treatment in preventing tooth decay. This event can be accomplished by verifying the proposed butterfly effect philosophy (during the prediction procedure), using the data collected from the research. The outcome of the study may better equip teachers to educate young children regarding the importance of using toothpaste with suitable fluoride content, or to undergo fluoride treatment for tooth decay. Education is the best method to prevent caries in kindergarteners since they are most likely to listen to a trusted adult (i.e. their teachers). Moreover, the collected data could be used to help predict behavioral patterns which cause caries (this event can be examined by hypothesis testing: Post-test and pre-test), thereby, preventing caries-causing practices. Finally, one may obtain the deserved Bayesian regression from the simulation of Bayesian inference, with estimated mean and variance that show normal distribution. Hence, a prediction model for caries development will be obtained. Indeed, with the use of the data (associated to different variables)

that collected from the research, one can establish the corresponding decision trees. If one can further compare these trees and select the best optimized path, the most feasible choice for a better decision-making will be applied. Then, one may develop and find the most suitable strategy for the regenerative medicine in teeth through our prediction model such as the need of subsidy for kindergarten tooth care – a machine generated policy.

Figure 12 depicts the flow of development for the research proposal as shown in the following diagram.

REFERENCES

1. Navarrete G, Mandel DR, editors. Improving Bayesian reasoning: What works and why? Lausanne Front Med 2016. Doi: 10.3389/978-2-88919-745-3.

2. Thierry M. Engineering Risk Management. Berlin, Munich: Walter de Gruyter GmbH; 2016.

3. Leung KT. Linear Algebra and Geometry. Hong Kong: Hong Kong University Press; 1974.

4. Lam S. Charging Our Electrical Devices in Anywhere and at any Time. Available from: https://www.ssrn.com/abstract=3326476 or http://dx.doi.org/10.2139/ssrn.3326476. [Last accessed on 2019 Jan 31].

5. Lam KS. Encouraging Statistical Learning Through Eastern and Western Mathematical History and Its Butterfly Effect Consequences. Available from: https://ssrn.com/abstract=3280934 [Last accessed on 2018 Nov 08].

6. Lam S. Predictive and Regenerative Medicine for Humans; 2019. Available from: https://www.ssrn.com/abstract=3510415 or http://dx.doi.org/10.2139/ssrn.3510415. [Last accessed on 2019 Dec 28].

7. Lam KS. Predictive and Regenerative Medicine for Humans. Available from: https://ssrn.com/abstract=3510415 [Last accessed on 2019 Dec 28].

8. Chen KJ, Gao SS, Duangthip D, Lo EC, Chu CH. Managing early childhood caries for young children in China. Healthcare (Basel) 2018;6:11.

9. Creswell JW. Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Upper Saddle, NJ: Pearson Education International; 2012.

10. Lam S. Evaluation of the Weather-Influenza Pattern with a Regression Model Approximation to Causality. Available from: https://www.ssrn.com/abstract=3563601orhttp://dx.doi.org/10.2139/ssrn.3563601. [Last accessed on 2020 Mar 29].

11. Seth AK. A MATLAB toolbox for Granger causal connectivity analysis. J Neurosci Methods 2010;186:262-73.

12. Ekholm O, Gundgaard J, Rasmussen NK, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. Scand J Public Health 2010;38:699-706.

13. Oni E. Public Policy Analysis. Concept Publications Ltd.; 2016.

14. Obi EA. Public Policy Analysis and Decision Making. CL Nwachukwu, AC Obiora; 2008.

15. Siu MK. The algorithmic and dialectic aspects in proofs and proving. In: Lin FL, Hsieh FJ, Hanna G, de Viller M, editors. Proceedings of ICMI Study 19 Conferernce: Proof and Proving in Mathematics Education. Vol. 2. Taipel: National Taiwan Normal University; 2009. p. 160-5.

16. Georg S. Fitting and Analysis Technique for Inconsistent Nuclear Data; 2018. Available from: https://www.arxiv.org/abs/1803.00960. [Last accessed on 2021 Dec 01].

17. Lau WW, Yuen AH. Exploring the effects of gender and learning styles on computer programming performance: implications for programming pedagogy. Br J Educ Technol 2009;40:696-712.

18. Yuen AH, Cheng JP, Chen MT. The Significance of Cultural Capital and Parental Mediation for Digital Inequity. Thousand Oaks, California: SAGE; 2016.

19. Mölder B. How philosophical models explain time consciousness. Proc Soc Behav Sci 2014;126:48-57.