Atlantic International University

A New Age for Distance Learning

Name：Muwanguzi Joshua Jjunju

Phone：256+776968958

Email：joshua2020muwanguzi@gmail.com

MODELING

Muwanguzi Joshua Jjunju
UB74035SCI83200

COURSE NAME:
MODELING

Assignment Title:
MODELING

ATLANTIC INTERNATIONAL UNIVERSITY
April/2022

Table of Contents

Introduction

In this course, the tasks faced by managers of civil engineering systems and by the systems engineers is the description of some system attributes such as the system structure, the way it works, or its condition or performance in the past, at the current time, or at some specified future time. Such tasks related to analysis and evaluation at any phase requires engineers to use tools including statistical and modeling tools.

The models in civil engineering include: a model to predict the future demand for urban transit, a model that describes the magnitude and direction of factors that enhance remediation of contaminated soil, a model that predicts the cost of reconstructing a levee 20 years from now, a model that describes the rate of corrosion in a structural steel member, and a model that describes the factors influencing the effectiveness of a building insulation system.

These various models in civil engineering can be further classified in several ways to include the system attribute that is being modeled, the mathematical form, the intended general purpose of the model (prescriptive versus descriptive), the level of empiricism (deterministic versus stochastic), the nature of the model (numerical vs. analytic, and the phase of civil engineering system development that is relevant to the model (such as demand models, cost models, deterioration models, operational policy effectiveness models, maintenance effectiveness models, and intervention decision models). Below is a focus on a specific type of analytic models: statistical regression models.

Body of Assignment

The statistical models are developed in more of an art procedure than a science. This is because there is no exact answer and the best

model is often obtained after several trials. It is often dependent on the experience of an analyst, agency
practices and culture, data availability and also the intended use of the model. The steps to develop a model include the following;

Step 1: Definition of Objective
The first step is always a definitive statement of what the model is meant to
describe or predict. This helps to prevent any communication problems between the data collectors and the modeler. It also helps the modeler to establish any a priori expectations of the model outcomes and capabilities.

Step 2 Sampling
This can also be called Data collection. The sample must be not only random (to avoid bias) but also a close replica of the population so that any inferences made from the data are applicable to the population.

Step 3 Specifying the Response Variable
This should reflect the objective in step1. It is an important step because it influences the mathematical form to be used at during the modeling process. The choice of the response variable is dictated by the characteristics of the units of measurement. The response variable can also be a survival probability or hazard probability depending on the likelihood that a system situation does not remain or remains up to a certain point in time.

Step 4 Selection of the Explanatory Variables
These are characteristics of the system or its environment that influence the response that is being to described using the statistical model. A simple model may involve only a few basic independent variables while more complicated models may include several independent variables. A number of explanatory variables are

grouped by the development phase at which the model is applied. This could be a continuous or a binary explanatory variable.

Step 5 Carry Out Preliminary Analysis of the Data

This step is to identify interesting trends or relationships between the dependent and independent variables. The tools used in this step include scatter diagrams, box plots, stem-and-leaf plots, pie charts, analysis of variance, pairwise t tests, among others.

Step 6 Model Specification

Model Specification is the most important step of the statistical model development process. Its various categories include the discrete versus continuous, linear versus nonlinear, single equation versus multiple equation, cross-sectional versus time series, and duration versus non-duration. The choice of functional form or mathematical form to be used depends on the type of response variable, the nature of the response data, and other considerations.

Step 7 Final Selection of Independent Variables

This involves dropping some of the excessive number of independent variables in order to simplify analysis. It is important for one to consider the plots in step 6 and analyse in order to find certain independent variables have little or no impact on the dependent variable and thus could be dropped without jeopardizing the efficacy of the model.

Step 8 Separate Your Data Set into Two

It is often prudent to break up the original modeling data set into two sets: one for the model calibration and a smaller portion for the model validation.

Step 9 Model Calibration

Calibration means determining the best function passing through the points. Mathematically, a “best function” could mean an equation that passes through the points such that the sum of the vertical deviations of various points from the regression line is minimized. Such a line is naturally the best unbiased and efficient line that passes through the points.

Step 10 Model Evaluation
It is under this step that the following tests could be used in order to ascertain how good the developed model is: the coefficient of determination (R2), the level of significance, the standard error (or t-statistics or p values) of the estimate, the heteroscedasticity of variance, and normality tests.

Step 11 Model Validation
The validity of a model is adjudged on how closely it fits to empirical observations and how well it extrapolates to situations or data other than those originally used in the model. A technique that is able to substitute the values of the independent variables from a validation data set into the calibrated model and determine the corresponding predictions of the response variable should be used. This allows the values to be compared with the actual observed values of the response for those independent variables.

During the development of a model, Errors in Systems Modeling and Suggested Precautions can be faced. These Errors occur due to uncertainties in the civil engineering systems management environment, such as material imperfections, variability in workmanship, climate/weather variations, economic uncertainties, equipment error, and human error or incompetence.

Model error could also take place due to misspecifications for example, omitting some key factors. A model is not truly complete until it adequately incorporates all relevant factors as well as the interactions between them. The model development process could be simplified through making appropriate simplifying assumptions, omitting certain duplicate factors, or considering only the aggregated effects of certain factors.

There are also some other important issues considered during Modeling. Among which is Addressing the Problem of Statistical Outliers. This outlier is an observation whose position is significantly distant from that of most observations in a given sample. It could be indicative of faulty data where the observations are simulated from a theoretical relationship which may not be valid under certain circumstances. An outlier could also be due to changes in system behavior or its natural or anthropogenic environment, human error, or instrument error. The outlier detection is useful in systems management because it can identify and draw attention to unexpected or overlooked system deficiencies or fraud before they develop into a state where they could have consequences that threaten the system, its users, or the environment. The detection can be carried out through identifying the outliers and considering the outlier In-laws or Outlaws which involve Retention and Deletion.

I have been able to learn from the Latin phrase cum hoc ergo propter hoc which is interpreted as “after this, therefore because of this.” The fact that two variables are correlated could not imply that one is a cause of the other and the conclusion that an event Y must have been caused by event X is a logical fallacy because there could be one of at least three counter explanations why they are highly correlated. In the first place, it could be a Reverse Causation where by correlation exists in the opposite direction. It could also be an Effect of the Lurking Variable where there may be a third unknown factor, Z (called a lurking or common-causal variable), that may actually be influencing both X and Y. By Coincidence, the seeming causation of one event by another may be purely coincidental.

Conclusion

The purpose of all models is to increase our understanding of the civil engineering systems. The so many important issues considered in making a Model also lay emphasis on the Limitations on the Response Variable, Exogenous versus Endogenous Variables and the Dangers of Extrapolation. These are less detailed in the above essay but their consideration should not miss in a true model.

Bibliography

Introduction to Civil Engineering Systems

https://students.aiu.edu/student/AIUWebinar.aspx?p=1&id=356
https://www.youtube.com/watch?v=I-IEWabgf0I&t=19s