An Application of Logistic Regression for the Efficiency Improvement of Bored Pile Casting in Wet Process System of Large Scale Construction Project of Sample Company

Thawatchai Thampetrarak¹ and Arnat Watanasungsuit²

Abstract

This paper presents the application of Logistic Regression Analysis to determine on the factors affecting the quality of the bored pile above pile cut-off level of large scale construction project. The 465 sets of quality data collected form sample construction project of which the response variables were dichotomous could not appropriately be analyzed by the Design of Experiment method. Results showed that factors of worker preparing bentonite solution and supervisor had the significance to the quality of bored pile at p-value 0.033 and 0.006 respectively and at goodness-of-fits test by Pearson’s method of 0.98. Finally the training for worker and supervisor including the establishment of work standard of bentonite solution preparation were proposed to the project manager in order to reduce bored pile waste.

Keywords: Logistic Regression, bored pile in wet system, waste reduction

¹ The student of Master of Engineering (Engineering Management) Program, Industrial Engineering and Management Department, Faculty of Engineering and Industrial Technology, Silpakom University, Nakompathom 73000, Thailand.
² The associate professor of Industrial Engineering and Management Department, Faculty of Engineering and Industrial Technology, Silpakom University, Nakompathom 73000, Thailand. Corresponding author: amat@gtmthailand.com
1. Introduction

Currently, the applied statistics are useful investigation tools in engineering applications in order to determine the significant parameters, especially, affecting the quality of products and services. Design of Experiments (DoE) is the vital methodology. Well-designed experiments can often lead to a model of system performance; such experimentally determined models are called empirical models [1]. Many of researchers used to apply the design of experiments to their study, such as; K. Boonsotsornatham et al. [2] have studied the optimization of blowing parameters in plastic bottle injection process. The number of round per minute (rpm) of injected screw, the temperature of injection chamber, and the temperature of die were the independents parameters, and, the weight of injected bottle was dependent parameters. The Analysis of Variance (ANOVA) was a main method of suitable level parameters investigation. H. Singh [3] has studied to acknowledge an optimal setting of turning process parameters. The cutting speed, feed and depth of cut were the independent parameters, and, the tool life was the dependent parameters. The Taguchi’s design of experiments method was used as major investigated tool. Finally, E.-S. Lee et al. have proposed the new technology of an ultra-precision grinding system and process for the aspheric surface micro-lens. The optimization of grinding conditions, moreover, with respect to ground surface roughness and profile accuracy is investigated by design of experiments.

As discussed above, the same character of these researches is the independent parameters are continuous variables or discontinuous variables, while the dependent parameters are continuous parameters. Moreover, the data has gotten from the designed experiments. This is difficult to our study. The Bore Pile casting technology has developed to wet process for the large of construction project. As the past observed data, the properties of bentonite solution, pile diameter, operator, and supervisor were the suspect to be significant parameters that affect the quality of the bored pile. So, the design of experiment which was used in past engineering research is not usable because the response variables are dichotomous. The regression analysis is our next statistical method. The linear regression analysis, however, was not appropriate to this study, because, the response variable or dependent parameters of this analysis are the continuous variable. The Logistic Regression analysis is an interested method.

The Logistic Regression is a parametric model to analyze dichotomous dependent variable problems; it has been widely applied in medicine and biology areas, but seldom used in engineering application [5]. Accordingly, this study was proposed an application of logistic regression analysis to investigate the significant parameters which affect the quality of bored pile casting in wet process system. Its response variable was categorized in dichotomous dependent variable. The significant parameters, then, were analyzed and corrected to explore the optimum results.

2. Logistic Regression

Logistic Regression approach can be explained as follows:

Let \( Y \) be a dependent variable, where \( Y \) is typically code as 1 or 0 for its two possible categories. Let \( X_1, X_2, ..., X_k \) be an independent variables (predictor). The Logistic model describes the expected value of \( Y \) in terms of the following “Logistic” formula [15]:

\[
\pi(X) = \frac{\exp\left( \beta_0 + \sum_{i=1}^{k} \beta_i X_i \right)}{1 + \exp\left( \beta_0 + \sum_{i=1}^{k} \beta_i X_i \right)}
\]

(1)

where, \( \pi(X) = E(Y \mid X) \). Logistic Regression was assumed that \( \pi(X) \) is related to \( X \) by the logit function. It is easy to show that:

\[
\frac{\pi(X)}{1-\pi(X)} = \exp\left( \beta_0 + \sum_{i=1}^{k} \beta_i X_i \right)
\]

(2)

The quantity \( \exp (\beta_0 + \sum \beta X) \) on the right hand side of (2) was called the odds ratio. Note that, the (2) was not linear in the explanatory variable \( X \). So, Logistic (or logit) transformation was deployed to be (3).

\[
\text{logit}(\pi) = \log\left( \frac{\pi(X)}{1-\pi(X)} \right) = \beta_0 + \sum_{i=1}^{k} \beta_i X_i
\]

(3)

The equation is linear function in \( X \). It could also be written as [16]:

\[
\text{logit}(\pi) = \beta X
\]

(4)

Where,

\[
\pi = \begin{bmatrix}
\pi_1 \\
\pi_2 \\
\vdots \\
\pi_l
\end{bmatrix}, \quad \beta = \begin{bmatrix}
\beta_0 \\
\beta_1 \\
\vdots \\
\beta_k
\end{bmatrix}, \quad \text{and} \quad X = \begin{bmatrix}
1 & X_{11} & X_{12} & \cdots & X_{1l} \\
1 & X_{21} & X_{22} & \cdots & X_{2l} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & X_l & X_{2l} & \cdots & X_{ll}
\end{bmatrix}
\]
2.1 Parameters estimation

The parameters of the model are estimated by the maximum likelihood method (ML) or may be obtained by using a generalized weighted least square procedure. Although the first method was widely used to its general applicability, it uses asymptotic estimation because exact properties of estimators are difficult to be obtained, thus resulting in biased estimators for small sample data [6]. However, the number of data which observed from the large size project of the sample company was 465 set. So, The ML is appropriate for this study. The detail of ML was being shown in [7], and [8].

The estimated standard errors of estimated coefficients were denoted as:

\[ SE(\hat{\beta}_i) = \left[ \hat{\beta}^2(\hat{\beta}_i) \right]^{1/2}, \quad i = 1, 2, \ldots, k \]  

(5)

2.2 Variables selection

The test for the inclusion or exclusion of explanatory variables used can be achieved by the comparison of one model with another. One of the two models under comparison is the full model (where at the initiate stage, it contains all of the available variables) and another one is a current model (a variable may subtracted) [9]. This method was namely the stepwise selection. It can be used to identify the best subset of variables. The difference of the deviation between two models has \( \chi^2 \) distributions; so \( \chi^2 \) test could be used.

2.2 Goodness-of-fit

A goodness-of-fit measure method was analyzed by Pearson hypothesis test [10], it was given by,

\[ \chi^2 = \sum_{i=1}^{n} \left( \frac{y_i - \hat{y}_i}{\text{Var}_i} \right)^2 \]  

(7)

where, \( y_i \) is the observed response and \( \hat{y}_i \) is the predicted one for the \( i^{th} \) subject, and \( \text{Var}_i \) is the estimated variance of the response. Equation (7) also has an approximate \( \chi^2 \) distribution. Its degree of freedom is the number of covariate patterns minus the number of estimated parameters.

3. Bored pile casting parameters

The suspect dependent parameters were bored pile diameter (mm.), the characteristic of bentonite solution (solution viscosity, solution density, and PH of solution), operator, and supervisor. The characters were shown in table 1

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>Bored pile diameter</td>
<td>( 800 &lt; X_1 &lt; 1200 ) mm.</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>Solution viscosity</td>
<td>( 29 - 50 ) sec.</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>Solution density</td>
<td>( 1.15 - 1.10 ) g/ml.</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>PH of solution</td>
<td>( 7 - 12 )</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>Operator</td>
<td>A, B, C</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>Supervisor</td>
<td>D, E, F</td>
</tr>
</tbody>
</table>

Remark: The solution viscosity was tested by Mush Cone test method. The solution density was tested by Mud Density Balance test method. The solution PH was tested by PH indicator (paper strips).

4. Results

The analysis was conducted on the statistical software package “Minitab” with the observed data of 465 set from the project A from the sample company. Due to its company propriety, detailed data could not be described here. The Logistic Regression analysis results could be expressed below.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Z</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>0.719611</td>
<td>7.29705</td>
<td>0.10</td>
<td>0.921</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>0.0001319</td>
<td>0.0011688</td>
<td>0.11</td>
<td>0.910</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0.0275114</td>
<td>0.0498675</td>
<td>0.55</td>
<td>0.581</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>-0.916761</td>
<td>5.33020</td>
<td>-0.17</td>
<td>0.863</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>-0.436680</td>
<td>0.381046</td>
<td>-1.15</td>
<td>0.252</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>-0.521054</td>
<td>0.245004</td>
<td>-2.13</td>
<td>0.033</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>1.09319</td>
<td>0.396101</td>
<td>2.76</td>
<td>0.006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Coefficient</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval Lower</th>
<th>95% Confidence Interval Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>0.719611</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>0.0001319</td>
<td>1.03</td>
<td>0.93</td>
<td>1.13</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0.0275114</td>
<td>0.40</td>
<td>0.00</td>
<td>13772.60</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>-0.916761</td>
<td>0.65</td>
<td>0.31</td>
<td>1.36</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>-0.436680</td>
<td>0.59</td>
<td>0.37</td>
<td>0.96</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>-0.521054</td>
<td>2.98</td>
<td>1.37</td>
<td>6.49</td>
</tr>
</tbody>
</table>
The goodness-of-fit test was expressed as; Pearson method, 
\[ P(\chi^2 > 149.010) = 0.984 \text{ at DF} = 188. \] This shows that the logistic regression model provides a good fit.

4. Discussions

Obviously, the operators and supervisors were significant parameters in bored pile casting in wet process. On the other hand the tolerance of solution characteristic has not affected the quality of bored pile. The operators and supervisors, then, were analyzed in the detail, such as, education, experience, training record etc. The engineering management technique was applied to this project to propose the corrective action.

5. Conclusions

This study was proposed the Logistic Regression analysis to determine the significant parameters affecting the quality of bored pile casting in wet process of the large scale construction project. It was found that factors of worker preparing bentonite solution and supervisor had the significance to the quality of bored pile at p-value 0.033 and 0.006 respectively and at goodness-of-fits test by Pearson’s method of 0.98. After the implementation of the training for worker and supervisor including the establishment of work standard of bentonie solution preparation, the efficiency improvement of Bored Pile Casting was about 10% and every party was satisfied with the proposed corrective action plan.

6. Acknowledgment

The researchers would like to express the gratefulness to the employees of the sample company for their kind collaboration.

References