DERIVING ACCURATE COST CONTINGENCY ESTIMATE FOR MULTIPLE PROJECT MANAGEMENT

Jin-Lee Kim [1] and Ok-Kyue Kim [2]

[1] Ph.D. Candidate, Department of Civil and Coastal Engineering, University of Florida, Gainesville, FL, USA
[2] Associate Professor, Department of Architectural Engineering, Chungbuk National University, Cheongju, Korea
Correspond to jin5176@ufl.edu

ABSTRACT: This paper presents the results of a statistical analysis using historical data of cost contingency. As a result, a model that predicts and estimates an accurate cost contingency value using the least squares estimation method was developed. Data such as original contract amounts, estimated contingency amounts set by maximum funding limits, and actual contingency amounts, were collected and used for model development. The more effective prediction model was selected from the two developed models based on its prediction capability. The model would help guide project managers making financial decisions when the determination of the cost contingency amounts for multiple projects is necessary.

Key words : Cost contingency, Project Management, Estimate, Statistical Methods

1. INTRODUCTION

A cost contingency in cost estimates for construction projects is an amount of money allocated for unexpected events that may affect project costs. When adding or subtracting a certain amount of value from the base estimate to reflect an equal chance of underrun or overrun of cost, project planners and managers make an assessment of the amount of funds to balance the relative stability of the scope and assumptions based on the baseline estimate [1]. When allocating project funds into multiple projects, they usually include a contingency amount in their cost estimates in order to stabilize the programmed estimate. For this reason, contingency accounts are significant elements, especially in a multiple project situation since any saving can contribute to the profit. A contingency fund also plays an important role in budget planning because it keeps the overall cost for a project at a constant value. Without the contingency funds, the overall cost for a project will increase and project planners and managers will need to revise their budget planning to accommodate the increase due to limited funding.

Currently, one of the most simple and common methods is to add a percentage of the estimated cost, for example 5-10%, based on previous experience with similar projects [2]. Such an approach can induce project managers to make erroneous financial decisions due to uncertainty. In addition, it is difficult to determine a cost contingency amount based on subjective judgments of project managers because the amount depends on various factors, such as organization, type of construction, amount of completion, and so on. However, historical data of cost contingency need to be considered in the development of a cost estimate system for any organization that deals with multiple projects simultaneously. This requires researchers to develop a model to predict future cost contingency values and estimate a mean response.

In order to estimate total project costs more realistically and sufficiently and to cover any cost occurred by uncertainties, contingencies should be allocated effectively using either statistical methods based on historical data or best judgment based on the stage of project development. For this reason, an organization with multiple projects needs to develop a guideline that provides for assigning a contingency amount into a cost estimate system. However, little work exists on how to model historical data of cost contingency to predict future cost contingency value and estimate mean response in spite of the recommendation of probabilistic approaches. Therefore, there is a need to find a suitable prediction model using statistical methods based on historical data in order to help managers make financial decisions effectively.

This paper examines the possibility for the use of a prediction model that can be utilized by a project manager or estimator in an organization that deals with multiple projects. The goal is to find a suitable model to predict cost contingency amounts as a function of original contract amounts. Literature and current contingency practices were reviewed to identify possible factors that affect the determination of cost contingency. A prediction model using the least squares estimation method was developed based on historical data of cost contingency.

2. CURRENT CONTINGENCY PRACTICES

In cost estimate, the most common method for dealing with uncertainty is to use a contingency amount. Contingency aims to make cost estimates realistic and sufficient to allow for any cost incurred by risks and uncertainty. Several definitions can be documented and used to meet the requirements set by an organization. The following are definitions of contingency.

- A contingency is an amount of money added to an estimate in order to allow for unpredicted changes that will likely be required during construction. This amount may be derived either through statistical analysis of past project costs or by applying experience from similar projects. However, cost contingency does not usually include extreme changes in scope or unforeseeable events such as a strike, earthquake, etc. [3].
- A contingency reserve is a quantity that allows for future situations planned for risks from known unknown conditions. The Construction Industry Institute [4] defines risks from known unknown conditions as foreseeable and possible risks that may include severe weather, unusual difficulty with a client, extreme adverse labor activity, sudden labor shortages due to new and competing work activity in the project area, and commodity shortages due to embargos and regulatory interventions. For instance, rework is certainly known, but the amount of rework is not certainly known. Contingency reserves include cost, schedule, or both. They are considered to reduce the impact of projects that missed the objective of cost or schedule [5].
- A contingency is an amount budgeted to cover costs that result from incomplete design, unforeseen and unpredictable conditions, or uncertainties. It is determined based on the status of design, procurement, and construction and the complexity and uncertainty of the project. However, contingency is not a countermeasure intended to avoid making an assessment of expected costs [6].

The cost contingency is generally calculated as a percentage of the original contract amount, as accepted in many cost estimates. Florida Department of Transportation (Hereinafter Florida DOT), for example, has established the contingency supplemental agreement and initial contingency amount pay item. Field supplemental agreements are funded through an initial contingency amount pay item, while work orders are funded through a contingency supplemental agreement. According to the Florida DOT's construction project administration manual, maximum fund limits for these items have been established based on the original contract amount [7].

The contingency amount becomes a single line item amount in cost estimates even though all significant risk elements are considered. The Construction Industry Institute [4] pointed out two disadvantages of managing contingency as a single account. One is that there is a tendency to cut down the amount on a basis of a "first-come, first-served," which results in a delay of corrective actions. The other is that the project manager is fully responsible for managing and controlling every account. These heavy responsibilities cause contingency management to be neglected because the project manager is the only person who manages a project-wide account. Thomson and Perry [2] also indicated several disadvantages with regard to using a contingency amount as follows: (1) the percentage figure is not appropriate for the specific project, (2) the percentage added indicates the potential for detrimental or downside risk, (3) the percentage tends to disregard time, performance, and quality risks, and (4) the percentage may become routine.

3. MODEL DEVELOPMENT

3.1 Factors Affecting Cost Contingency

Project managers who allocate contingency funds into multiple projects are interested in how to determine a cost contingency for a specific project because the total contingency amount cannot exceed beyond certain funding limits within overall budgets during a certain time period. The total contingency amount, Y, for a certain time period, such as a fiscal year, can be defined as the summation of contingency for each project. The total contingency amount is also expressed by the summation of the product of the original contract amount and a contingency ratio, as shown in Eq. (1)

$$Y = \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \beta_i x_i \tag{1}$$

where, x is a random variable denoting the original contract amount of a certain project and $x = 1, 2, \ldots n$, β is the contingency ratio of a certain project and is another random variable that is of interest, which is generally equal to or grater than 0 %, and i is a project index and $i = 1, 2, \ldots n$.

It is a convenient practice to determine contingency as a percentage of the original budget of a project. The research team assumed that all of the historical data obtained for this study were determined based on this practice, as previously mentioned. Even though cost contingency amounts are determined based on various factors, such as organization, type of construction, amount of completion, and so on, the only factors considered in this paper are both the original contract amounts and the actual contingency amounts. The estimated contingency amount was considered a factor that affects the determination of the future contingency amounts. The estimated contingency amounts were calculated according to maximum fund limits established by Florida DOT. Maximum funding limits are shown in Table 1.

Since cost contingencies for future similar projects can be predicted based on known costs for previously constructed projects, historical data of cost contingency must be modeled automatically to answer "what if" scenarios defined by the user under multiple project situations. Therefore, a prediction model for cost contingency was developed using the least squares estimation method. One of the purposes of the least squares analysis is to predict future responses and estimate mean responses [8].

Table 1. Maximum funding limits for FDOT projects [7]

Original contract amount	Maximum limit options		Decision
	A	B	
$ 5,000,000 or less	5 % of the original contract amount	$ 50,000	Whichever is less
More than $ 5,000,000	1 % of the original contract amount	$ 150,000	

3.2 Data Collection and Splitting

Historical data for cost contingency are available from the Florida DOT database [9]. Florida DOT posted detailed data of cost and time for transportation projects by district. Data, 79 points, were extracted from the cost and time reports, ranging from 2003 to 2004. In addition, data, 53 points, were extracted from that report, ranging from 2004 to 2005. Two important variables, the original contract amount and the actual cost contingency by project, were extracted and obtained for this study. Both the original contract amount for a project and the estimated contingency amount set by maximum funding limits for a project were the only proposed factors to develop a prediction model based on the current practice, as previously described.

Splitting the data set is necessary to obtain a model and to validate the model in the case where it is impractical to obtain an adequate independent data set with which to validate a model [8]. This research did not split data because two different sets of historical data were available. The first set was used for model development and the other set was used for validation. Table 2 shows descriptive statistics of a set of historical data of cost contingency used for model development. Tr Mean stands for the 90% trimmed mean and StDev stands for the sample standard deviation.

Table 2. Descriptive statistics of cost contingency amounts

Variable	N	Mean	Median	Tr Mean
Cost contingency	79	49565	50000	47775
Variable	StDev	Min	Max	
Cost contingency	27818	5000	150000	

3.3 Linear Regression Models

Two different types of linear models were developed to predict future cost contingency, depending on the number of independent variables. One is ***Model 1*** that involves only one independent variable, which is the original contract amount, and the other is ***Model 2*** that involves two independent variables, which are the original contract amount and the estimated cost contingency amount based on funding limits, as previously described. These simple linear models state that the true mean of cost contingency changes at a constant rate as the value of the independent variable, which is either the original contract amount for Model 1 or the original contract amount and the estimated cost contingency amount for Model 2, increases or decreases.

Let us first consider Model 1 that contains only one variable, which is the original contract amount. The function relationship between the true mean of Y_i, cost contingency of a project, and X_i, the original contract amount, is the equation of a straight line, as shown in Eq. 2.

$$E(Yi) = \beta_0 + \beta_1 X_i \tag{2}$$

The observations of the dependent variable Y_i are assumed to be random observations from populations of random variables with the mean of each population given by the true mean of Y_i. The deviation of an observation Y_i from its population mean $E(Y_i)$ can be considered by adding a random error ε_i in order to give the statistical model, as shown in Eq. 3.

$$Yi = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{3}$$

The subscript i indicates the project index, $i = 1, 2, ..., n$. The X_i are cost contingency amounts per project and are assumed to be measured without error. That is, cost contingency amounts are assumed to be a set of known constants for this study. The Y_i and X_i are paired observations and they are measured on every project index.

The random errors ε_i have zero mean and are assumed to have common variance σ^2 and to be pair-wise independent. Since the only random element in the model is ε_i, these assumptions imply that the Y_i also have common variance σ^2 and are pair-wise independent. The random errors are also assumed to be normally distributed, which means that the Y_i are normally distributed, as expressed below.

$$\varepsilon_i \sim NID\ (0, \sigma^2)$$

The Pearson product moment correlation coefficient was calculated between the actual contingency amounts and the original contract amounts. The coefficient of correlation, r = 0.852, confirms the positive linear relationship between two variables.

4. ANALYSIS OF FINDINGS

4.1 Two Regression Models

Under these assumptions, two prediction models with a different number of independent variables were developed as candidates in order to select a better model. The same assumptions as Model 1 were made for developing Model 2. Historical data of cost contingency were performed using Minitab software program [10]. Model 1 is written in Eq. 4.

$$\textbf{Model 1}: \hat{Y} i = 28822 + 0.00735 \hat{X}_1 \tag{4}$$

where, $\hat{Y}_i$ is an estimate of the response, cost contingency and $\hat{X}_i$ is an estimate of the independent variable, the original contract amount. The equation is also expressed as *Cost Contingency = 28822 + 0.00735 Estimated Contingency*. The *y*-intercept, 28822, indicates the point at which the line crosses the *y*-axis. The slope, 0.00735, is the estimated change in the mean value of cost contingency for every original contract amount. Model 2 is also written in Eq. 5.

Model 2:

$$\hat{Y}i = 5017 + 0.88\hat{X}_1 + 0.00141\hat{X}_2 \quad (5)$$

where, $\hat{Y}_i$ is an estimate of the response, cost contingency and $\hat{X}_i$ and $\hat{X}_2$ are estimates of the independent variables, the estimated contingency amount and the original contract amounts, respectively. The equation is also expressed as *Cost Contingency = 5017 + 0.880 Estimated Contingency Amounts+0.00141 Original Contract Amounts*. The *y*-intercept, 5017, indicates the point at which the line crosses the *y*-axis. The slopes, 0.880 and 0.00141, are the estimated changes in the mean value of cost contingency for every estimated contingency amount and original contract amount. Table 3 shows the results of parameter estimates for two models. The coefficients of determination R^2 are 72.5 % and 84.3 % for Model 1 and Model 2, respectively (See Table 5). The high coefficient of determination means the dependant variable is well explained by the independent variables in the two models.

Table 3. Regression table for Model 1 and Model 2

Predictor	Coefficient	Standard Error	T-value	P-value
(1) Model 1: Cost Contingency = 28822 + 0.00735 Original Contract Amounts				
Intercept	28822	2202	13.09	0.000
Original Contract Amounts	0.0073529	0.0005162	14.25	0.000
(2) Model 2: Cost Contingency = 5017 + 0.880 Estimated Contingency Amounts + 0.00141 Original Contract Amounts				
Intercept	5017	3577	1.40	0.165
Estimated Contingency	0.8796	0.1167	7.53	0.000
Original Contract Amounts	0.0014108	0.0008812	1.60	0.114

Table 4 shows the result of the analysis of variance. The hypothesis that the true value of the linear regression coefficient and the slope is zero is of interest. The null hypothesis is written as H_o: $\beta_1 = m$, where m is any constant of interest and can be equal to zero. The alternative hypothesis is H_a: $\beta_1 \neq m$. The *F*-statistic was used for tests of significance. Using the two-tailed alternative hypothesis and α-value (0.05), a critical value of $F_{(0.05,1,77)}$ = 3.92 was compared against F-value to determine whether *MS(Regression)* is sufficiently larger than *MS(Residual)*. Since *F* = 202.94 is greater than 3.92, the conclusion that the data provide conclusive evidence of a linear effect of cost contingency was made for Model 1. In the same manner as Model 1, the same conclusion was made for Model 2 because *F* = 203.35 is greater than $F_{(.05,2,76)}$ = 3.07.

Table 4. Analysis of Variance for Model 1 and Model 2

Source	Degree of Freedom	Sum of Squares	Mean Squares	F-value	P-value
(1) Model 1: Cost Contingency = 28822 + 0.00735 Original Contract Amounts					
Regression	1	43758408557	43758408557	202.94	0.000
Residual Error	77	16602946943	215622688		
Total	78	60361355500			
(2) Model 2: Cost Contingency = 5017 + 0.880 Estimated Contingency Amounts + 0.00141 Original Contract Amounts					
Regression	2	50857513633	25428756816	203.35	0.000
Residual Error	76	9503841867	125050551		
Total	78	60361355500			

The estimated contingency amounts that were established by maximum funding limits, the actual contingency amounts that were actually used for a project, and the point estimate amounts that were calculated using Model 2 developed in this study were compared in Figure 1. The result shows that the point estimate amounts using Model 2 are much closer than the estimated contingency amounts to the actual contingency amounts.

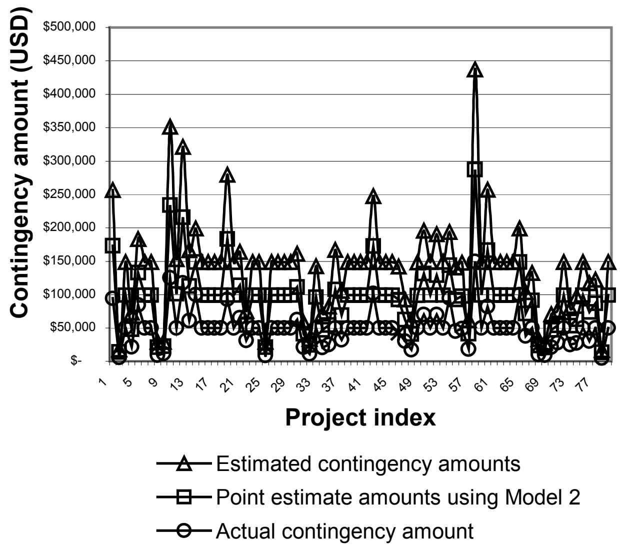

Figure 1. Comparison of the estimated, actual, and point-estimated contingency amounts

4.2 Better Prediction Model Selection

A better prediction model has to be selected to increase the effectiveness of the model as a tool. Two models were evaluated based on R^2, $R^2_{(adj.)}$, PRESS, and S (Standard deviation). The coefficient of determination, R^2, is simply the square of the correlation coefficient and the percentage of variance between data and their regression model. The adjusted coefficient of determination, $R^2_{(adj.)}$, is more effective than R^2 over models involving different numbers of parameters because $R^2_{(adj.)}$ need not always increase as variables are added to the model, as opposed to R^2. $R^2_{(adj.)}$ also tends to stabilize around some upper limit as variables are added [8]. The PRESS statistic, the sum of squares of these discrepancies was calculated to detect inadequacies in the model. The PRESS value is a measure to select the best model that represents data [11]. Table 5 shows the results to be used for a better regression analysis. The better model of the two was selected based on the values of $R^2_{(adj.)}$, S, and PRESS. Criteria for judgment are a high $R^2_{(adj.)}$, a low S value, and a low PRESS value. Therefore, Model 2 is more effective than Model 1 and the research team concludes that Model 2 has the predictive capability.

Table 5. Results for better regression analysis

Model	No. of variables	R^2	$R^2_{(adj.)}$	S	PRESS
1	1	72.5	72.1	14684.1	17893183545
2	2	84.3	83.8	11182.6	10253885730

4.3 Model Validation

Model 2 was selected as the more effective one from the two developed models based on its prediction capability. For the purpose of validation, the data set that has 53 data points was used to perform the models' prediction capabilities. Validation of the model using an independent sampling of the population is more desirable than the use of estimates of mean squared error of prediction (MSEP) using the original sample data [8]. The model with lower MSEP was considered a better model in predicting the dependent variable.

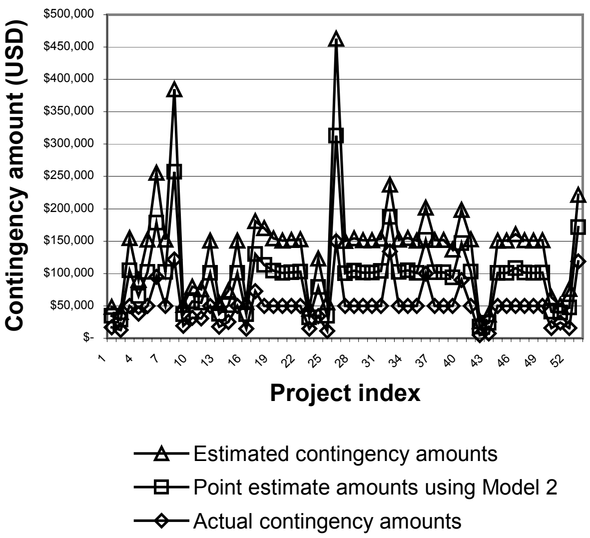

Figure 2. Comparison of the estimated, actual, and point-estimated contingency amounts

Model 2 meets the previous criteria and has more predictable capability than Model 1, proving its superiority. In addition, using a different data set, a comparison was made in Figure 2. The result shows that the point estimate amounts using Model 2 indicate closer estimate amounts than estimated contingency amounts that were established by the maximum funding limits, meaning that the result in Figure 2 is same as the result in Figure 1.

4.4 Limits of Study

Factors that affect the determination of cost contingency amounts need to be identified more accurately. The research team only considers the original contract amount and the estimated contingency amounts. These two parameters are not sufficient to develop a more effective model to predict and estimate cost contingency. Multicollinearity among variables needs to be considered for future study.

5. CONCLUSIONS AND RECOMMENDATIONS

The research conducted in this paper was intended to develop a model that predicts cost contingency amounts for an organization that deals with multiple projects. The original contract amounts, the estimated contingency amounts set by maximum funding limits, and the actual contingency amounts were collected and used for model development. The more effective prediction model was selected from the two developed models based on its prediction capability. The model provides a good indication of how to determine cost contingency amounts for future projects as a function of the original contract amounts and the estimated contingency amounts set by maximum funding limits. Therefore, the model would help guide project managers making financial decisions when the determination of the cost contingency amounts for multiple projects is necessary. It is important to note that the model can be used for guidance but an individual model for any specific organization should be developed.

REFERENCES

[1] Lorance R. B., and Wendling, R. V. (1999). "Basic Techniques for Analyzing and Presentation of Cost Risk Analysis," *AACE International Transactions*.

[2] Thompson, A., and Perry, J. G. (1992). *Engineering Construction Risks: A Guide to Project Risk Analysis and Risk Management*. Thomas Telford, London.

[3] Association for the Advancement of Cost Engineering (AACE). (1990). *Standard Cost Engineering Terminology, Cost Engineering Notebook*.

[4] Construction Industry Institute (CII). (1989). *Management of Project Risks and Uncertainties*, Publication 6-8, Austin, Tex.

[5] Project Management Institute (PMI). (2002). *Project Management Body of Knowledge (PMBOK)*, Upper Darby, PA.

[6] US DOE (1994). *Cost Guide*, Washington D.C.

[7] Florida Department of Transportation (FDOT). (2002). *Construction Project Administration Manual*, FL.

[8] Rawlings J. O., Pantula, S. G., and Dickey, D. A. (1998).

Applied Regression Analysis: A Research Tool. 2nd Ed., Stringer-Verlag New York, Inc., New York.

[9] Florida Department of Transportation (FDOT). *Construction office cost and time report*, Retrieved from http://www.dot.state.fl.us/construction/CONSTADM/reports/cost&timeNEW/construction_office_report.htm on May 4, 2005.

[10] MINITAB (2000). *MINITAB StatGuide*, MINITAB, Inc., State College, Pennsylvania.

[11] Draper, N., and Smith, H. (1998). *Applied Regression Analysis*. 3rd Ed., J. Wiley & Sons. NY.