COST GROWTH ESTIMATION IN LARGE SCALE PROJECTS USING INVERSE BAYES FORMULAE

Jeffrey H. Smith; Sol Price School of Public Policy, University of Southern California, Ralph and Goldy Lewis Hall, Los Angeles, CA 90089-0626, jeffs@usc.edu

ABSTRACT

The science and exploration missions undertaken by NASA are major project developments that have historically experienced cost growth due to technology challenges, schedule risks, and programmatic risks (funding stability), and other external events. The aim of this paper was to apply Inverse Bayes Formulae (IBF) in reverse from known cost growth at the end of a project to calibrate the prior distribution of cost growth that *should* have been used to estimate the total project cost at the beginning. Cost data were obtained from fifteen NASA space missions that included Earth orbiters, planetary probes to Mars and Jupiter, telescopes, and comet and asteroid flyby missions. Comparisons with alternative metrics are presented that show reduced cost estimation error due to IBF.

Keywords: Project Cost Growth, Bayes, IBF

Introduction

Assessing a cost growth factor is challenging from a number of perspectives [1]. The commonly accepted approach to constructing the initial cost estimate is to estimate the probability distributions of the elemental cost elements and use simulation methods to compute a probabilistic summation of the element costs to the total cost [2]. The resulting total typically has a normal distribution via the Central Limit Theorem. This feature is useful for the exercise presented herein since the Bayesian computations are facilitated by the conjugate properties of the normal distribution.

Approach

Historical cost growth (CG) factors are often used to adjust cost estimates for new projects. The basis for these calculations can be based on the actual (known) completion cost at the end of the project or, the initial estimate at the beginning of the project. The former are termed "historical" CG factors since they are computed after (*posteriori*) the project is completed. The latter is used when the actual completion cost is unknown during the development phase and the CG values evolve over time beginning with the initial cost estimate as the basis and subsequent estimates measured to this basis. This approach is termed the "forward" CG method since it is *a priori* in nature—the final outcome cost growth is unknown until the project is completed. Both of these views of cost growth are used for the Inverse Bayes Formulae (IBF) [3] technique used here. The method works in reverse from known cost growth at the end of a project to determine the prior distribution of a cost growth factor that *should* have been used at project start to estimate the total project cost. The first step was to use the mean and variances of the project data to calibrate normal distributions for the prior and posterior computations (forward and backward) using the properties of normal conjugates [4]. These formulae are used for the forward computation of the posterior cost growth using the initial cost as basis for the sample portion and the historical growth based on the actual cost outcome for the priors. The process is then reversed using the computed posteriors to start the process at project completion and, with the historical cost growths for the sample data, working backward toward project start.

The dataset consisted of 15 NASA projects that included Earth orbiters, planetary, space telescopes, and comet and asteroid missions. Because the focus of this study was on estimating project development cost growth, only development costs were analyzed. The data were extracted from the ONCE database available through NASA [5].

The cost growth mean and standard deviation at each milestone was computed as the sample estimate for the inverse Bayes formulae. The prior means and standard deviations were calculated via three iterations of IBF working in reverse from launch to initial project start. The cost estimates at project start in the dataset were then increased by average cost growth and IBF cost growth factors to produce forecasts at launch. The forecasts were compared to the actual costs using: mean square error, mean absolute deviation, mean absolute percent deviation, and "leave-one-out" for the lowest and highest development cost projects. The IBF cost estimates were significantly better than use of historic average cost growth factors.

Discussion and Observations

While the intent of the study was to explore the potential of IBF for cost growth estimation, the small number of projects in the study (15) limits to some extent, the generality of this application to other projects. Another limitation could be the variety of project types which raise questions about the homogeneity of the dataset and while the same basic project development processes were applied to each of these projects, the disparity in their objectives, scope, schedules, and costs could be an explanation why the variance and range of values of the average cost growths are large. It may be the IBF approach reduces the impact of large aggregate variance in the historical dataset through the tendency of Bayes Theorem to iterate to a lower posterior variance after combining the sample and prior variances.

The following observations were made:

- The IBF cost growth factor led to closer estimates of actual cost at project start than applying an historical average cost growth factor.
- While the IBF estimates were closer as a group and for many individual projects, the residuals for some projects were still large albeit less so than for the average CG approach.
- The conjugate properties of the normal distribution assumption facilitated the IBF calculations, however, the form and existence of the prior distributions is limited by the negative term in the mean and variance parameter calculations. If specific terms are large enough, the parameters are undefined (e.g., negative variance).

REFERENCES

[1] Arena, Mark V., R.S. Leonard, S.E. Murray, and O. Younossi, "Historical Cost Growth of Completed Weapon System Programs, Rand Technical Report No. TR-343, Rand Corporation, Santa Monica, CA, 2006, pp. 1-8.

[2] NASA Cost Estimating Handbook Version 4.0, Doc. No. NASA CEH v.4.0, National Aeronautics and Space Administration, Washington, D.C., February 2015.

[3] Ng, Kai Wang, and Howell Tong, "Inversion of Bayes Formula and Measures of Bayesian Information Gain and Pairwise Dependence," Research Report No. 477, University of Hong Kong Department of Statistics and Actuarial Science, Hong Kong, November 2010.

[4] Winkler, R.L., An Introduction to Bayesian Inference and Decision, Second Edition, Probabilistic Publishing, Gainesville, FL.

[5] Johnson, J.K., E. Plumer, M. Blandford, and J. McAfee, "One NASA Cost Engineering Database (ONCE), 2015 NASA Cost Symposium, Ames Research Center, Moffett Field, CA, August 25-27, 2015.