Further Study of the Normality of CPI and SPI(t)

By Walt Lipke
PMI® Oklahoma City Chapter

ABSTRACT – Understanding the statistical distribution of the periodic cost and schedule indexes from Earned Value Management (EVM) and Earned Schedule (ES) is critical to forecasting, planning, estimating, process modeling, and simulation. Within the last 20 years, two independent studies examined the normality of the periodic cost variance (CV). Thereafter a study, was performed examining the EVM and ES indexes as well as CV. This last study corroborated the finding of non-normality for CV and provided evidence that the logarithm of the EVM cost index and the ES schedule index periodic values are normally distributed. Using additional project data, the normality of the indexes is examined, once again.

Introduction

My search began in 2002 for a normal distribution representation for the EVM cost performance index (CPI)[a] and the schedule performance index (SPI(t))[b] from Earned Schedule. During the trial application of Statistical Process Control (SPC) unreasonable and inconsistent results were obtained. Believing the application of SPC would provide useful project control information, literature research was performed seeking a resolution to the problem. A reference providing guidance on the application of SPC indicated that a data transformation was needed to approximate normality [Pitt, 1995].

Further research indicated the use of the log-normal distribution might be appropriate. "Skewed distributions are particularly common when mean values are low, variances large, and values cannot be negative. Such skewed distributions often closely fit the log-normal distribution" [Stahel, et al, 2001]. Figure 1 provides a visual helpful to this discussion.

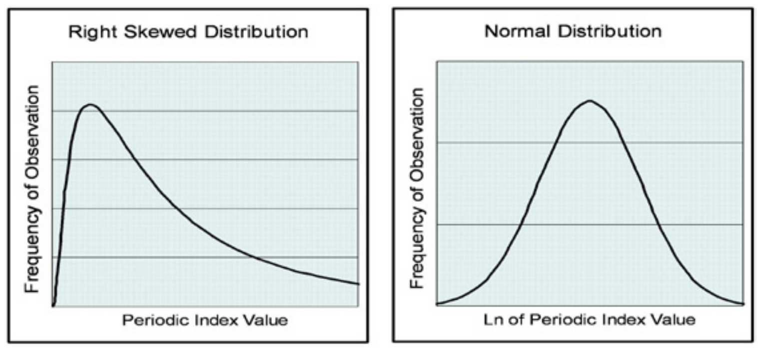


Figure 1. Right-Skewed & Normal Distributions

The characteristics of the periodic values for CPI and SPI(t) used in the SPC trial appeared to suit the application of the natural logarithm. The cumulative values for CPI and SPI(t) from several projects were generally near 1.0, the standard deviation varied from 0.2 to 0.7, and the index values observed were not less than zero.

From the problems in the trial application of SPC, I had arrived at the position that identical SPC signals[c] should be discovered independent from how the index is analyzed [Lipke, 2002 (March]. For example, identical results should be obtained from reciprocal data representations, such as CPI and CPI^{-1}. From that beginning point, the logarithm transformation needed to satisfy four behavior characteristics:

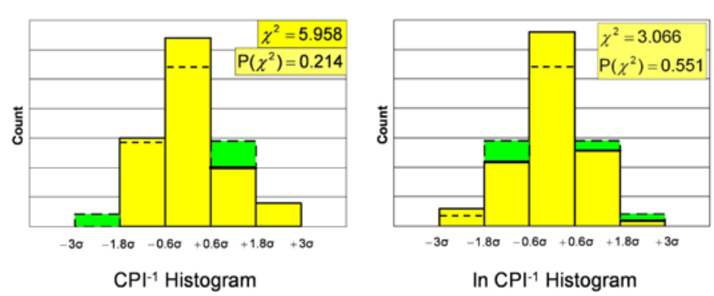
1. The antilogarithm values of the average of the logarithms for the inverted and non-inverted data representations must be reciprocals.
2. Values for the standard deviation[d] determined from the logarithms of the inverted and non-inverted data are equal.
3. SPC signals identified from the logarithms from the inverted data representation are found in the non-inverted, and vice-versa.
4. The logarithm values of the data for both inverted and non-inverted representations approximate a normal distribution.

The first three characteristics are completely and exactly satisfied using logarithms; the antilogarithm of the average values are reciprocals, the values for the standard deviation are equal, and the signals found are identical. Mathematically, these characteristics must occur from the application of the natural logarithm function.

For the fourth characteristic, it was seen from the creation of histograms that the logarithms of the indexes, in general, were more normally distributed than the indexes, themselves. An example of the histograms is illustrated in Figure 2. The yellow columns are counts of the number of observations occurring within five regions of equal size. Although the logarithm data representation visually appeared more normal, I could not affirmatively state that it satisfactorily approximated a normal distribution. Statistically based hypothesis testing of normality was needed to provide more conclusive evidence.

The subsequent study, tested the hypothesis of normality for the periodic values of CPI, SPI, CV, ln CPI and ln SPI(t), where "ln" preceding the index is an abbreviation for logarithm. The five indicators were tested using 222 observations from ten software projects. Each indicator was tested for each project using three hypothesis tests for normality: Normal Probability Plot (NPP), Anderson-Darling (A-D), and Shapiro-Wilk (S-W). Thus, 150 tests were made and analyzed in the initial study.

The results from the testing rejected the hypothesis of a normal distribution for the indicators, CPI, CV, and SPI. However, it was shown that there is insufficient evidence to reject normality for both ln CPI and ln SPI(t). Furthermore the probability of normality was shown to be very high for the logarithmic representations, and very low for the untransformed indicators [Lipke, 2002 (December)].[e] This determination satisfied the fourth characteristic sought for the method of transformation.



Note: Dashed lines indicate the count for a normal distribution

Figure 2. Histograms

The study of additional project data undertaken in this paper is to further examine the conclusions reached in the initial paper. Are the findings from the first study coincidental, or are they corroborated from the analysis using new EVM data? Is there additional evidence that the periodic values of CPI and SPI(t) are log-normally distributed?

Methodology

From the 2002 study of normality it was seen that there was little difference in the hypothesis test results from the NPP, A-D, and S-W methods. Therefore, for this study, only the Anderson-Darling test is utilized [Annis, 2011]. As before, the hypothesis testing is performed at the 0.05 level of significance.[f]

To preclude false rejections of normality, SPC is applied to the data to identify and remove any signals. In this application, an observation is identified as a signal when its value exceeds, plus or minus, three standard deviations from the mean. The signals are identified using the logarithm values in the project data set. The observation removed from the logarithm representation is also removed from the un-transformed data set. For example, when the observation $\ln CPI_k$ is removed, CPI_k is also removed from the A-D analysis.

In summary, SPC signals are first identified and removed from the project data sets. Using the in-control observations, each of the four data representations for each project is tested for normality using the A-D test method. The results are compiled and the probability of normality is calculated using the binomial distribution.

Project Data

The project data for this study comes from EVM applications that are primarily labor only. The data sources are not related organizationally; they represent very different types of work and are located in four different countries.

Eight projects totaling 209 observations are included in the study. Four projects are aligned with information technology, two are from high technology, one is a maintenance application, and one is a construction project. The projects are small to moderate size, ranging in duration from one to two years.

Analysis

To determine whether the observed periodic values of CPI, ln CPI, SPI(t), and ln SPI(t) from individual projects are distributed normally, the hypothesis to be tested is formed:

Ho: The data are distributed normally
Ha: The data are not distributed normally

where Ho is termed the Null Hypothesis and Ha, the Alternate Hypothesis.

The hypothesis test of the data is made by comparing its calculated test statistic to the critical value associated with 0.05 level of significance; i.e., 0.752 [Annis, 2011]. For the testing of fewer than 30 observations, the test statistic is increased as a function of the number in the sample. When the test statistic value is less than the critical value, Ho is not rejected. However, when the test statistic exceeds the critical value, Ho is rejected in favor of Ha.

Results

Table 1 displays the count of signals identified from applying SPC to the periodic values of $\ln CPI^{-1}$ and $\ln SPI(t)^{-1}$. As shown, of the eight projects four had no signals. For the 209 observations, seven signals were identified for cost with eight for schedule. Thus, the application of SPC did not significantly reduce the number of observations and much of the data was preserved for hypothesis testing.

Observations & SPC Signals			
Project	Nr Observations	Cost Index Signals	Schedule Index Signals
#1	65	2	1
#2	12	0	0
#3	20	1	4
#4	20	3	0
#5	24	0	0
#6	24	0	0
#7	26	0	0
#8	18	1	3
Totals	209	7	8

Table 1. Observations & SPC Signals

The composite of the hypothesis testing results is shown in Table 2. The logarithm data representations strongly support the 2002 normality study conclusion; i.e., applying the natural logarithm to the periodic values for the EVM cost index and the ES schedule index transforms them such that they approximate a normal distribution. Also, supporting the findings from the 2002 study, normality was inconsistent for the untransformed indexes, CPI^{-1} and $SPI(t)^{-1}$.

Normality Hypothesis Test Results @ 5% Significance Ho: Normal distribution of data is not rejected (Test Statistic < 0.752) / Ha: Normal distribution of data is rejected								
	CPI^{-1}		$\ln CPI^{-1}$		$SPI(t)^{-1}$		$\ln SPI(t)^{-1}$	
Project	Ho / Ha	Test Statistic	Ho / Ha	Test Statistic	Ho / Ha	Test Statistic	Ho / Ha	Test Statistic
#1	Ha	1.4940	Ho	0.5940	Ha	2.7655	Ha	0.7567
#2	Ha	1.4581	Ho	0.6719	Ha	1.0181	Ho	0.2294
#3	Ha	1.3292	Ho	0.5079	Ha	1.1257	Ho	0.3075
#4	Ho	0.2013	Ho	0.4933	Ha	2.8854	Ho	0.5611
#5	Ha	1.8510	Ho	0.4718	Ha	1.2138	Ho	0.6903
#6	Ha	1.9733	Ha	0.7724	Ha	1.3401	Ho	0.7381
#7	Ha	0.9552	Ho	0.2664	Ha	2.0193	Ho	0.2758
#8	Ho	0.3632	Ho	0.3527	Ha	0.7710	Ho	0.4448
Count Ho	2		7		0		7	
Probability	14.5%		99.6%		0.4%		99.6%	

Table 2. Normality Hypothesis Test Results

For seven of the eight projects, the null hypothesis was the result from testing the logarithm data representations. Ho was rejected once each for $\ln CPI^{-1}$ and $\ln SPI(t)^{-1}$; in each instance the test statistic value was marginally larger than the critical value. However, the results are much different for the untransformed indexes. The periodic observations of CPI^{-1} from six of the eight projects were rejected for normality, whereas all eight $SPI(t)^{-1}$ data sets were rejected by the testing.

From the analysis of the eight projects, the probability that the logarithm transformation yields a normal distribution is extremely high, 99.6 percent for both $\ln CPI^{-1}$ and $\ln SPI(t)^{-1}$. Oppositely, the probabilities for CPI^{-1} and $SPI(t)^{-1}$ are very low, 14.5 and 0.4 percent, respectively.

Conclusion

The null hypothesis, indicating the data approximate a normal distribution, is supported at the 0.05 level of significance for the $\ln CPI^{-1}$ and $\ln SPI^{-1}(t)$ data representations. The hypothesis of normality is not a reasonable assumption for CPI^{-1} or $SPI(t)^{-1}$.

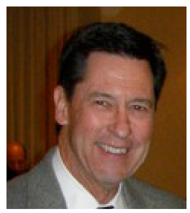
The result of testing the additional data from eight projects strongly corroborates the study performed in 2002. There is sufficient evidence to reasonably assume that the use of logarithms transforms the periodic values of CPI and SPI(t) (and their reciprocals) to approximate a normal distribution, when the EVM application is primarily labor only.

As recommended in the 2002 normality study, improvements in accuracy and reliability are foreseen from the use of ln CPI and ln SPI(t) in the statistical applications associated with forecasting, planning, estimating, process modeling, and simulation.

CONTINUED ON PAGE 18

Footnotes:

a) CPI is defined as the earned value (EV) divided by the actual cost (AC) [Humphreys, 2002].
b) SPI(t) is equal to the earned schedule (ES) divided by the actual time (AT) [Lipke, 2011]
c) "SPC signals" are data values that are separated from the population mean by more than three standard deviations [Pitt, 1995].
d) The standard deviation is a measure of the variation in the observed values [Pitt, 1995]
e) For the logarithmic representations, the probability of normality ranged from 86.1 to 100 percent, whereas the probabilities for the indexes and CV ranged from 0.8 to 5.2 percent.
f) The 0.05 level of significance indicates there is a five percent probability of the test indicating the distribution is not normal when, in fact, it is [Crow, et al, 1960].



About the Author – Walt Lipke retired in 2005 as deputy chief of the Software Division at Tinker Air Force Base. He has over 35 years of experience in the development, maintenance, and management of software for automated testing of avionics. During his tenure, the division achieved several software process improvement milestones, including the coveted SEI/IEEE award for Software Process Achievement. Mr. Lipke has published several articles and presented at conferences, internationally, on the benefits of software process improvement and the application of earned value management and statistical methods to software projects. He is the creator of the technique Earned Schedule, which extracts schedule information from earned value data. Mr. Lipke is a graduate of the USA DoD course for Program Managers. He is a professional engineer with a master's degree in physics, and is a member of the physics honor society, Sigma Pi Sigma (ΣΠΣ). Lipke achieved distinguished academic honors with the selection to Phi Kappa Phi (ΣΣΣ). During 2007 Mr. Lipke received the PMI Metrics Specific Interest Group Scholar Award. Also in 2007, he received the PMI Eric Jenett Award for Project Management Excellence for his leadership role and contribution to project management resulting from his creation of the Earned Schedule method. Mr. Lipke was selected for the 2010 Who's Who in the World.

References:

Annis, C. "Goodness of Fit Tests for Statistical Distributions," http://www.statisticalengineering.com/goodness.htm, 2011

Crow, E. L., F. A. Davis and M. W. Maxfield. Statistics Manual. New York: Dover, 1960

Humphreys, G. Project Management Using Earned Value, Orange, CA: Humphreys & Associates, 2002.

Lipke, W. "Concept Description," http://www.earnedschedule.com/, 2011

Lipke, W. "A Study of the Normality of Earned Value Management Indicators," The Measurable News, December 2002

Lipke, W. "Statistical Process Control of Project Performance," CrossTalk, March 2002

Pitt, H. SPC for the Rest of Us, Reading, MA: Addison-Wesley, 1995

Stahel, W. A., E. Limpert and M. Abbt. "Log-normal Distributions across the Sciences: Keys and Clues." BioScience, May 2001: 341-352.