Distribution of Cost Growth in Robotic Space Science Missions

Christopher Swan
Jet Propulsion Laboratory, Pasadena, CA

Abstract:

Cost growth characterization is a critical factor for effective cost risk analysis and project planning. This study analyzed low level budget changes in Jet Propulsion Laboratory-managed space science missions, which occurred during the development of the project. The data was then curve fit, according to cost distribution categories, to provide a reference set of distribution parameters with sufficient granularity to effectively model cost growth in robotic space science missions.

1.0 INTRODUCTION:

Cost growth characterization is instrincally valuable for cost risk analysis and project planning. In addition to providing crucial information for project planning, it also provides cost growth distributions for use in cost risk exercises. These distributions, are then grounded in historical basis, from which more accurate cost risk analyses can be made early in a project's life cycle.

The critical factor in cost growth characterization as it relates to cost risk analysis is that it enables risk specification in the form of probability distributions. These distributions (typically normal, lognormal, and triangle) delineate a curve that the probable costs will follow and can vary greatly according to the parameters describing the curve. Figure 1 illustrates the effect of varying parameters on the shape of the probability distribution. The parameters (derived from this study's data) show that both the average cost growth and the variance of cost is clearly greater for one distribution versus another.

Cost-risk analyses combine the distributions of each element in the risk analysis and presents the results in a cumulative distribution function (also called an "S-curve"), which gives the probability of the final total costs being less than or equal to a predicted cost. An example of an S-curve showing the commonly accepted "low risk" point of 70% is shown in Figure 2.

By analyzing historical costs, the parameters of these distributions can be determined empirically, thus providing some much needed credibility to these analyses.

This study does not attempt to explain the reasons behind cost growth or to analyze the actions taken to address the growth. While, these are both important subjects, they are more difficult to address in a quantitative

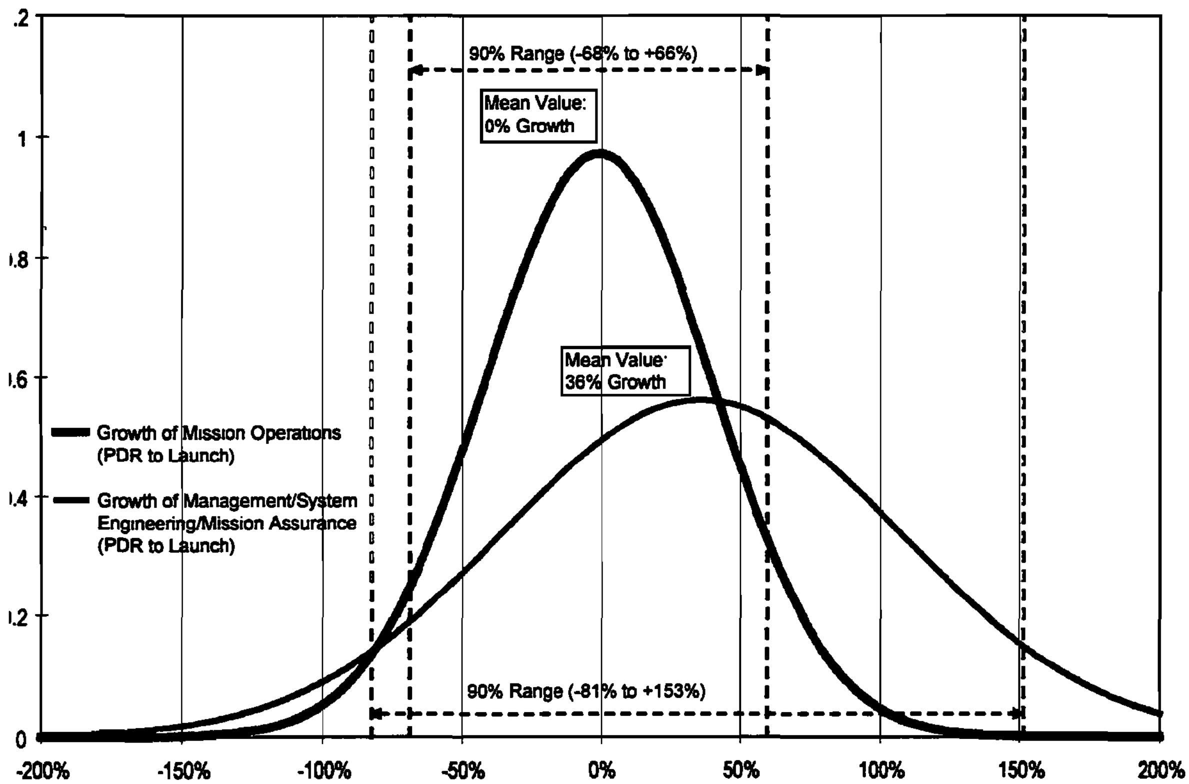

Figure 1. Sample of the impact of various parameters for the same distribution.

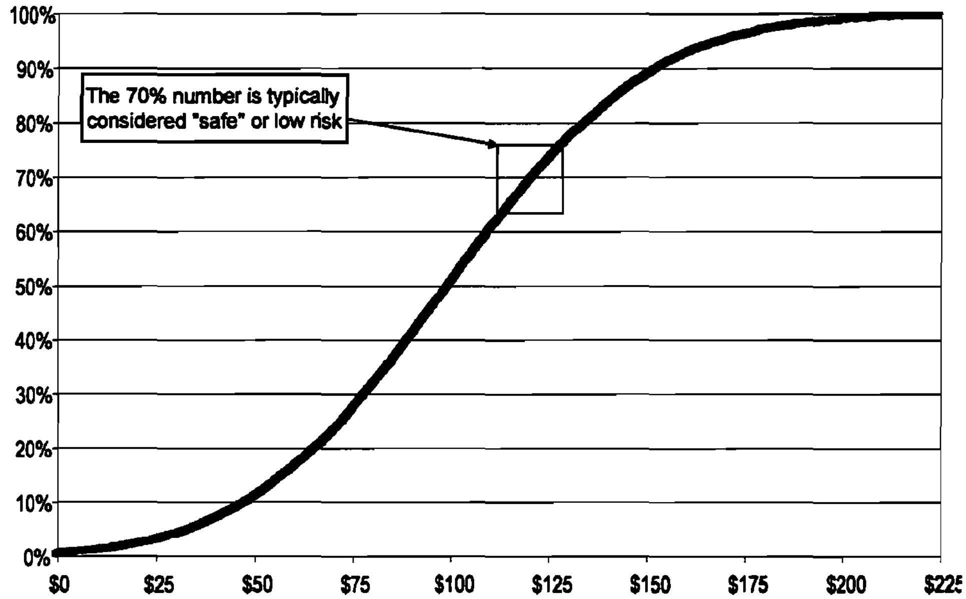

Figure 2. Sample S-curve showing the 70% low-risk cost.

manner and do not impact the core question addressed in this study: how did costs grow?

2.0 METHODOLOGY:

The main challenge to performing cost growth analysis is the availability of sufficient data to obtain statistically valid results. This study took advantage of the Jet Propulsion Laboratory's (JPL) New Business Systems (NBS), which archives budgetary information at the cost account level for each budget version. Although the aforementioned level of detail is only available for post-1996 projects (projects that began after the system's implementation in 1996), there are three JPL missions (comprising 374 accounts) that meet this criteria: NBS contains the complete development cost history for the Mars Exploration Rover (MER), Mars Reconnaissance Orbiter (MRO), and Cloudsat projects. The budgetary information from these projects was then sorted against major project milestone dates—Preliminary Design Review (PDR), Critical Design Review (CDR), and Launch)—to determine the current budget at those given points in time. Each version of the project budget consisted of two distinct parts, actual costs to date and estimated costs to complete, seemlessly integrated by NBS for every project account. In addition, every budget version also has an effective budget date that was matched to the major milestone dates obtained from CADRE[2] and JPL's internal scheduling resources.

What About Reserves?

Although project cost reserves are integral to the financial plan for any project, they are not considered in this study, because budget information (both actual and estimated) only contains expended reserves; reserves that have not been expended are not tracked. Fundamentally, this study explored cost growth over time against the current best estimate (CBE) of a project; how the project dealt with the cost growth (i.e., applying reserves or descopes) was not analyzed.

After obtaining the cost growth data at the account level, the tasks were grouped according to criteria that provided additional granularity in the results according to both a JPL-

2 Cost Analysis Data Requirement is a NASA required document produced at major project milestones documenting the current cost, schedule, and technical baseline of a project.

specific work breakdown structure (WBS—Table 1) and a generic "type of work" categorization (outlined in Table 2). The grouping of the tasks according to these two categories was performed by three separate cost engineers; any discrepancies were resolved to ensure that the most accurate binning was achieved.

Table 1. WBS Categories and Definitions

Category	Description
Management/System Engineering/Mission Assurance	Tasks related to project management, project system engineering, or mission/quality assurance.
Payload	Tasks involving designing, building, and testing the science payload of the mission.
Flight System	Effort and procurements needed to build the spacecraft(s) for the mission.
Science	The technical and management efforts of directing and controlling the Science investigation aspects of the project
Mission Operations—Ground Data System	Tasks needed to build the hardware and software to operate the mission.

Table 2. "Type of Work" Categories and Definitions

Category	Description
Management	Management of people and tasks.
Contract	Account is primarily a subcontract (may include contract technical management).
Hardware	Procurements and labor required to produce flight and ground hardware.
Engineering	Engineering effort and analysis not directly related to hardware production (system engineering, modeling, software, algorithms, analysis, etc.)
Science	The technical and management efforts of directing and controlling the Science investigation aspects of the project

The cost data was normalized to fiscal year 2006 dollars by using a constant inflation rate of 3.1%. All subsequent cost growth information was then reduced to percent growth from PDR to Launch and CDR to Launch using equations 1 and 2, respectively.

$$\%growth = \frac{LaunchCost - PDRcost}{PDRcost} \times 100\% \quad (1)$$

Equation for cost growth from PDR to Launch

$$\%growth = \frac{LaunchCost - CDRcost}{CDRcost} \times 100\% \quad (2)$$

Equation for cost growth from CDR to Launch

After being reduced to percent growth, statistical software—@Risk Professional developed by Palisade Corporation—was used to fit distributions to each data category. The BestFit, @Risk Professional's curve fitting module, uses the Maximum Likelihood Estimator method to fit distributions to data. The software was allowed to dynamically adjust the number of Chi-squared bins and the size of each bin to provide the best fit for each distribution (equalprobable bins).

Each data set was matched with commonly used distributions in cost-risk analysis—Triangle, Normal, and Lognormal. To determine which distributions would be the best fit for the provided data, a goodness-of-fit test was applied to each distribution. This was achieved via the chi-squared statistic (shown in equation 3), which provides a p-value to indicate the likelihood of a curve matching a particular data set.

$$X^2 = \sum_{i=1}^{K} \frac{(N_i - E_i)^2}{E_i} \quad (3)$$

K = The number of bins

N_i = The observed number of samples in the i^{th} bin

E_i = The expected number of samples in the i^{th} bin

Chi-squared statistic

The p-value ranges from 0 (no match) to 1.0 (perfect match) and in instances where the p-value was zero for all distributions, the ranking was done according to an @Risk proprietary fit statistic called "*the critical value*."

2.1 DATA SOURCE ERROR

A small number of tasks were removed from the analysis because they seemed to illustrate accounting errors rather than cost growth. Any account that showed a growth rate of less than -100% (i.e., negative budget) or accounts that showed cost growth more than 1000% (indicating that the project did not originally budget) were removed. All told, a total of three accounts were removed from both the PDR's 334 account analysis and the CDR's 361 account analysis.

3.0 RESULTS

Before analyzing the individual distributions, it was important to view the different categorizations of data in context with one another. Table 3 illustrates the overall metrics of the data set with regards to cost growth and the total value of the accounts studied. The number of accounts increased over time as some tasks simply were not budgeted until late in the project. This is not a cost growth analysis issue, as the majority of the accounts are budgeted by CDR and the value of the unbudgeted accounts represents less than 3% of the total value.

Using the "type of work" category, Table 4 shows that the majority of the tasks fall under contract, hardware, or engineering with the other two categories accounting for less than 10% of the total dollar value. The average value of accounts in the contract category is much higher than the others due to several large contracts in the data set (MRO/Cloudsat flight system contract, etc.). This categorization does a good job of illustrating the dramatic risk of developing hardware in contrast to the relatively low risk of management and science. The contract and engineering categories have some disadvantages as contracts include a wide variety of work; the large number of engineering accounts included in the analysis suggests that additional engineering sub-categories could provide further granularity.

Similar metrics for the WBS categories are shown in Table 5. Accounts related to the development of the flight system dominate this categorization, accounting for almost 80% of the total dollars.

The resulting distributions for each category are illustrated in data Appendix I, figures I-1 through I-22 with the overall growth distributions shown in figure 2. Each distribution has an associated p-value, which describes how accurately it matches the data it is fitted to.

Table 3. Metrics on total number of accounts and overall growth

	Number of Accounts	Average Value of Each Account	Total Value	Average % Growth to Launch
PDR	334	$3,278,394	$1,094,983,466	36%
CDR	361	$3,109,709	$1,122,605,088	19%
Launch	374	$3,011,109	$1,126,154,847	N/A

Table 4. Metrics on the "type of work" categorization

Type of Work	Number of Accounts	Average Value of Each Account	Total Value	% Growth from PDR to Launch			% Growth from CDR to Launch		
				Min.	Avg.	Max.	Min.	Avg.	Max.
Management	44	$1,156,592	$50,890,026	-94%	15%	159%	-100%	1%	139%
Contract	24	$16,810,092	$403,442,216	-97%	32%	152%	-38%	18%	150%
Hardware	82	$3,753,918	$307,821,297	-63%	94%	430%	-63%	45%	249%
Engineering	191	$1,765,675	$337,243,853	-100%	23%	486%	-79%	16%	441%
Science	33	$810,832	$26,757,455	-97%	-1%	201%	-98%	-4%	111%

Table 5. Metrics on the work breakdown structure categorization

WBS	Number of Accounts	Average Value of Each Account	Total Value	% Growth from PDR to Launch			% Growth from CDR to Launch		
				Min.	Avg.	Max.	Min.	Avg.	Max.
Mgmt/SE/MA	56	$1,315,841	$73,687,081	-94%	36%	325%	-99%	13%	160%
Payload	50	$3,102,830	$155,141,496	-56%	69%	486%	-56%	35%	345%
Flight System	161	$4,931,686	$794,001,393	-100%	54%	430%	-63%	31%	441%
Science	36	$679,531	$24,463,099	-97%	-2%	201%	-100%	-6%	111%
MOS/GDS	71	$1,110,729	$78,861,778	-99%	-10%	159%	-79%	-3%	209%

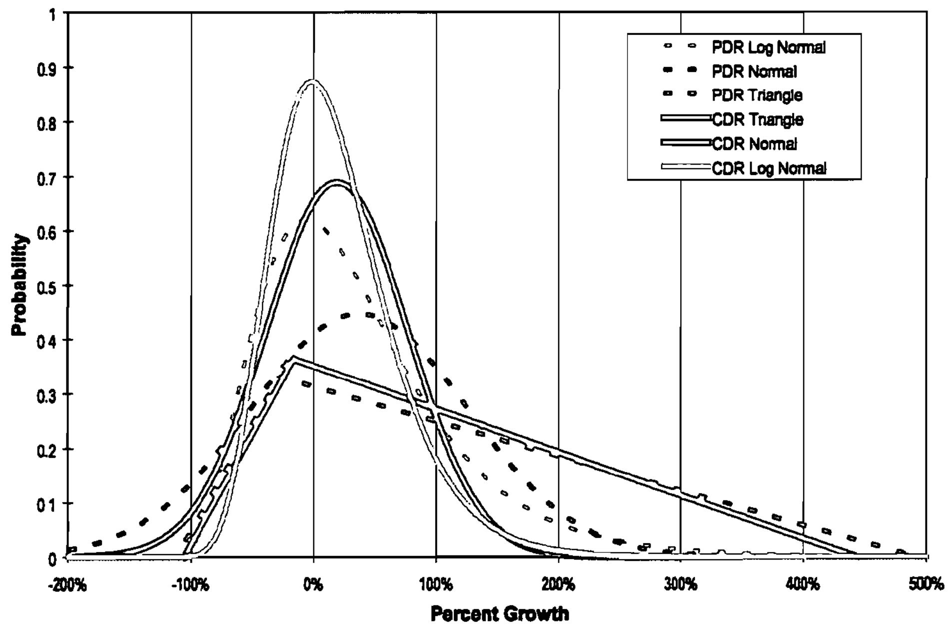

Figure 2. Various fitted distributions showing cost growth from both PDR and CDR to launch2

4.0 DISCUSSION

The results of this study illustrate the need for conservative risk assessments with distributions that can account for the full variance of cost growth. The data clearly indicates that cost growth can be as low as -100% and as high as almost 500% (Tables 4 and 5), which is a much wider than typical cost risk analyses assume. It also shows that real differences exist in the risks associated with different types of work, thus promoting a need for more granularity in cost risk analyses to capture risks more effectively.

In reviewing the distributions generated by this study (Figs. I-1–I-22), it becomes apparent that, in many cases, the outlying data points govern the fit of the data. This results in triangle distributions that are highly conservative and lognormal/normal distributions that fail to effectively capture some of the outliers, but are typically a better fit to the more nominal cost growth data. This suggests that either the outlying data points are true anomalies and should not be included in the determination of distributions or the outliers are simply a product of the available data.

Regardless of the issues with outlying data points, the data produced a set of distributions which are cataloged according to fit (Table 6) or with their associated histogram (I-1-22). While additional data and validation is needed before declaring these results definitive, these parameters provide an important, historical point of reference for cost risk analyses.

Table 6. Equations for each of the best-fit distributions[2]

Growth from Milestone	Category	Distribution	Mean	Standard Deviation	Shift (lognormal only)	p-value
PDR	Contract	Normal	0.316	0.631		0.733
PDR	Engineering	Lognormal	1.713	0.692	-1.49	0
PDR	Flight System	Lognormal	2.229	1.699	-1.699	0.019
PDR	General	Normal	0.37	0.891		0
PDR	Hardware	Lognormal	2.01	1.088	-1.075	0.789
PDR	Management	Lognormal	2.813	0.465	-2.658	0.28
PDR	Mgmt/SE/MA	Normal	0.359	0.71		0.3053
PDR	MOS/GDS	Normal	-0.009	0.409		0.0036
PDR	Payload	Lognormal	1.551	0.87	-0.87	0.2371
PDR	Science	Lognormal	1.629	0.515	-1.641	0.47
PDR	WBS Science	Lognormal	1.576	0.508	-1.598	0.2365
CDR	Contract	Lognormal	0.932	0.368	-0.753	0.206
CDR	Engineering	Lognormal	1.171	0.473	-1.023	0
CDR	Flight System	Lognormal	1.209	0.522	-0.909	0
CDR	General	Normal	0.189	0.578		0
CDR	Hardware	Lognormal	1.472	0.555	-1.029	0.061
CDR	Management	Normal	0.001	0.45		0.029
CDR	Mgmt/SE/MA	Lognormal	3.103	0.469	-2.97	0.0457
CDR	MOS/GDS	Lognormal	1.326	0.337	-1.355	0
CDR	Payload	Lognormal	1.123	0.579	-0.785	0.1367
CDR	Science	Normal	0.004	0.388		0.237
CDR	WBS Science	Normal	-0.006	0.428		0.0653

5.0 REFERENCES

Black, Hollis. "Fudamentals of Cost Risk Assessment in Aerospace Programs," AIAA Space 2001—Conference and Exposition, Albuquerque NM, 2001.

D'Agostino, Ralph B. and Stephens, Michael A. "Goodness-of-fit-techniques",CRC Press. Boca Raton, FL, 1986.

2004 NASA Cost Estimating Handbook (CEH), NASA Procedural Requirements (NPR) 7120.5C.

Software Manual: Guide to Using @RISK, Risk Analysis and Simulation Add-In for Microsoft® Excel, Version 4.5, Palisade Corporation, February, 2002, Newfield, New York.

Interview and discussion with Palisade developer, January 25, 2007, People in teleconference: Developer—Erik Westwig; Sales Representative—David Doran.

6.0 ACKNOWLEDGEMENTS

Thanks are extended to the following individuals for contributing to the production of this report: Tom Coonce, Nagin Cox, Julie Wertz-Chen, Corey Harmon, Leigh Rosenberg, and Jeanné Washington.

2 Triangle is not included in this table because it is never the "best fit". Refer to Figures I-1 through I-22 for all distribution parameters including triangle.

I. DATA APPENDIXES

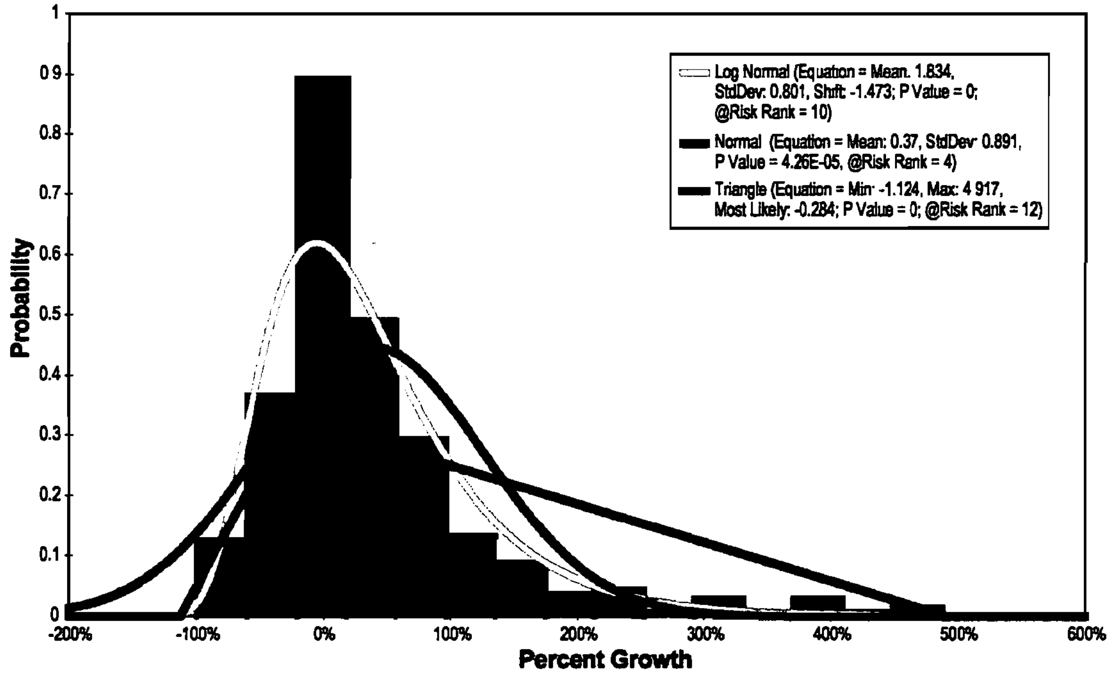

Figure I-1. Histogram with various distributions for Cost growth from PDR to Launch

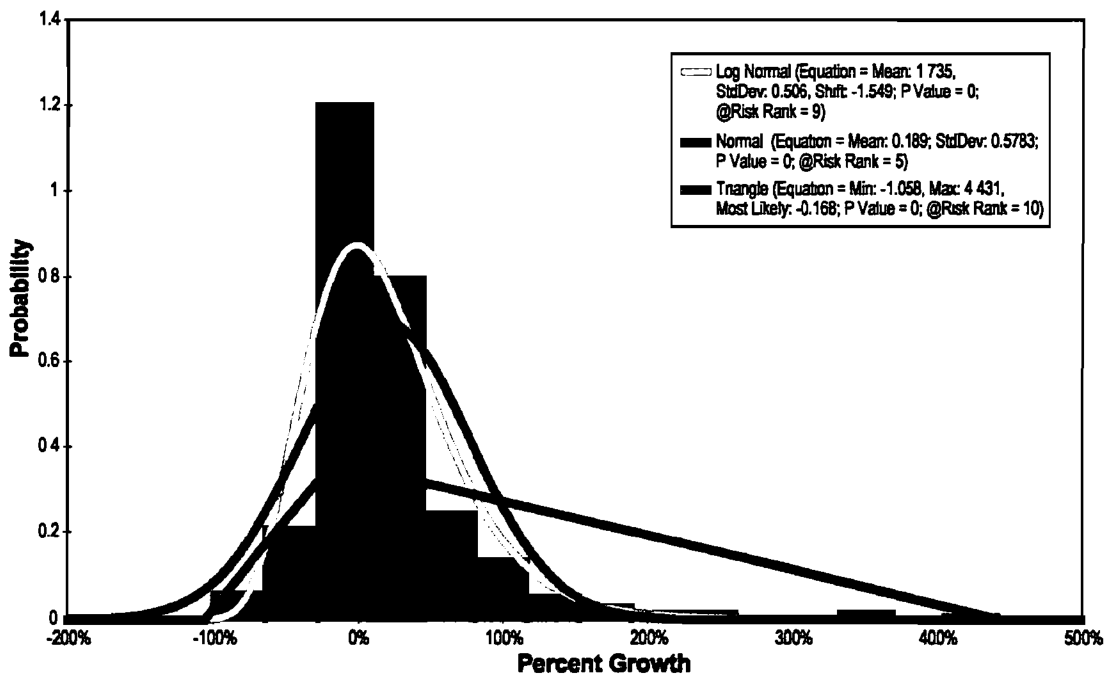

Figure I-2. Histogram with Various Distributions for Cost Growth from CDR to Launch

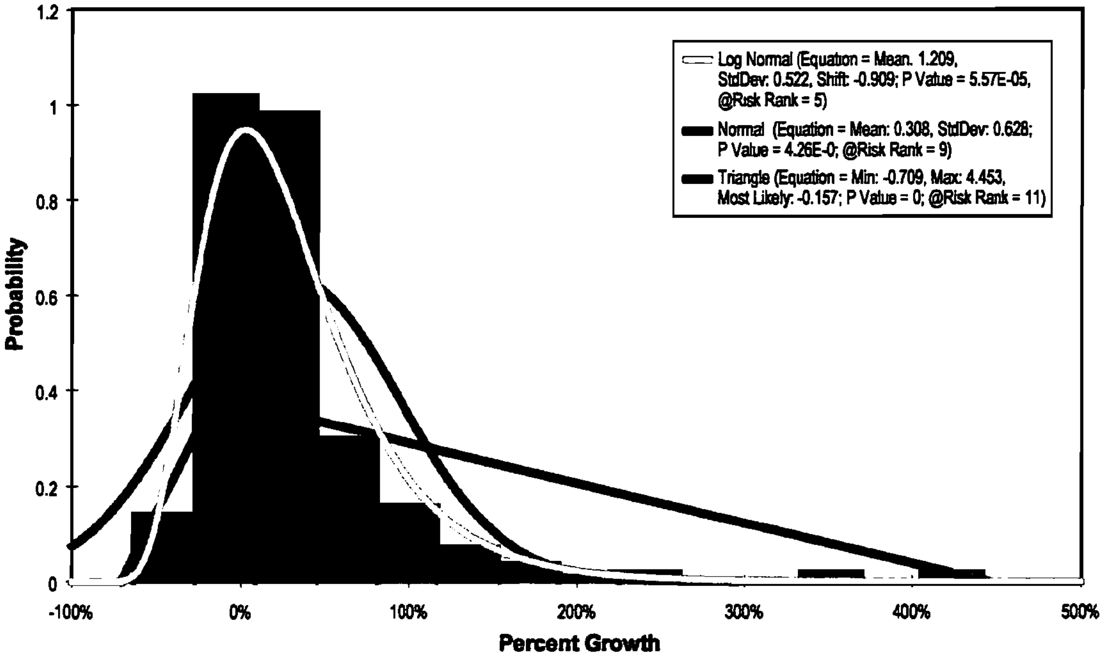

Figure I-3. Histogram with various distributions for Flight System Cost Growth from CDR to Launch

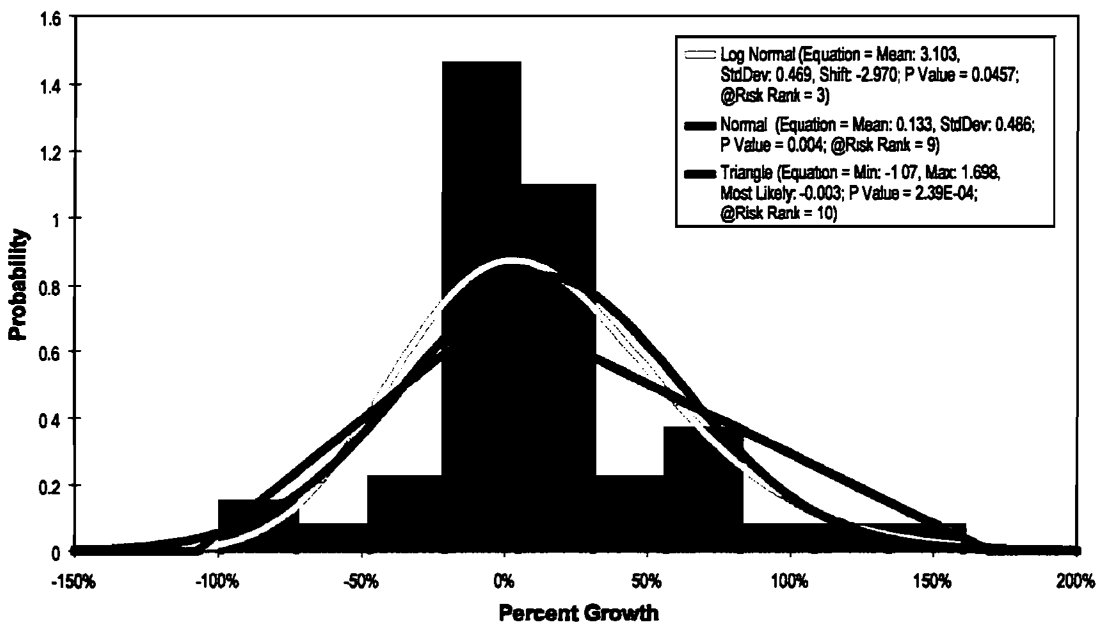

Figure I-4. Histogram with various distributions for Mgmt/ SE/MA Cost Growth from CDR to Launch

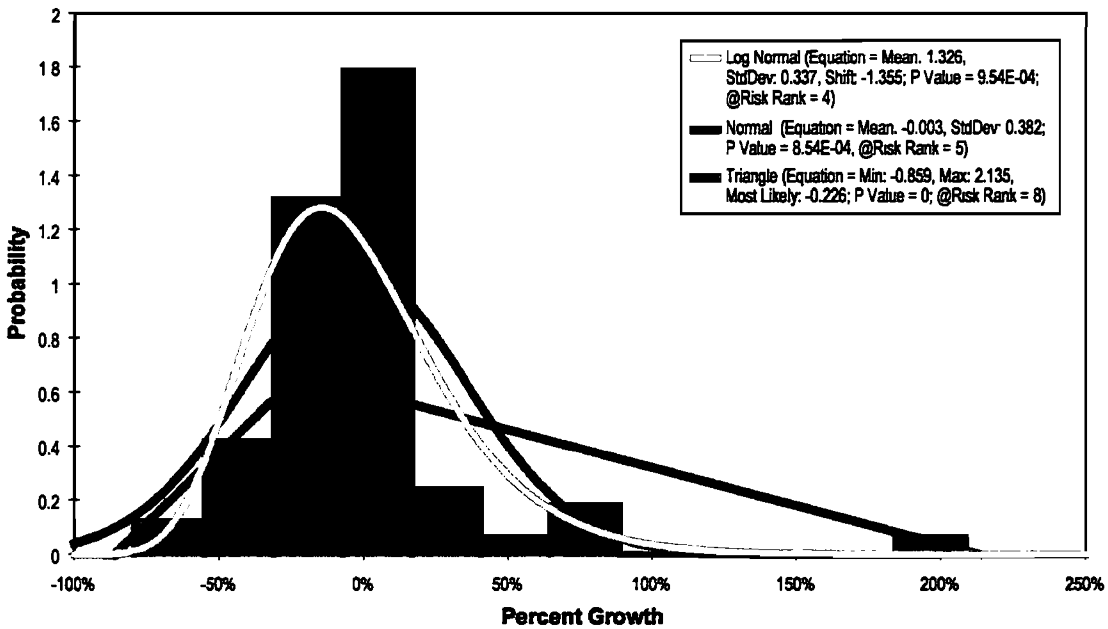

Figure I-5. Histogram with various distributions for MOS/GDS Cost Growth from CDR to Launch

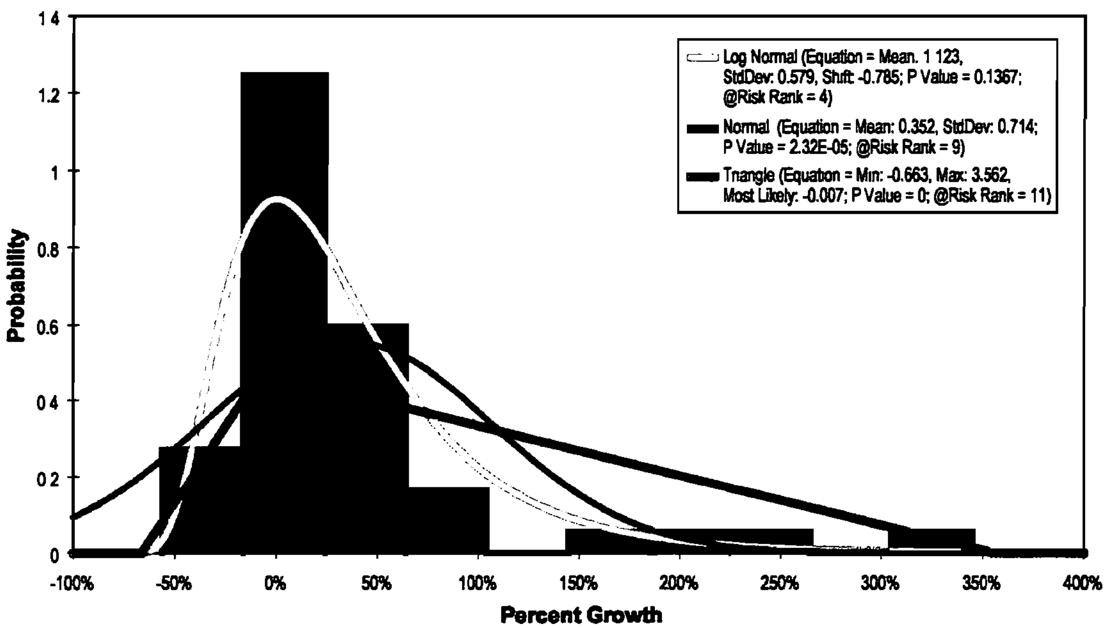

Figure I-6. Histogram with various distributions for Payload Cost Growth from CDR to Launch

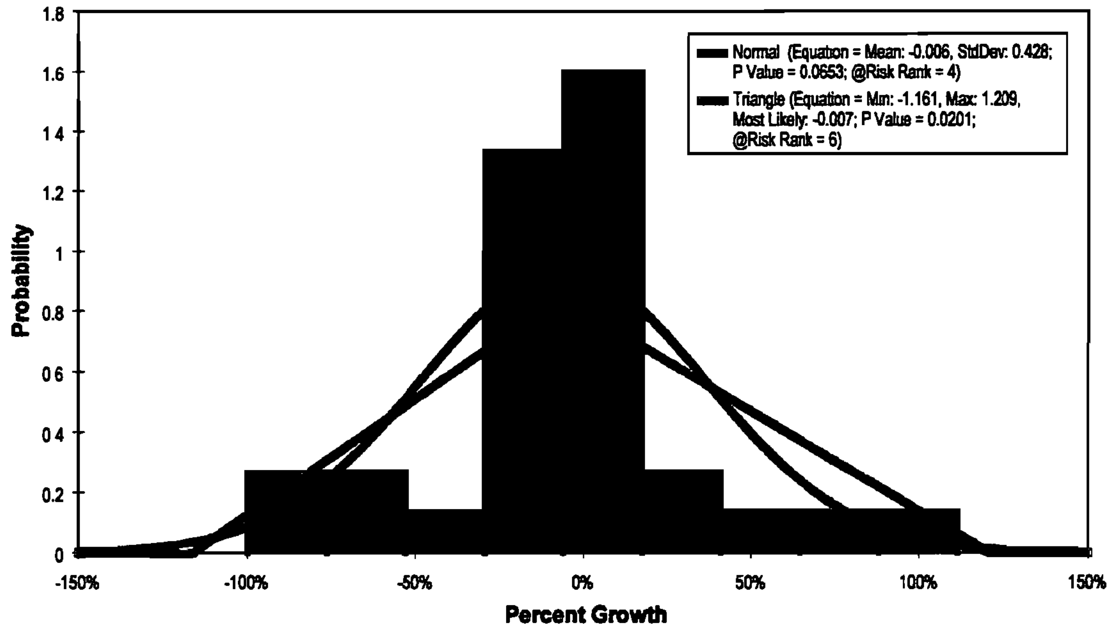

Figure I-7. Histogram with various distributions for WBS Science Cost Growth from CDR to Launch

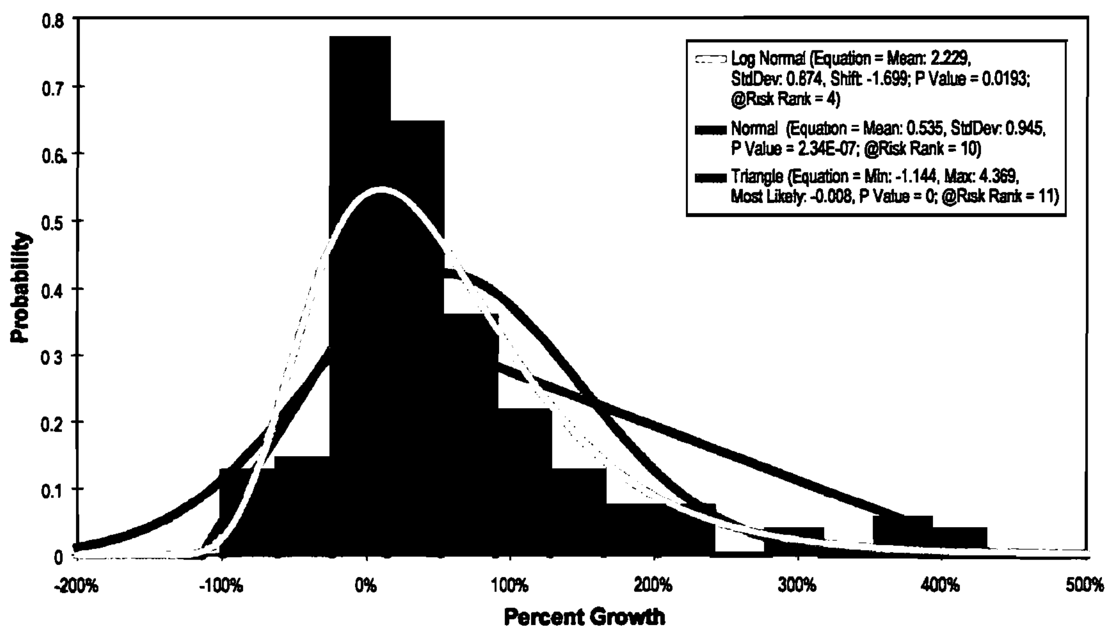

Figure I-8. Histogram with various distributions for Flight System Cost Growth from PDR to Launch

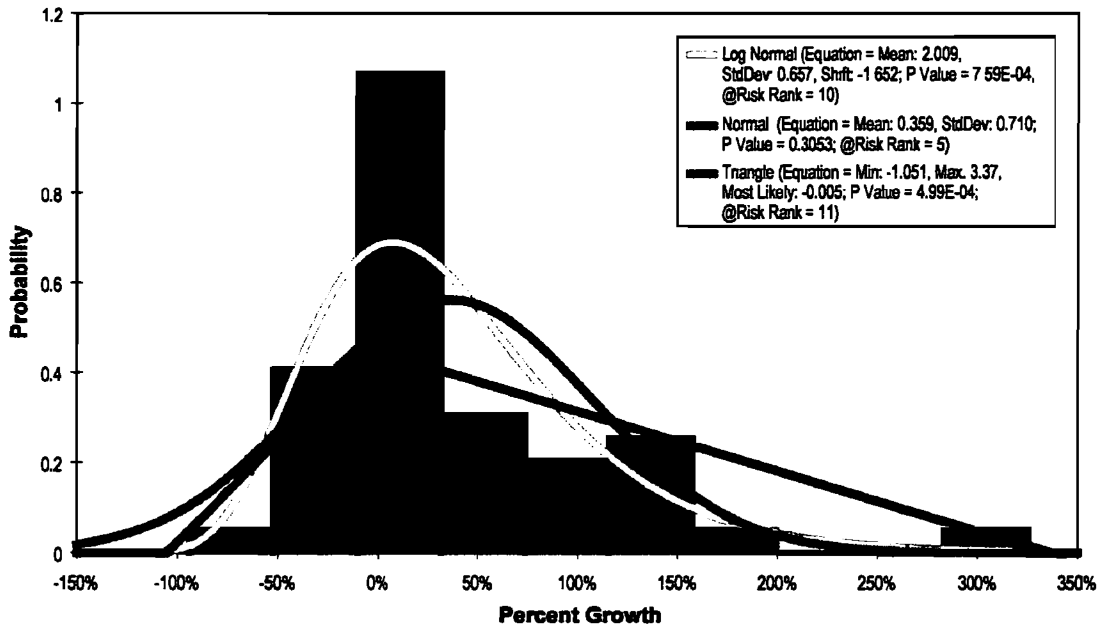

Figure I-9. Histogram with various distributions for Mgmt/ SE/MA Cost Growth from PDR to Launch

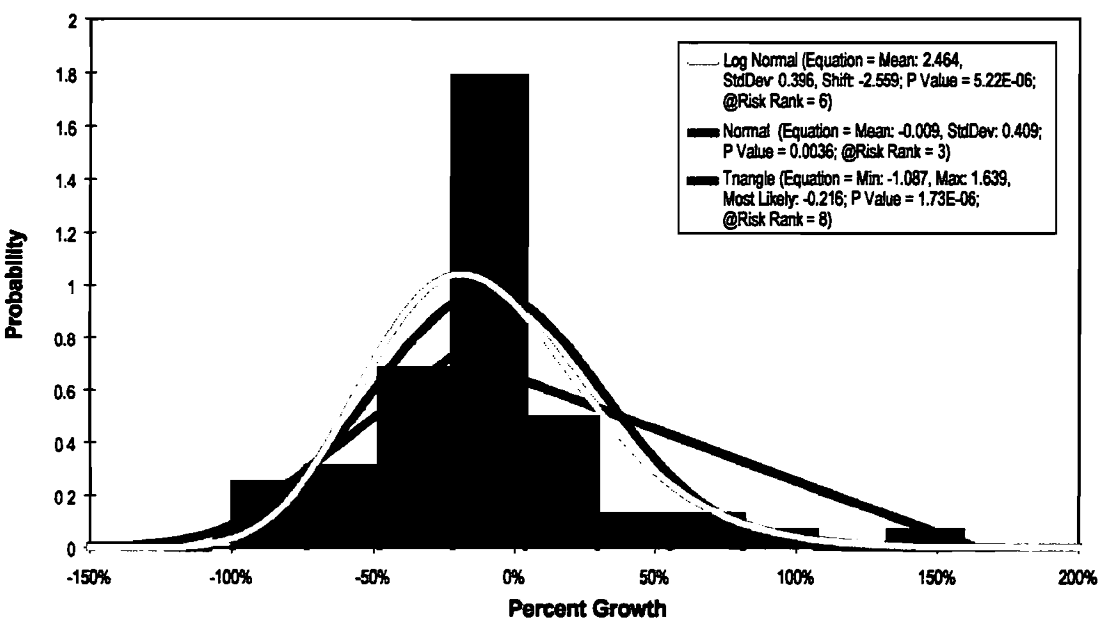

Figure I-10. Histogram with various distributions for MOS/ GDS Cost Growth from PDR to Launch

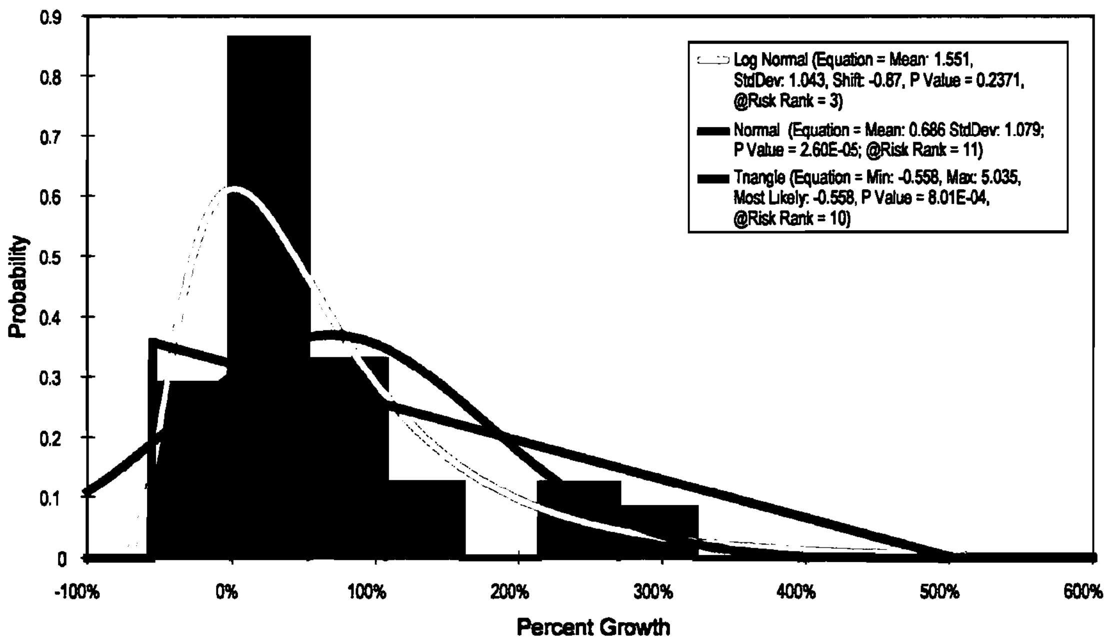

Figure I-11. Histogram with various distributions for Payload Cost Growth from PDR to Launch

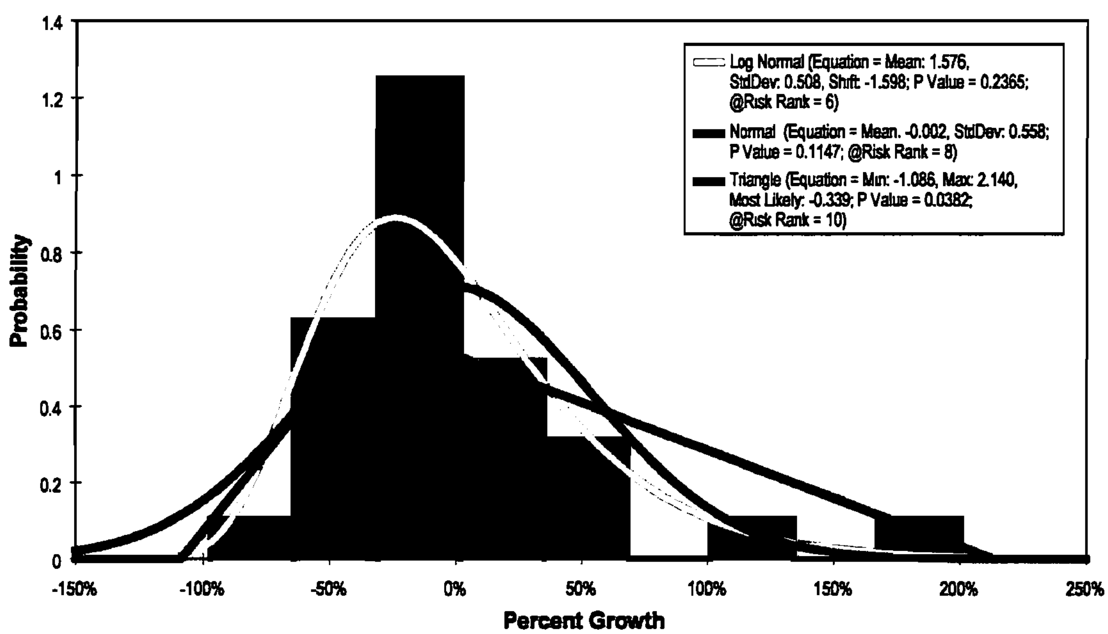

Figure I-12. Histogram with various distributions for WBS Science Cost Growth from PDR to Launch

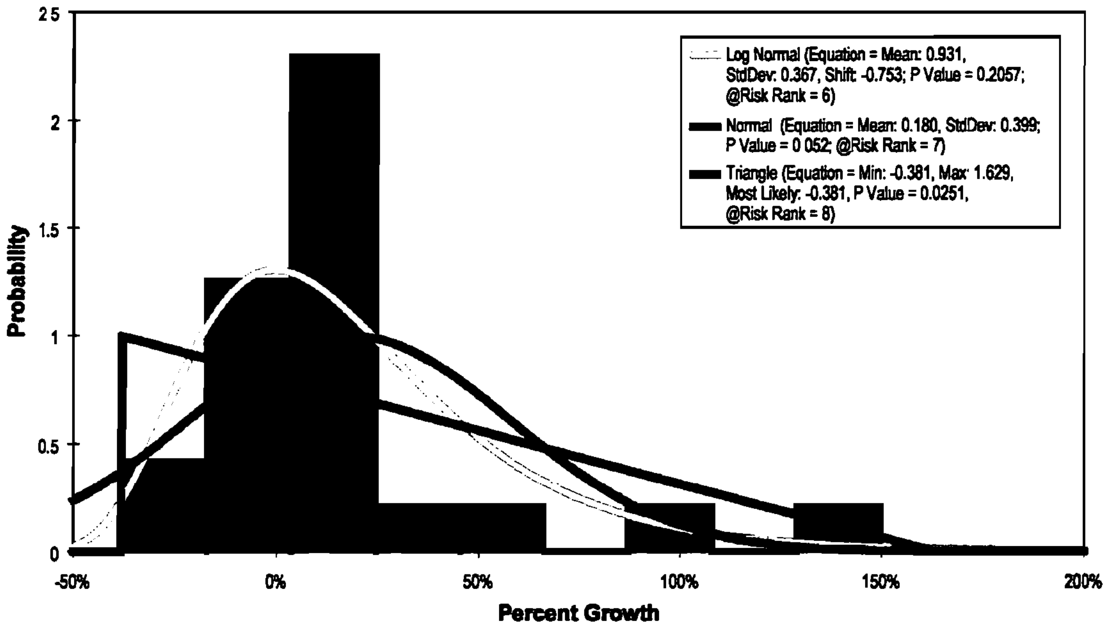

Figure I-13. Histogram with various distributions for Contract Cost Growth from CDR to Launch

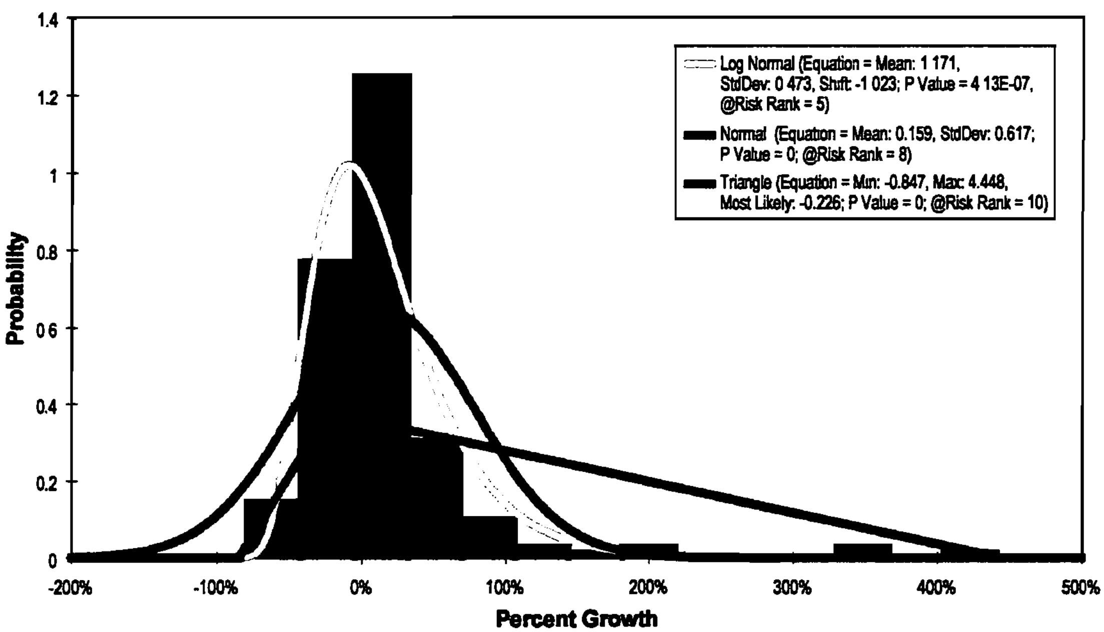

Figure I-14. Histogram with various distributions for Engineering Cost Growth from CDR to Launch

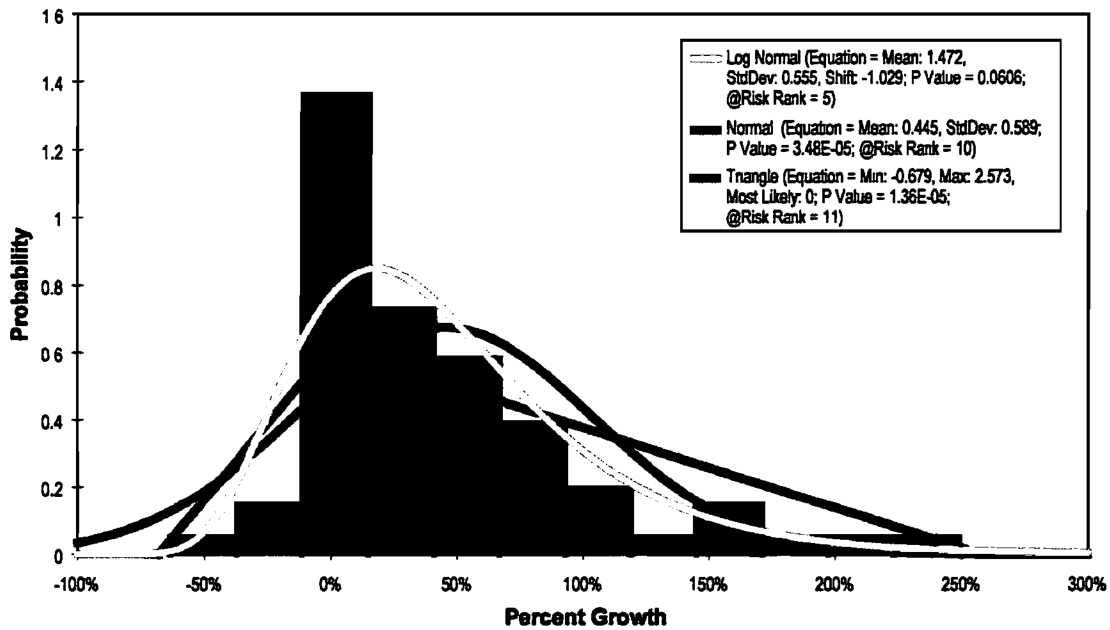

Figure I-15. Histogram with various distributions for Hardware Cost Growth from CDR to Launch

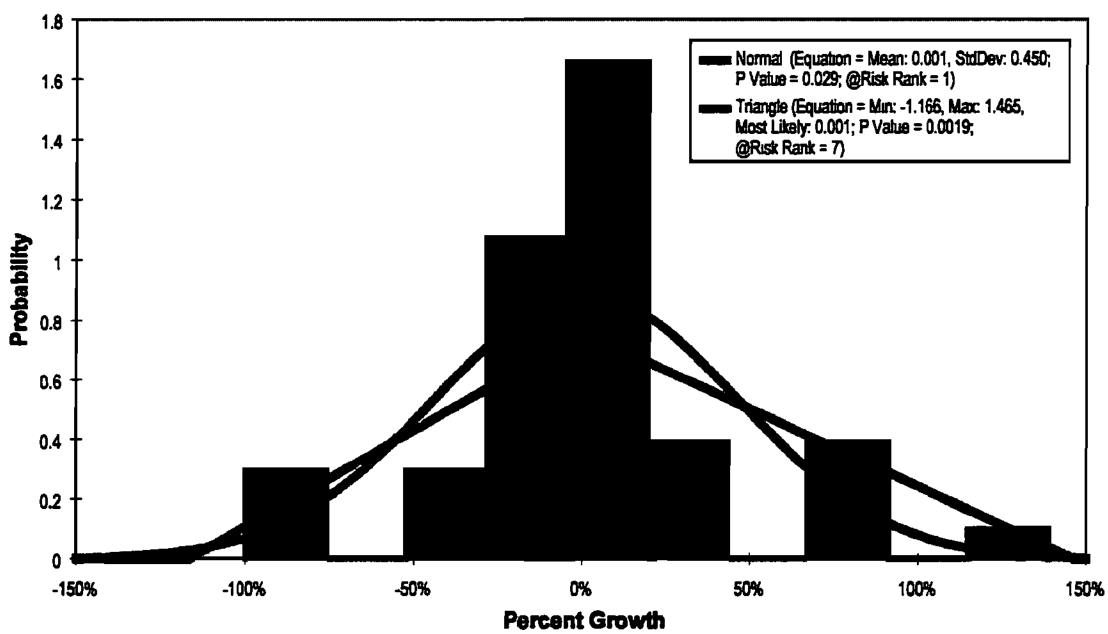

Figure I-16. Histogram with various distributions for Management Cost Growth from CDR to Launch

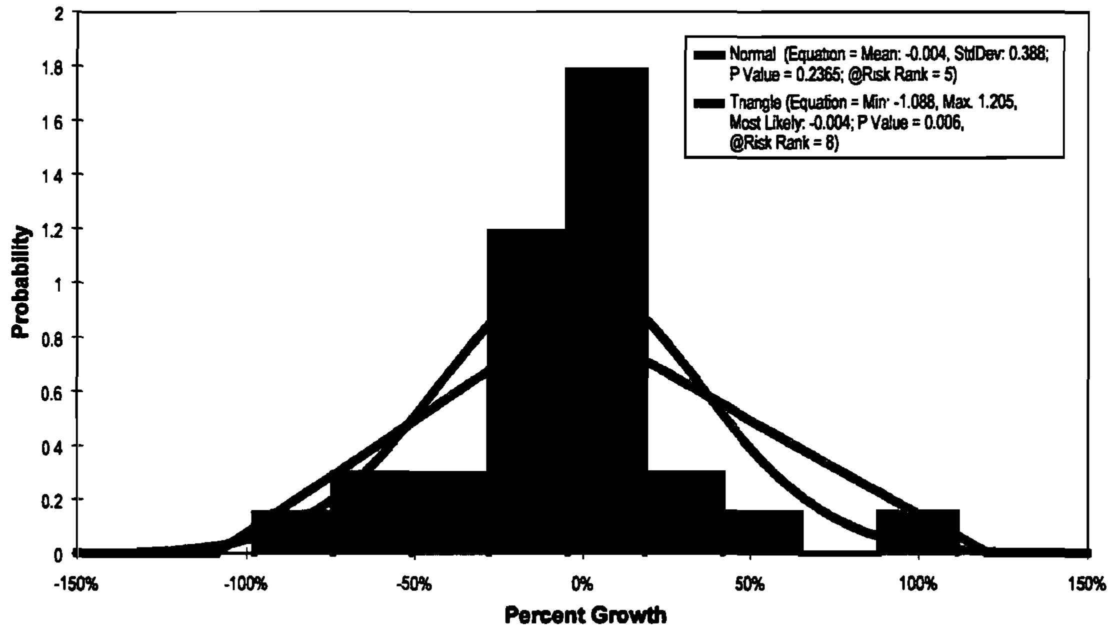

Figure I-17. Histogram with various distributions for Science Cost Growth from CDR to Launch

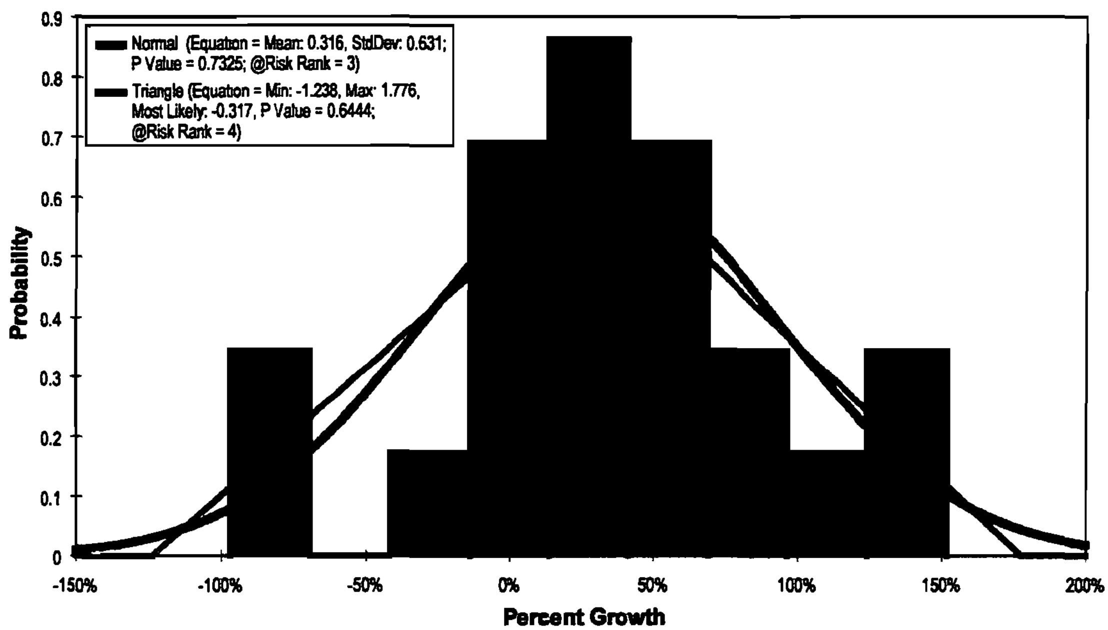

Figure I-18. Histogram with various distributions for Contract Cost Growth from PDR to Launch

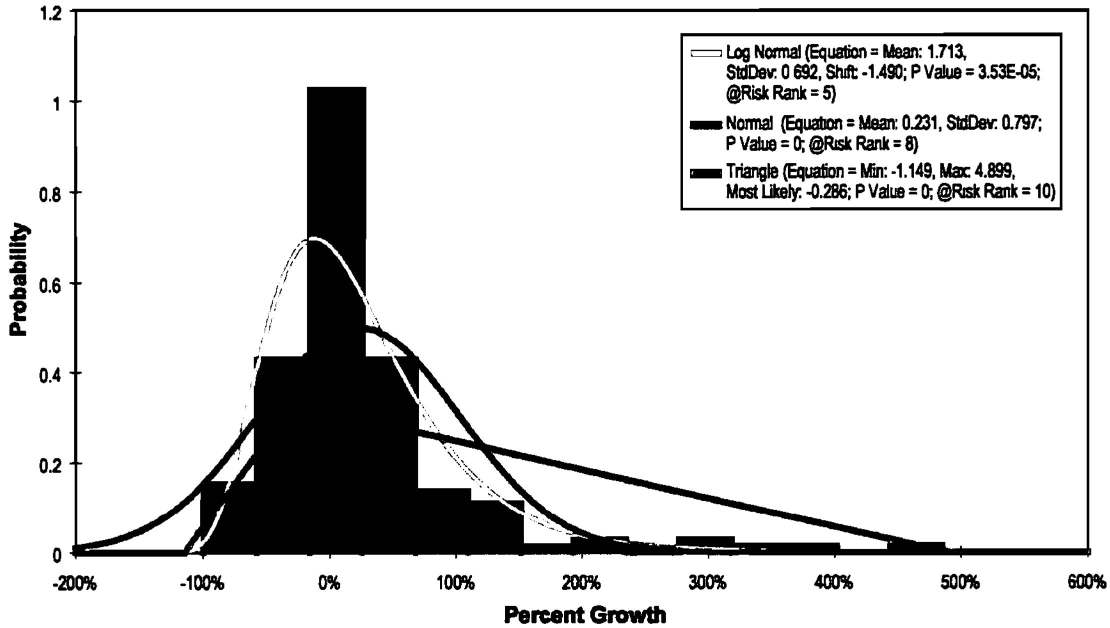

Figure I-19. Histogram with various distributions for Engineering Cost Growth from PDR to Launch

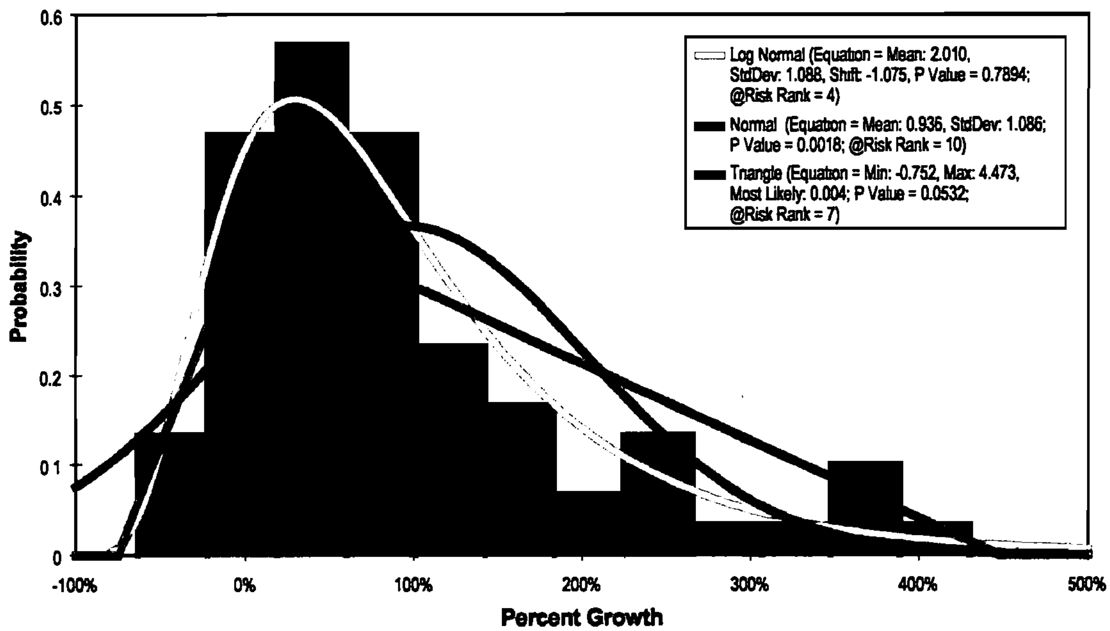

Figure I-20. Histogram with various distributions for Hardware Cost Growth from PDR to Launch

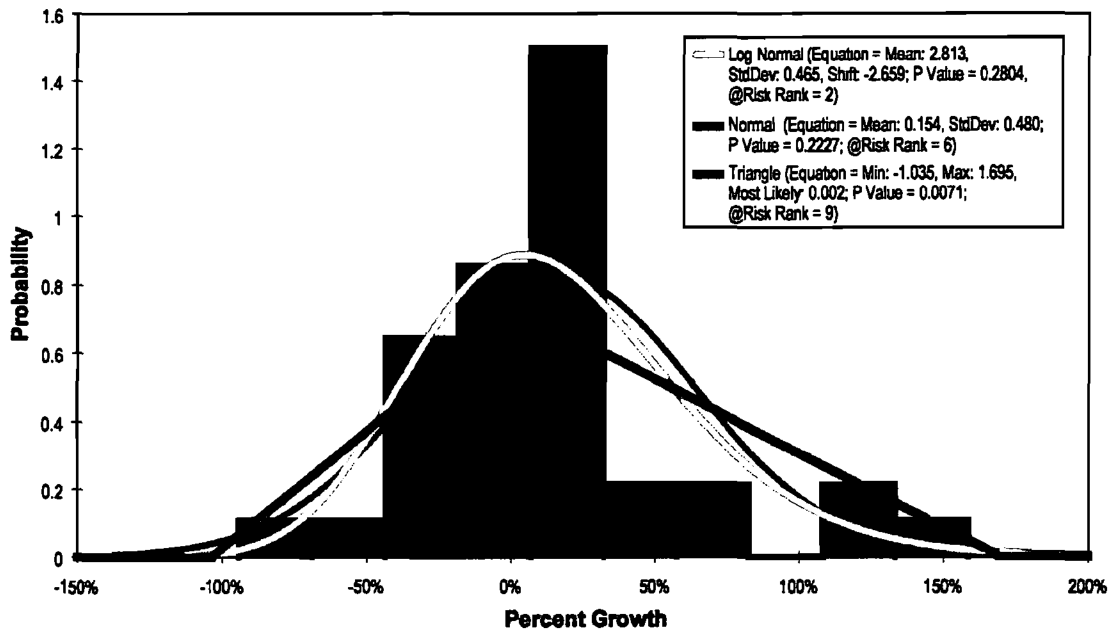

Figure I-21. Histogram with various distributions for Management Cost Growth from PDR to Launch

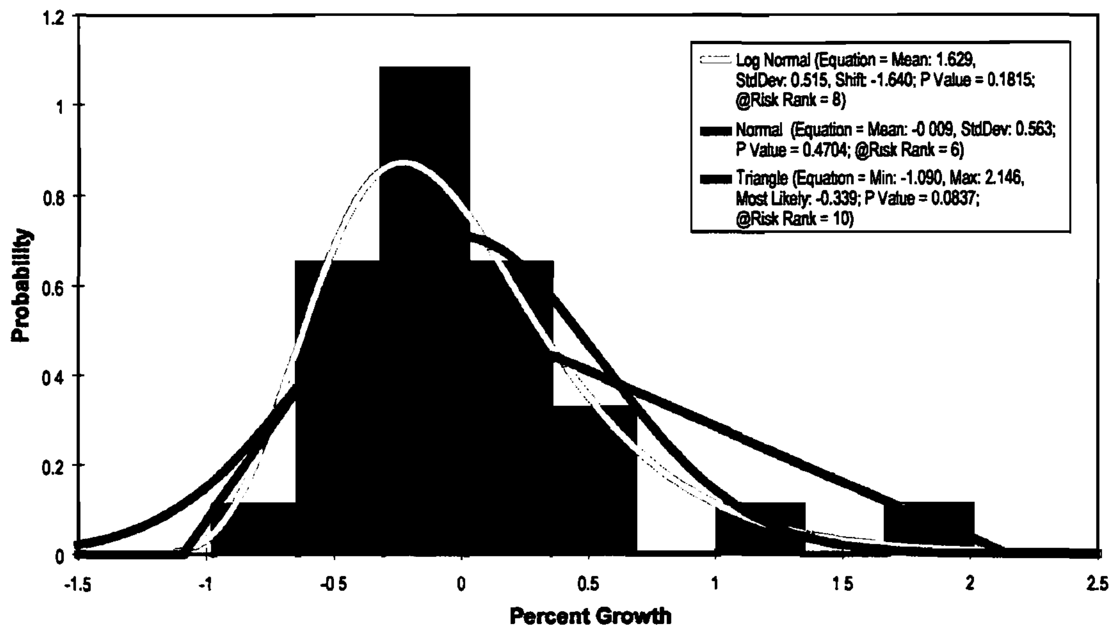

Figure I-22. Histogram with various distributions for Science Cost Growth from PDR to Launch

End of File