NORTHROP GRUMMAN
Information Technology

SCEA
The Society of Cost Estimating and Analysis

Rayleigh Curves – A Tutorial

Heather F. Chelson
Richard L. Coleman
Jessica R. Summerville
Steven L. Van Drew

SCEA 2004 – Manhattan Beach, CA
June 2004

Outline

- **Background**
- **Description**
- **Application**
- **The N-R Curve Generation Tool**
  - Risk Analysis considerations
- **Refining the Rayleigh after Program Start**
- **Fitting the N-R Curve in Mature Programs**
- **Conclusions**

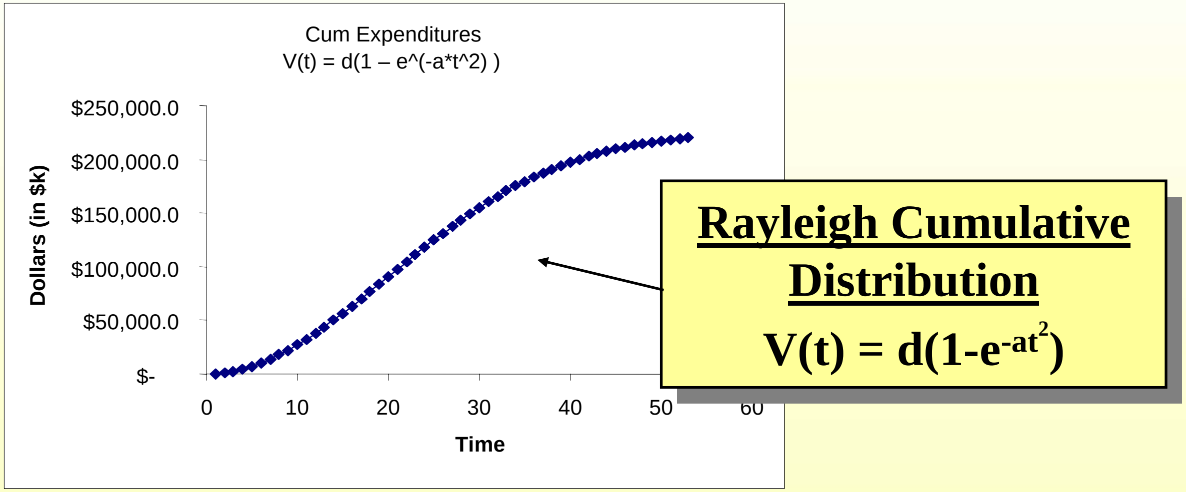
richard.coleman@ngc.com (703) 402-3702

Background

- **Studies done by Norden, Lee and others have shown that the cumulative costs of R&D projects, derived from earned value systems, typically follow the Rayleigh distribution[1] quite closely**

$$V(t) = d(1-e^{-at^2})$$

- **The Rayleigh distribution models the buildup, peak and taper of a development program's effort over time**
- **Using the Rayleigh curve, forecasting EACs, given sufficient earned value data, is a matter of predicting the d and a variables in the above equation to yield a value for $V(t_{final})$.**



1. *Norden-Raleigh Analysis: A Useful Tool for EVM in Development Projects*, **David Lee, Logistics Management Institute, The Measurable News, March 2002**

richard.coleman@ngc.com (703) 402-3702

Detailed Description

richard.coleman@ngc.com (703) 402-3702

Norden-Rayleigh Model

- **Cumulative distribution function for the Rayleigh:**

$$V(t) = d(1-e^{-at^2})$$

V(t) = Total effort expended

t = Time

d = Scale factor of the distribution

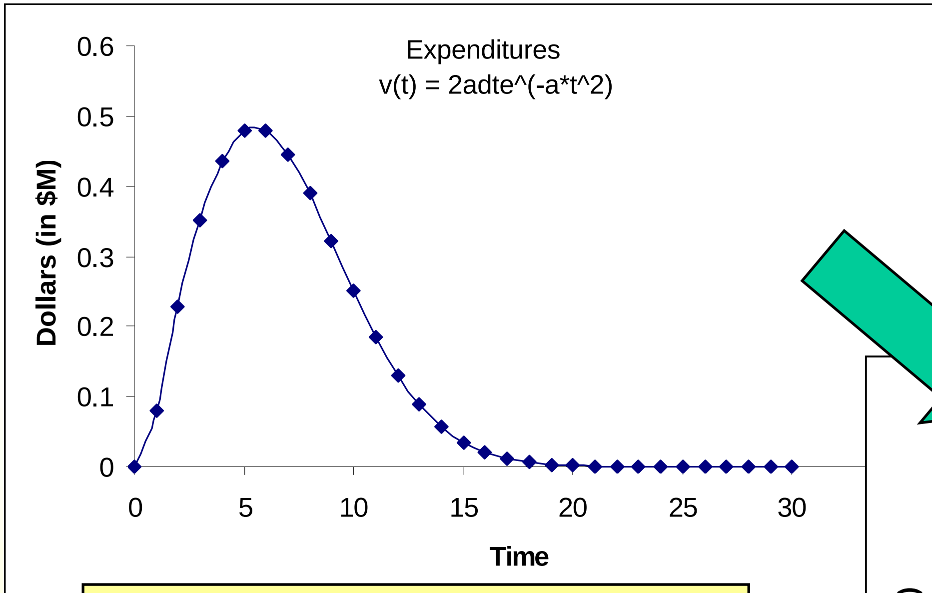
a = Shape parameter

- **Probability density function for the Rayleigh:**

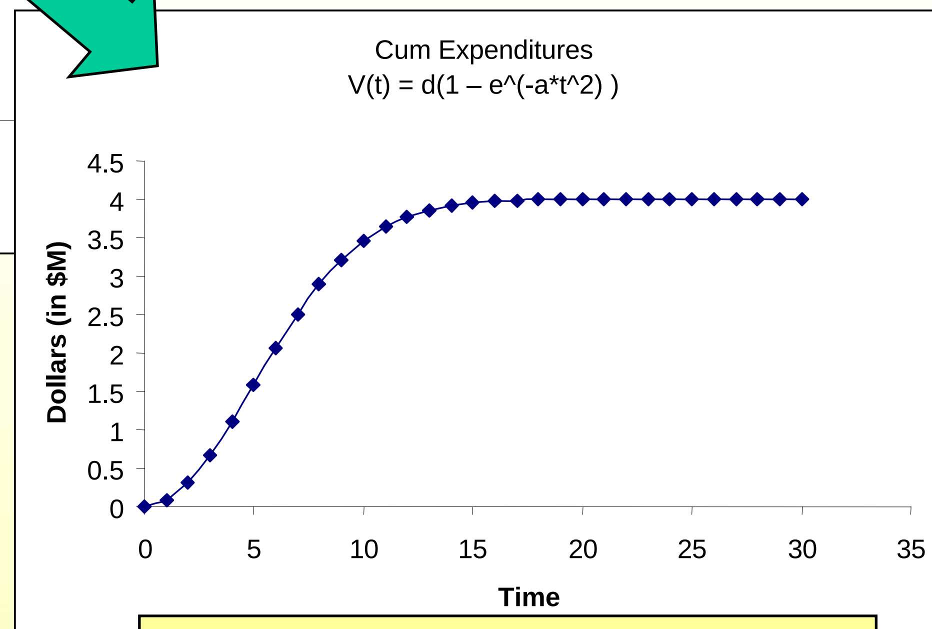
$$v(t) = 2adte^{-at^2}$$

richard.coleman@ngc.com (703) 402-3702

Rayleigh Curve – Use in Modeling Funding Profiles



Funding Profile Over Time



Cumulative Funding Over Time

richard.coleman@ngc.com (703) 402-3702

The Norden-Rayleigh Funding Model

- **Models time-phasing of expenditures for Development programs**
  - Given expenditures vs. time data, useful for forecasting
    - Cost-to-go
    - Time-to-go
  - Models typical programs that rapidly ramp-up labor efforts and then taper off
  - Ideally reflected in manufacturing programs as well as incremental software development efforts

richard.coleman@ngc.com (703) 402-3702

Application

richard.coleman@ngc.com (703) 402-3702

Application of the Rayleigh Curve

- **Valid tool for assessing funding and cost of Development programs**
  - Assessing funding profiles:
    - Rayleigh Model offers a standard of comparison for the reasonableness of a project's planned funding phasing
  - Assessing cost:
    - An assumed scale (*d*) and shape factor *(a)* can be used to build a profile
      - But uncertainties attached to the project end time, or t_f means that the Rayleigh Curve methodology cannot reasonably predict cost until there is sufficient earned value data to estimate *d* and *a*
- **Valid tool for generating an EAC**
  - Must have the following information
    - Computed *d* and *a* from the ACWP data already completed

richard.coleman@ngc.com (703) 402-3702

When the Rayleigh Model Does Not Apply

- **When the schedule contains a great deal of uncertainty**
- **When programs* are comprised of distinct sub programs with starts and stops, e. g.:**
  - When a contract funds more than one development program within the same funding profile
  - Software programs that release periodic versions or upgrades within the same funding profile

* If a program is an aggregation of sub-programs, and cannot be predicted *in toto*, it must be broken into independent component sub-programs, and the Rayleigh applied to each sub-program

richard.coleman@ngc.com (703) 402-3702

Benefits and “Endorsements”

- **Benefits**
  - Good cross check to EAC
  - Fast
- **The methodology is in use elsewhere**
  - AFCAA
  - OSD
  - ASC

richard.coleman@ngc.com (703) 402-3702

The N-R Curve Generation Tool

richard.coleman@ngc.com (703) 402-3702

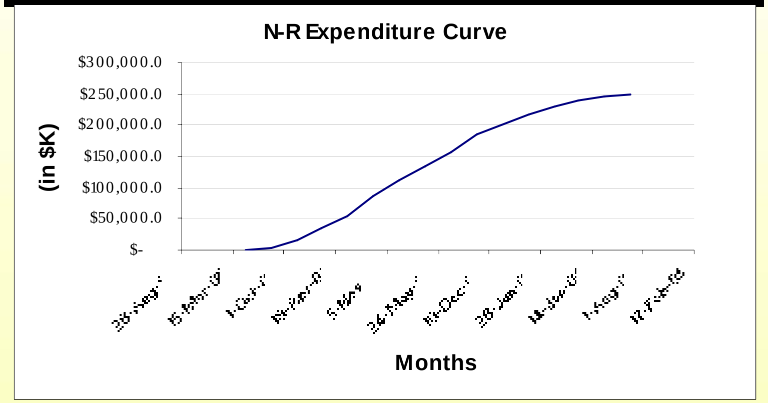
N-R Curve Generation Tool

- **This N-R Curve Generation Tool is a basic tool that can be used early in the program to generate a program's total funding profile**
  - Useable at outset to develop or check the planned funding profile
  - Usable throughout a program as a cross check or early indicator
    - Early in the program (before ~20% complete) the plot will provide a good cross check when plotted against the immature ACWP profile, and is an early indicator of trends
      - According to Christensen, *et al*, it is at 20% that a program stabilizes to a degree that the claim can be made that the Cum CPI will not change by more than 10% from its value at the 20% point.[1]
      - The 20% point is a forward looking point … the actual percent complete is unclear until later, but the thumb rule is still valid
    - This tool is also useful at any point in the program to provide a cross-check on EVM data that may appear suspect

INPUTS	
Norden-Rayleigh Tool Expenditure Curves	
Program Estimate	$ 250,000.0
Program Start Date	1-May-00
Program Completion Date	15-Jun-04
t_f	50
a (Shape Parameter)	0.0014
d (Scale Parameter computed as a cross check of total funding)	257,732

yellow cells are for input

OUTPUT



Double-click on the picture to launch the N-R tool.

1. *Is the CPI-Based EAC a Lower Bound to the Final Cost of Post A-12 Contracts?*, David S. Christensen, Ph.D., David A. Rees, Ph.D., The Journal of Cost Analysis and Management, Winter 2002.

Determining *a* and *d* (Early in the Program)

Early in the program (because the ACWP is immature), the pdf parameters – *a* and *d* – can only be "found" from the schedule variables. Below are the equations for calculating *a* and *d.*

$$V(t) = d(1 - e^{-at^2}),$$

$$\text{at } t_f,\ V(t_f) = d(1 - e^{-at_f^2})$$

Given $\mathbf{V(t_f) = .97d}$, solve for ***a***…

Because V(t) does not reach v_0 in finite time, the project's end time is usually[1] defined as the time at which:

$$V(t_f) = 97\% \text{ of } v_0,$$

or, $\mathbf{V(t_f) = .97d}$

1. Analyzing Development Programs' Expenditure with the Norden-Rayleigh Model, David Lee, 32nd ADoDCAS, February 1999, p21.

$$V(t_f) = d(1 - e^{-at_f^2})$$

$$.97d = d(1 - e^{-at_f^2})$$

$$.97 = (1 - e^{-at_f^2})$$

$$e^{-at_f^2} = .03$$

$$-at_f^2 = \ln(.03)$$

$$\mathbf{a = -\ln(.03) / t_f^2}$$

The authors recommend using this computation only as a rough cross check to the program plan, particularly for the curve generation.

A mismatch between this derivation of *d* and the program funding should be viewed as an indicator of schedule and funding misalignment

$$V(t) = d(1 - e^{-(-\ln.03/t_f^2)t^2})$$

$$\mathbf{d = V(t) / (1 - e^{-(-\ln.03/t_f^2)t^2})}, \text{ where } t_f \text{ is known}$$

Warning: SDD Completion Date is difficult to estimate, and therefore t_f is almost always unknown as is evidenced by the existence (in fact commonness) of schedule growth. This limits the reliability of the Norden-Rayleigh method until sufficient data are available.

richard.coleman@ngc.com (703) 402-3702

Use of the Curve Generator for Risk

- **The previous tool will produce a Norden-Rayleigh curve when program planning data are input**
  - Start date
  - End date
  - Total budget
- **A cross check of total funding is available, computed from t_f, or t_{final}, but it is not considered reliable**
- **The same tool can produce useful outputs for risk estimates**
  - If a risk estimate is done, in either cost or schedule or both, different values for end date and total funding will yield an alternative profile
  - Even if a formal risk analysis is not done, nominal (average) growth factors can be applied to yield a profile with “typical” growth

richard.coleman@ngc.com (703) 402-3702

Refining the Rayleigh after Program Start

richard.coleman@ngc.com (703) 402-3702

Refining the Raleigh Curve

- **As the program begins to gather stable ACWP data, the Rayleigh curve should be updated to reflect the improved availability of information**
- ***a* and *d* can be further refined by finding the peak of the funding profile**
  – Finding *a* and *d* in terms of the peak of the pdf (t_{peak}) firms up the value of *a* and *d*
    – Due to the previously noted volatility in schedules, t_{final} is a poor basis
    – *a* and *d* dependent on t_{final} should only be used when t_{peak} cannot be determined
  – *(derivation on following slide …)*

richard.coleman@ngc.com (703) 402-3702

Refining *a* and *d*

To determine when funding is at the max, we must find the point (t_p, or *t-peak*) at which the first derivative of the pdf is zero (standard math technique):

Start with the pdf

$$v(t) = 2adte^{-at^2}$$

Taking the first derivative

$$v'(t) = 2ad * [e^{-at^2} * t * (-2at) + e^{-at^2}]$$
$$= 2ad * (e^{-at^2} * -2at^2 + e^{-at^2})$$
$$= 2ade^{-at^2} * (-2at^2 + 1)$$
$$\text{Set } v'(t) = 0$$
$$0 = 2ade^{-at_p^2} * (-2at_p^2 + 1)$$

Solving, we get

$$t_p = 1 / \sqrt{2a}$$

So,

$$a = 1 / (2t_p^2)$$

← By definition, time is greater than 0, so *a* must be greater than 0.

And,

$$d = v(t) / 2t_p te^{-(1/2t_p^2)t^2} \textit{ or } d = V(t) / (1 - e^{-(1/2t_p^2)t^2})$$

← Solving for *d* in terms of t_p, since time is greater than 0 as is also v(t) [funding], so *d* must be greater than 0.

Computing the 2nd derivative yields a negative number (given that *a* and *d* are greater than 0), indicating that t_p is at the max point vs. a min point of the curve:

$$v''(t) = a^2dte^{-at^2}(8at^2-12),$$
$$\text{substitute } t_p = 1/(\text{sqrt}(2a)) \Rightarrow$$
$$v''(t) = -8a^2d/(\text{sqrt}(2ae))$$

richard.coleman@ngc.com (703) 402-3702

Fitting the N-R Curve in Mature Programs

Presented at the 2004 SCEA-ISPA Joint Annual Conference and Training Workshop - www.iceaaonline.com

richard.coleman@ngc.com (703) 402-3702

Fitting the N-R Curve in Mature Programs

- **After a program is 20% complete, earned value data should be sufficient to fit a Rayleigh distribution to the data**
  - The 20% point is not empirically demonstrated, but the authors believe that EACs are sufficiently stable at this point to use the method based on work by Christle, Abba, Christensen and others
- **The parameters *a* and *d* are found by fitting a curve to the data using least squares. This is difficult given that the equation has two unknowns.**

  - **Solutions: to best fit a Rayleigh curve to the earned value data, the analyst needs additional tools that will make these computations**

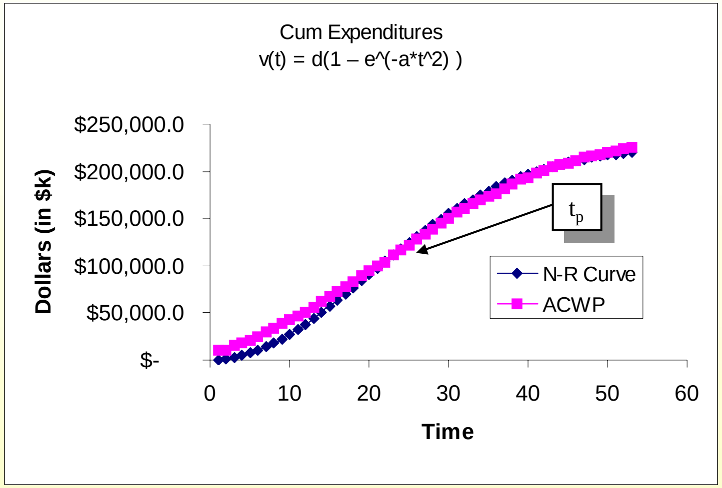
COTS software solutions:

***Rayleigh Analyzer*, Logistics Management Institute *Premium Solver Platform Versions 5.0 or 5.5*, Frontline Systems Inc.**

- Used with Microsoft Excel

***Solver DLL Platform*, Frontline Systems Inc.**

- Used with Visual Basic and C++



Warnings:

1) Excel Solver uses an algorithm that finds <u>local</u> optimal solutions based on the inputted start points for the decision variables (changing cells) in non-linear equations. The answers provided may not be the <u>global</u> optimal solutions.

2) The 20% point is a forward looking calculation. It may prove inexact, but is sufficient for use of the thumb rule

richard.coleman@ngc.com (703) 402-3702

Conclusions

richard.coleman@ngc.com (703) 402-3702

Conclusions

- **The Norden-Rayleigh model can be a valid tool for assessing performance (cost and schedules) of DoD Development programs and offers tests for the reasonableness of a project's planned earned value phasing**
  - Caveat: the reliability of the model is dependent on the maturity of the earned value data to estimate *a* and *d* (the shape and scale parameters)

A Summary of the Different Methodologies

	Beginning of program	Stabilized Program	Mature Program
ACWP data availablitity	Not available or insufficient	Initital data available	Mature, stable and available
Basis of *a* and *d*	*a* is based on an assumed schedule – the critical *t-final* – *d* is based on program plans and checked with *t-final*	*a* and *d* based on a known curve – the critical *t-peak* – to compute the curve	*a* and *d* found by fitting a curve to the data using least squares method
Concerns	Actual *t-final* is unknown due to the reality of schedule variability	Actual *t-peak* is difficult to determine until ACWP profile is well beyond the peak	Difficult because the equation has two unknowns (*a* and *d*)
Comments	Good for early planning	*t-peak* can be sketchy if determined too early	Needs lots of data (program past 20%)

richard.coleman@ngc.com (703) 402-3702

References (also see footnotes)

- *Analyzing Development Programs' Expenditure with the Norden-Rayleigh Model*, David Lee, 32nd ADoDCAS, February 1999
- *The Rayleigh Analyzer*, John Dukovich, Scott Houser, and David Lee, LMI Report At902C1, October 1999
- *Familiar Metric Management – Effort, Development Time, and Defects Interact*, Lawrence H. Putnam, Ware Myers, Quantitative Software Management, Inc.
- *Norden-Raleigh Analysis: A Useful Tool for EVM in Development Projects*, David Lee, Logistics Management Institute, The Measurable News, March 2002
- *ASC/FMC Rayleigh Curve Overview*, Ross Jackson, 60th ASC Industry Cost and Schedule Workshop, April 2003
- *Is the CPI-Based EAC a Lower Bound to the Final Cost of Post A-12 Contracts?,* David S. Christensen, Ph.D., David A. Rees, Ph.D., The Journal of Cost Analysis and Management, Winter 2002.

richard.coleman@ngc.com (703) 402-3702