Lattice Model of System Evolution

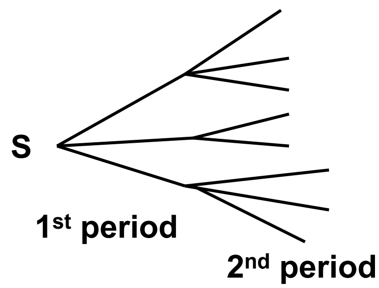
Richard de Neufville
Professor of Engineering Systems
and of
Civil and Environmental Engineering
MIT

Outline

- "Curse of Dimensionality"
- Binomial Lattice Model for Outcomes
  - Linear in Logarithms
  - Binomial lattice.xls
- Binomial Lattice Model for Probabilities
  - Normal distribution in logarithms
- Fitting to a known distribution
  - General: use pdf parameters to solve for u, d, p
  - "Financial": Assumptions and solution

System Evolution

- Consider system evolution over time
  - Starts at in a configuration (≡ a "state") S
  - In 1st period, could evolve into "i" states S_{1i}
  - In 2nd, each S_{1i} could evolve into more states…



Etc, etc…

"Curse of Dimensionality"

- Consider a situation where each state:
  - Evolves into only 2 new states…
  - Over only 24 periods (monthly over 2 years)
  - How many states at end?

ANSWER: 2, 4, 8 => $2^N = 2^{24}$ ~ 17 MILLION!!!

This approach swamps computational power

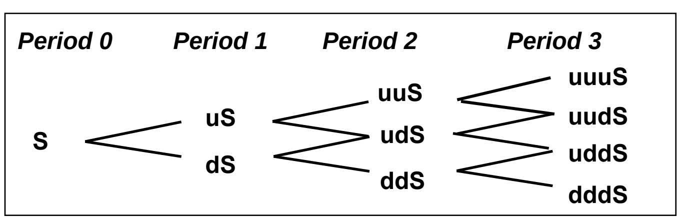
Binomial Lattice Model – 1st Stage

- **Assumes**
  - Evolution process is same over time (stationary)
  - Each state leads to only 2 others over a period
  - Later state is a multiple of earlier state
  - S => u S and d S (by convention, <u>u</u>p > <u>d</u>own)
- **For one period:**

What happens over 2nd , more periods?

Engineering Systems Analysis for Design
Massachusetts Institute of Technology
Richard de Neufville ©
Lattice Model Slide 5 of 48

Binomial Lattice: Several periods



- **States coincide**
  - path "up then down" => d(uS) = udS
  - same as "down then up" => u(dS) = udS
- **States increase linearly (1,2, 3, 4 => N) not exponentially (1, 2, 4, 8…) = 2^N**
  - After 24 months: 25 states, not 17 million

Engineering Systems Analysis for Design
Massachusetts Institute of Technology
Richard de Neufville ©
Lattice Model Slide 6 of 48

Main Advantage of Binomial Model

- Eliminates "Curse of Dimensionality"
- Thus enables detailed analysis
  - Example: A binomial model of evolution every day of 2 years would only lead to 730 states, compared to ~17 million states resulting from 'decision tree' model of monthly evolution
- The jargon phrase is that Binomial is a recombinatiorial model…

Non-negativity of Binomial Model

- The Binomial Model does not allow shift from positive to negative values: lowest value (d^n S) is always positive
- This is realistic – indeed needed -- in many situations:
  - Value of an Asset; Demand for a Product; etc.
- Is non-negativity always realistic?
- NO! Contrary to some assumptions
  - Example: company profits! Easily negative

Path Independence: Implicit Assumption

Pay Attention – Important point often missed!

- **Model Implicitly assumes "Path Independence"**
  - Since all paths to a state have same result
  - Then value at any state is independent of path
  - In practice, this means nothing fundamental happens to the system (no new plant built, no R&D , etc)

When is "Path Independence" OK?

- **Generally for Market items (commodities, company shares, etc). Why?**
  - Random process, no memory….
- **Often not for Engineering Systems. Why?**
  - If demand first rises, system managers may expand system, and have extra capacity when demand drops.
  - If demand drops then rises, they won't have extra capacity and their situation will differ
  - Process – and result -- then depends on path!

Example for "path independence"

- Consider the case when gasoline prices rose from $2 to nearly $5 / gallon
- Is consumer behavior "path independent" ?
- Depends…
  - Yes, if individuals behave as before, when prices dropped back down to original level
  - No, if shock of high price gas meant they switched to use of public transport, sold car and cut driving, etc.

Easy to develop in Spreadsheet

- Easy to construct by filling in formulas
- Class reference: Binomial lattice.xls
  - Allows you to play with numbers, try it
- Example for: S = 100; u = 1.2 ; d = 0.9

OUTCOME LATTICE

100.00	120.00	144.00	172.80	207.36	248.83	298.60
	90.00	108.00	129.60	155.52	186.62	223.95
		81.00	97.20	116.64	139.97	167.96
			72.90	87.48	104.98	125.97
				65.61	78.73	94.48
					59.05	70.86
						53.14

Relationship between States

- The relative value between a lower and the next higher is constant = u /d

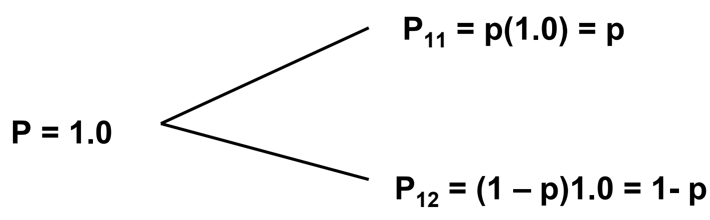
  S => uS and dS ; Ratio of uS /dS = u/d

- Thus results for 6^{th} period, u/d = 1.2/.9 = 1.33

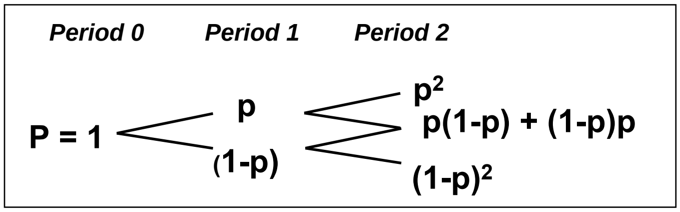
	Step	(u/d)exp[step]	0utcome/lowest
298.60	6	5.62	5.62
223.95	5	4.21	4.21
167.96	4	3.16	3.16
125.97	3	2.37	2.37
94.48	2	1.78	1.78
70.86	1	1.33	1.33
53.14	0	1.00	1.00

Application to Probabilities

- Binomial model can be applied to evolution of probabilities
- Since Sum of Probabilities = 1.0
  - Branches have probabilities: p ; (1- p)



Important Difference for Probabilities



- A major difference in calculation of states:
  - Values are not “path independent”
  - Probabilities = Sum of probabilities of all paths to get to state

Spreadsheet for Probabilities

- Class reference: Binomial lattice.xls
- Example for: p = 0.5 ; (1 –p) = 0.5
- => Normal distribution for many periods

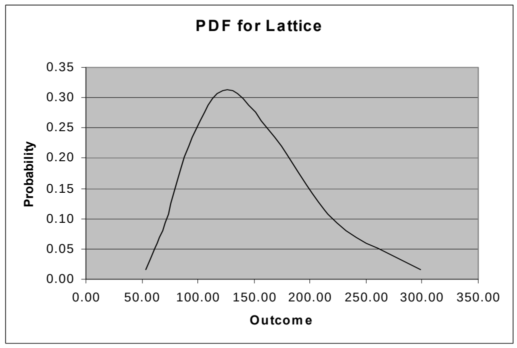
PROBABILITY LATTICE

1.00	0.50	0.25	0.13	0.06	0.03	0.02
	0.50	0.50	0.38	0.25	0.16	0.09
		0.25	0.38	0.38	0.31	0.23
			0.13	0.25	0.31	0.31
				0.06	0.16	0.23
					0.03	0.09
						0.02

Outcomes and Probabilities together

- Applying Probability Model to Outcome Model leads to Probability Distribution on Outcomes
- In this case (u = 1.2 ; d = 0.9; p = 0.5):

AXES	
Outcome	Prob
298.60	0.02
223.95	0.09
167.96	0.23
125.97	0.31
94.48	0.23
70.86	0.09
53.14	0.02

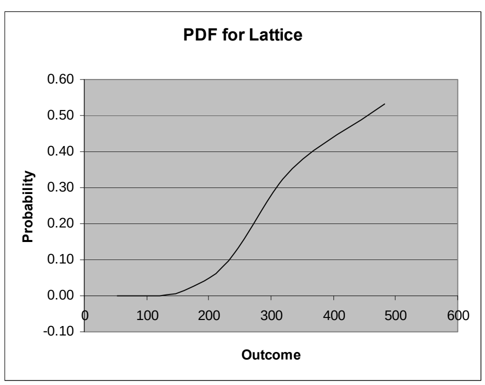


Engineering Systems Analysis for Design
Massachusetts Institute of Technology

Richard de Neufville ©
Lattice Model Slide 17 of 48

Many PDFs are possible

- For example, we can get "triangular right" with u = 1.3 ; d = 0.9 ; p = 0.9

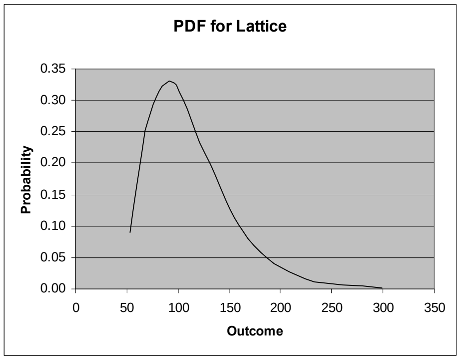


Engineering Systems Analysis for Design
Massachusetts Institute of Technology

Richard de Neufville ©
Lattice Model Slide 18 of 48

Many PDFs are possible

- … or a "skewed left"
  with u = 1.2 ; d = 0.9 ; p = 0.33



Let's try it

- An interlude with Binomial lattice.xls

Calibration of Binomial Model

- Examples show that Binomial can model many PDF
- Question is:

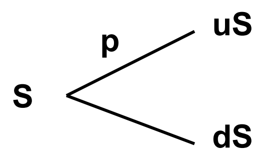
  Given actual PDF, what is Binomial?

Binomial Calibration: Outline

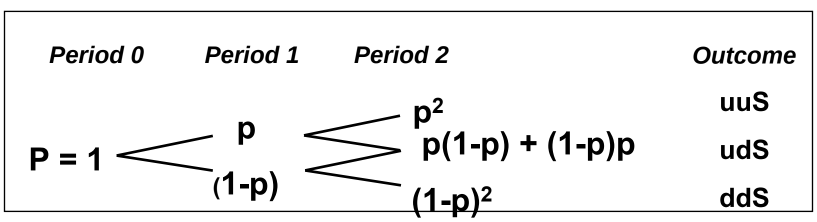
- General Procedure
  - One example
- Financial or Log Normal Approach
  - Concept
  - Procedure
  - Application to detailed data
  - Application to assumed conditions
- Appendix
  - Theoretical justifications

General Procedure for Calibrating Binomial to Data

- 3 unknowns: u, d, and p
- Need 3 conditions, for example:
  - Mean, Variance, Range
  - Most Likely, Maximum, Minimum
- Use more stages for more accuracy



A Simple Example



Average or Most likely =
$= S\,[p^2u^2 + 2p(1-p)\,ud + (1-p)^2d^2]$

Maximum $= S\,[uu] \Rightarrow u = \sqrt{(Max/S)}$

Minimum $= S\,[dd] \Rightarrow d = \sqrt{(Min/S)}$

Numerical Results

Projected demand for new technology

Starting = 100 **Maximum = 500**

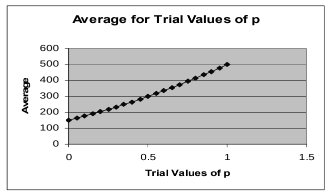
Most likely = 200 **Minimum = 150**

Solution

u = 2.24

d = 1.22

p ~ 0.20



Binomial expansion

100	224	500
	122	274
		150

"Financial" or Log Normal Approach

- **The standard used for financial analysis**
  - **Justified on Theory concerned with markets**
- **Can be used for engineering analysis**
  - **Justified on Practical Basis as a reasonable approach (theory based in market economics does not apply)**

"Financial" or Log Normal Concept

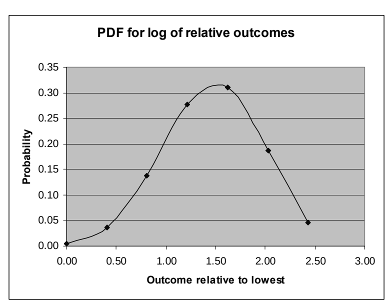
- Idea concerns distribution of deviations from trend line
  - Assumed that percent deviations have a "Normal" pdf
  - This is a Geometric Distribution
- This contrasts with assumption used in standard regression analysis
  - Deviations (or measurement errors) from trend are "Normal" (random)

Engineering Systems Analysis for Design
Massachusetts Institute of Technology
Richard de Neufville ©
Lattice Model Slide 27 of 48

Log Normal Illustration

- pdf of log of outcomes is "Normal"
- "natural logs of outcomes are normally distributed"
- For: u = 1.2 ; d = 0.8 ; p = 0.6
- Note linearity of ln (outcomes):

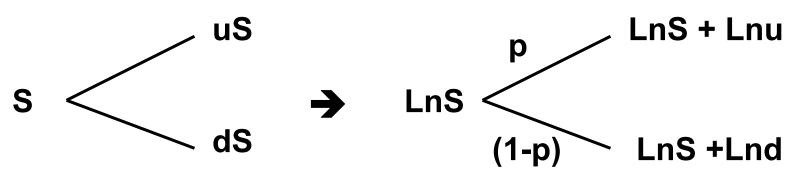
outcome	Ln[(out/low)/LN(low)	**LN(outcome/lowest))**
298.60	6	2.43
199.07	5	2.03
132.71	4	1.62
88.47	3	**1.22**
58.98	2	0.81
39.32	1	0.41
26.21	0	0.00



Engineering Systems Analysis for Design
Massachusetts Institute of Technology
Richard de Neufville ©
Lattice Model Slide 28 of 48

Calibration for Log Normal

- **Transform data on outcomes to logs**
- **Solve – as indicated on next slides**



Three Conditions to be met

- **Average increase over period:**

  $v\Delta T = p \, Lnu + (1 – p) \, Lnd$

- **Variance of distribution**

  $\sigma^2 \, \Delta T = p \, (Lnu)^2 + (1- p) \, (Lnd)^2 – [p(Lnu) + (1-p)Lnd]^2$

  = weighted sum of squared observations

  -- [(average) squared]

  = second moment around mean

- **Normal pdf of log outcomes**
  - **Up and down variation equally likely, so**
  - **Lnu = - Lnd equivalent to u = 1/d**

- **This has 3 equations and 3 unknowns (u, d, p)**

Solution for u ; d ; p

The previous equations can be solved, with a lot of "plug and chug" to get

$u = e \exp(\sigma \sqrt{\Delta t})$

$d = e \exp(-\sigma \sqrt{\Delta t})$

$p = 0.5 + 0.5 (v/\sigma) \sqrt{\Delta t}$

The calculated values can be used directly

Notice that these formulas imply that v and σ have been determined from other data, such as an analysis over many periods

Practical Approaches

1. Application to detailed data

2. Use for assumed conditions (as for new technologies, for which actual data not available)

Solving with Actual PDF Data

Two Elements of Observational (or Assumed) Data – Variance and Average Trend

- **Variance of PDF = σ^2 = (standard deviation)^2**
- **Average trend, v, generally assumed to be growing at some rate per period: $S_T = S\, e^{vT}$**
- **Rate depends on length of period: 12%/year = 1%/month etc**
- **v and σ expressed in terms of <u>percentages!</u>**

Baseline Estimation Procedure

- **Parameters v and σ can be derived statistically using observations over time (e.g., of oil price)**
- **v, the average rate of exponential growth, e^{vt}, is the best fit of LN(data) against time**
  - **Since $S_T = S\, e^{vT}$; $vT = \ln(S_T)$ - constant**
- **σ, standard deviation, defined by differences between the observations and average growth**

NOTE: using past data assumes future ~ past. May be better to base v and σ on direct estimates of future from knowledge or theory

Three Conditions to be met

- **Average increase over period:**
  $v\Delta T = p\ Lnu + (1 - p)\ Lnd$
- **Variance of distribution**
  $\sigma^2\ \Delta T = p\ (Lnu)^2 + (1 - p)\ (Lnd)^2 - [p(Lnu) + (1-p)Lnd]^2$
  = weighted sum of squared observations
  -- [(average) squared]
  = second moment around mean
- **Normal pdf of log outcomes**
  - Up and down variation equally likely, so
  - $Lnu = - Lnd$ equivalent to $u = 1/d$
- **This has 3 equations and 3 unknowns (u, d, p)**

Example Solution for u ; d ; p

- **Assume that S = 2500 (e.g., $/ton of Cu Fine)**
  $v = 5\%$ $\sigma = 10\%$ $\Delta t = 1$ year
- **Then**

$$u = e\ \exp(\sigma \sqrt{\Delta t}) = e\ \exp(0.1) = 1.1052$$

$$d = e\ \exp(-\sigma \sqrt{\Delta t}) = 0.9048 = (1/u)$$

$$p = 0.5 + 0.5\ (v/\sigma) \sqrt{\Delta t} = 0.75$$

Note: everything varies with Δt

Using Assumed Conditions

- In Design, we may not have historical data from which we can derive v and σ
- May have forecasts or estimates of future states, such as demand for a product
- For example, suppose our estimate is that demand would grow 20% +/- 15% in 5 years

How do we deal with this?

Dealing with Assumed Conditions (1)

- First, keep in mind that v and σ are <u>yearly</u> rates. If you use any other period, you must adjust accordingly.
- Given 20% growth over 5 years, v ~ 4%
- [Strictly, the rate is lower, since process assumes exponential growth. However, the accuracy implied by 20% growth does not justify precision beyond 1st decimal place]

Dealing with Assumed Conditions (2)

- σ can be estimated in a variety of ways

- Reasoning that uncertainty grows regularly, then +/- 15% over 5 years => +/- 6.7% in 1 year
- As follows: $5 (\sigma_1)^2 = (\sigma_5)^2 \Rightarrow \sigma_1 = \sigma_5 / \sqrt{(5)}$
  [assuming a process without memory]

- With 2 observations, a statistical estimate for σ is somewhat speculative. Within the accuracy of this process, however, the assumptions in the forecast imply $\sigma \sim 6.7\%$

Estimates of p

- Estimates of v and σ may present a problem…

- With high growth and low variance, for example v = 5% and σ = 3%. If we insert these values in slide 31, using $\Delta t = 1$
  $p = 0.5 + 0.5 (v/\sigma) \sqrt{\Delta t} \Rightarrow p > 1.0$!!!
  This is impossible. What to do?

- Solution: use shorter time, such as 3 months [Δt = ¼], where v and σ also scale down.
- More broadly, accuracy increases as step size, Δt , is smaller

Broader extensions

- Normal processes are also common in nature, and can in any case be good approximations
- Also, economic model may be relevant in any engineering situations. This needs to be verified, and cannot simply be assumed.

Use of Lattice Model in practice ?

- Lattice Model has limitations (see future "lattice vs. decision analysis" presentation)
- Most important of these are:

1. Non-stationary systems, specifically "jumps" when there is some kind of "trend breaker" – such as new President elected, new regulations passes, etc.
2. Multiple decisions or flexibilities to be considered simultaneously

Summary

- Lattice Model similar to a Decision Tree, but…
  - Nodes coincide
  - Problem size is linear in number of periods
  - Values at nodes defined by State of System
  - Thus "path independent" values
- Lattice Analysis widely applicable
  - With actual probability distributions
  - Accuracy depends on number of periods -- can be very detailed and accurate
- Reproduces uncertainty over time to simulate actual sequence of possibilities

Appendix: Note on Solution to Financial, Log Normal Binomial

- System of 3 equations and 3 unknowns is not linear, so a unique solution is not guaranteed
- Solution presented works
- Note that the "standard finance" solution (using real probabilities) -- Arnold and Crack
  - matches the moments for v and σ
  - in the limiting distribution as
  - The number of steps => infinity and
  - the Binomial => Normal Distribution

Assumptions to Justify Log Normal Approach

1. Based on Theory of how markets function
   -- basis for financial analysis

2. Broader extensions

3. Based on Practical Assumptions

Theory about how markets function

A 3-phase argument, from point of view of investors -- different from managers of individual projects

1. Investors can avoid uncertainties associated with individual Projects – they can spread their money over many projects (i.e., diversify so risks cancel out). Thus only looks at market risk

2. Markets are efficient, have "full information". This is a sufficient (but not necessary) condition for error to be random or "white noise"

3. Randomness in pdf of % changes right description

When are above assumptions reasonable?

- When markets exist – does not apply for many engineering projects
- When markets are "efficient" [which means that there is continuous activity and no one has special, "inside" information] -- again does not apply to most engineering projects (and often not to actual markets)
- When changes can be described in % terms – not when outcomes can go from positive to negative (as for profits, not for prices)

Use of Normal Distribution in practice ?

- Assumption about randomness of percent changes may be acceptable
- A "standard" procedure, well documented – Not a good justification!
- Alternative is to chose a combination of u, d, p that matches assumed form of pdf