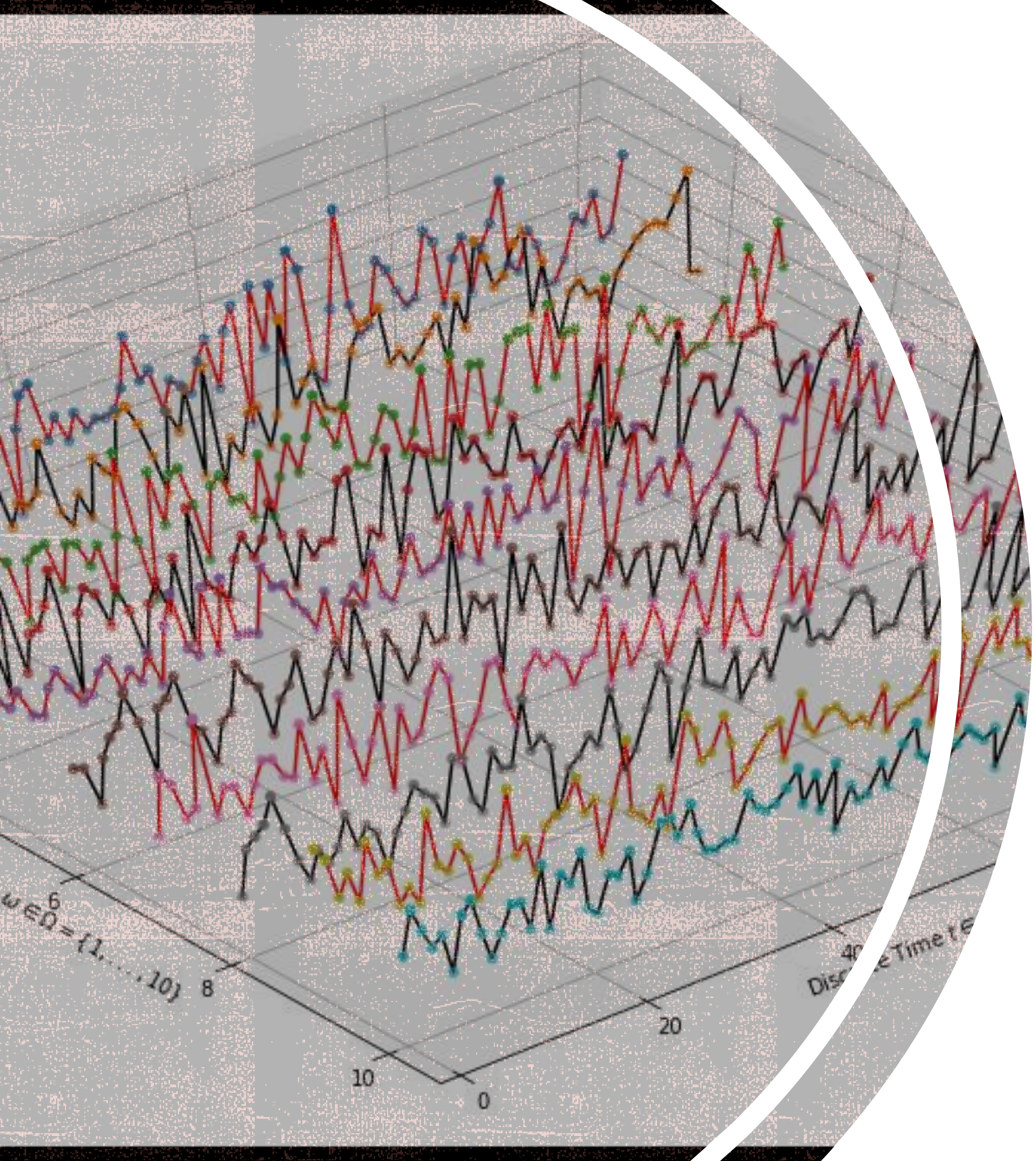


STOCHASTIC PROCESSES IN FINANCE

Video Lecture VL06

Valuation for Financial Engineers 6103

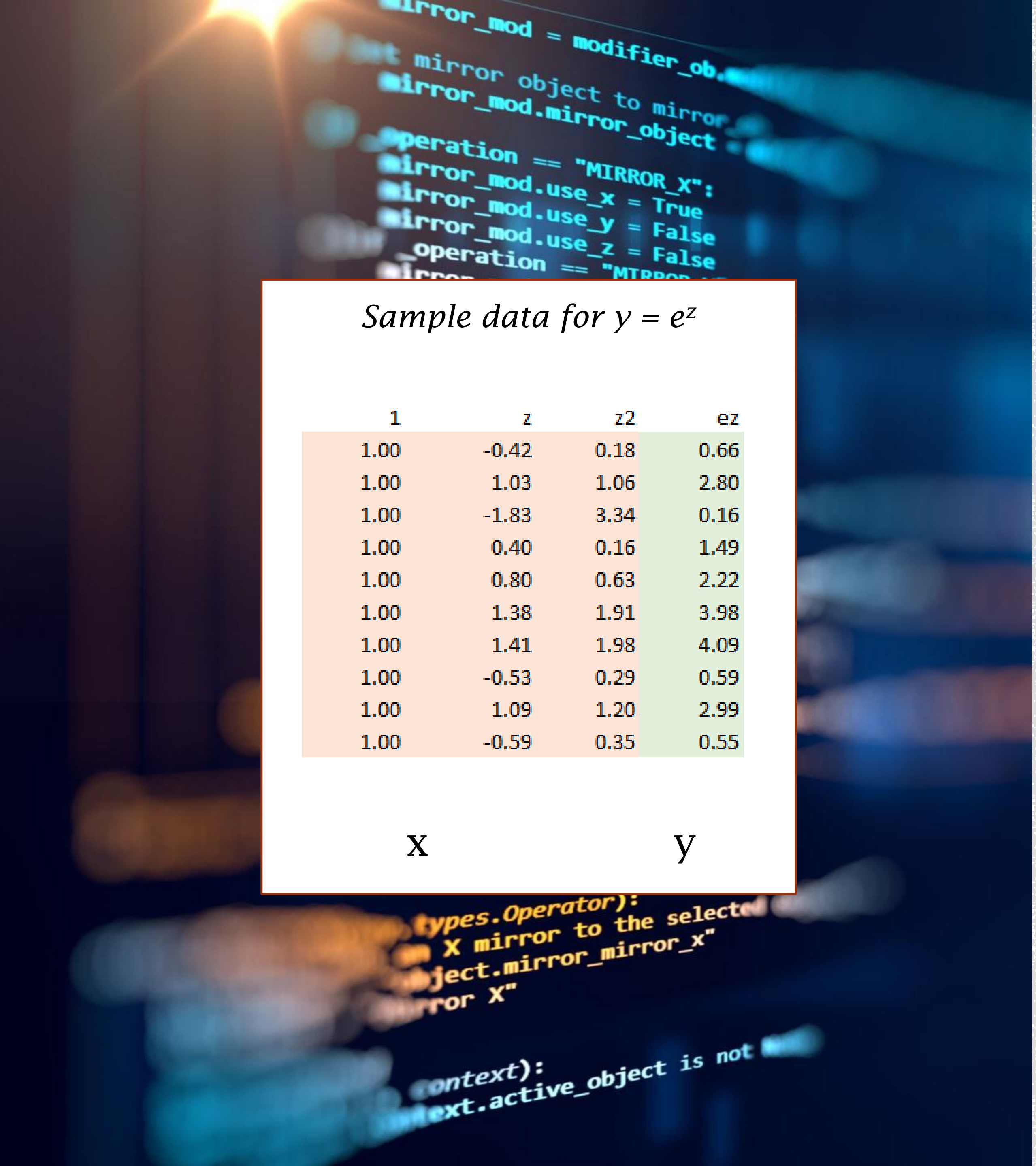
Professor David Shimko

TODAY'S TOOLKIT

- Review of simulation and simulation errors
- Four types of stochastic processes
  - GBM: Exponential growth random walk
  - ABM: Linear growth random walk
  - MR: Mean-reverting random walk
  - JUMP: Applied to any other random walk
- Forecast statistics
- Manipulation of stochastic processes
- Parameterization of stochastic processes
- Optimization

REVIEW OF SIMULATION

- Any random variable can be simulated by the function $F^{-1}(u)$, where F^{-1} is the inverse CDF of u, and u is a uniformly distributed random variable
- To simulate the value of a function of number of uncorrelated random variables (any distributions), follow these steps
  - Generate N values and calculate the mean
  - Calculate the standard deviation, and divide by $\sqrt{N}$ to get the standard error
  - Report the mean as your estimate and your “maximum error” as 2 times the standard error
  - Increase the sample size if more precision is needed
- To correlate normally distributed random variables, use independent random simulations and multiply the data vector by the transpose of the Cholesky decomposition of the desired covariance matrix
  - Correlation is only defined for normal random variables
- To correlate random variables of any distribution, convert correlated normal random numbers to uniform distributions, and then use $F^{-1}(u)$ to produce the desired random variables (this is the Gaussian copula)



Sample data for $y = e^z$

1	z	z2	ez
1.00	-0.42	0.18	0.66
1.00	1.03	1.06	2.80
1.00	-1.83	3.34	0.16
1.00	0.40	0.16	1.49
1.00	0.80	0.63	2.22
1.00	1.38	1.91	3.98
1.00	1.41	1.98	4.09
1.00	-0.53	0.29	0.59
1.00	1.09	1.20	2.99
1.00	-0.59	0.35	0.55

x y

REVIEW OF SIMULATION ERRORS: CONTROL VARIATE

- In an ordinary simulation, if the simulated values are given by a vector y, calculate the standard deviation estimate for y and divide y the square root of the sample size
- In a (linear) control variate situation
  - Define the data matrix X as the pink box to the right
  - Let $b' = (b_0 \quad b_1 \quad b_2)$ be the regression coefficients estimated from the simulation results
  - Let $X_h' = (1 \quad 0 \quad 1)$, the expected values of the columns in the example
- Then the estimated mean of y is $\mathrm{E}[y] = X_h' b$.
- The variance of the expectation is $Var(\mathrm{E}[y]) = s_e^2 X_h'(X'X)^{-1} X_h$
  - This is called the variance of the mean response
  - Its square root is the standard error of the estimate of $\mathrm{E}[y]$

STOCHASTIC PROCESSES IN FINANCE

Arithmetic random walk

- Continuous time: Arithmetic Brownian Motion
- Used for spreads, interest rates, some currencies
- Can become negative

Geometric random walk

- Continuous time: Geometric Brownian Motion
- Used for asset prices
  - Constant expected log returns and standard deviations
  - Positively skewed outcomes
  - Never reaches zero

Mean reversion

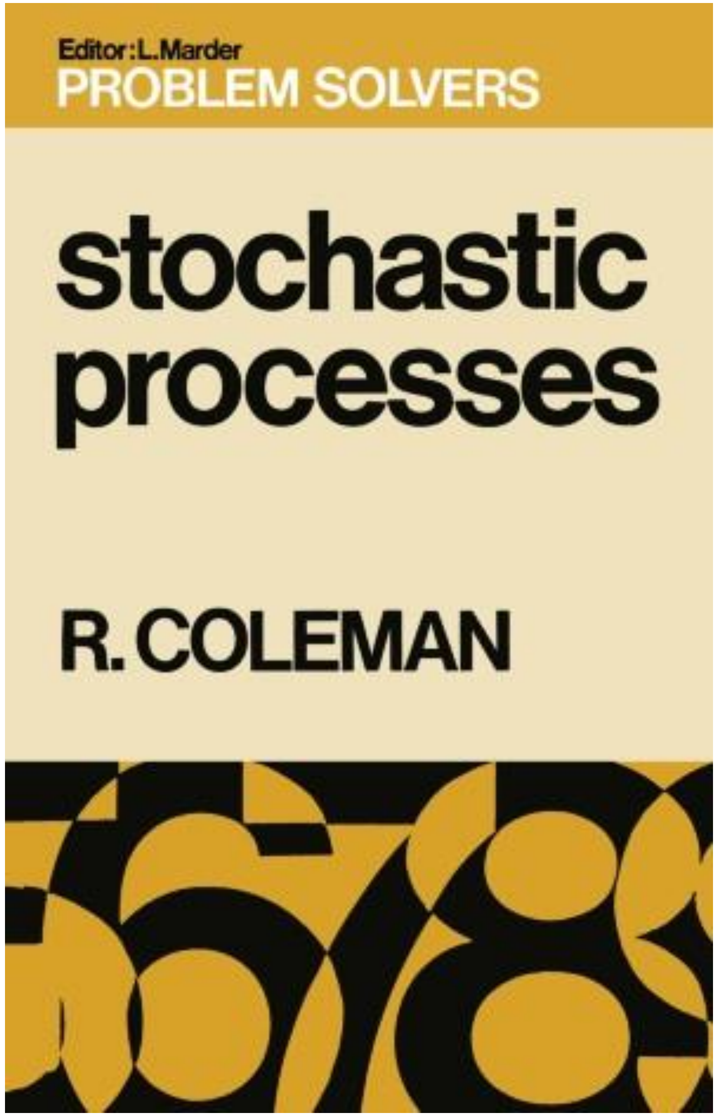
- Continuous time: Ornstein-Uhlenbeck process
- Used for prices of non-assets
  - Interest rates, Commodities, Foreign exchange
- Negative values occur but are tempered by mean reversion if the long run mean is positive

Jumps

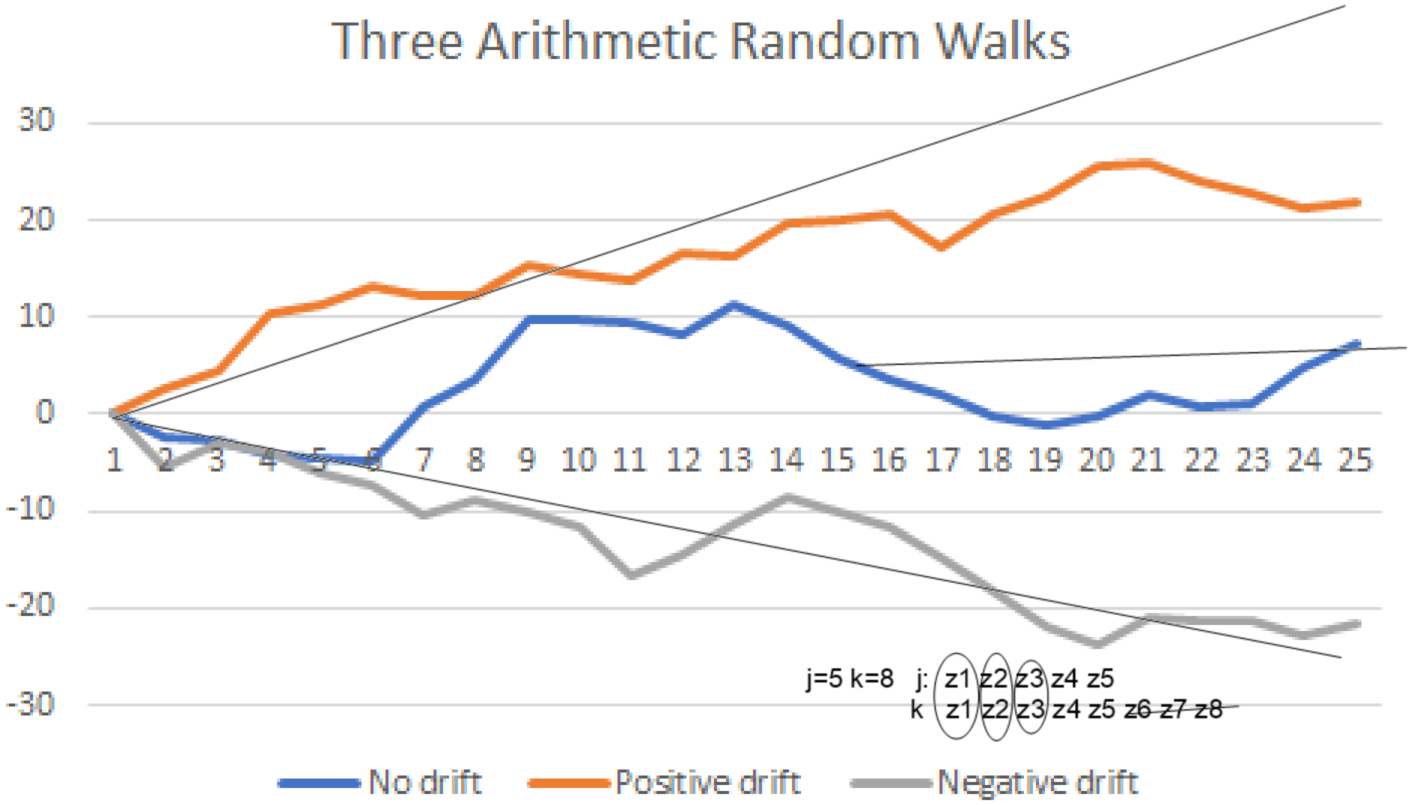
- Used with other models to superimpose discrete, random events

WHAT IS A STOCHASTIC PROCESS?

- The word *stochastic* is jargon for *random*. A *stochastic process* is a system which evolves in time while undergoing chance fluctuations. We can describe such a system by defining a family of *random variables*, $\{X_t\}$, where X_t measures, at time t, the aspect of the system which is of interest. For example, X_t might be the number of customers in a queue at time t. As time passes, customers will arrive and leave, and so the value of X_t will change. At any time t, X_t takes one of the values 0, 1, 2,…; and t can be any value in a subset of $(-\infty, \infty)$, the infinite past to the infinite future.
  - R. Coleman, *Stochastic Processes*
- Based on their mathematical properties, stochastic processes can be grouped into various categories, which include random walks,[33] martingales,[34] Markov processes,[35] Lévy processes,[36] Gaussian processes,[37] random fields,[38] renewal processes, and branching processes.[39] The study of stochastic processes uses mathematical knowledge and techniques from probability, calculus, linear algebra, set theory, and topology[40][41][42] as well as branches of mathematical analysis such as real analysis, measure theory, Fourier analysis, and functional analysis.[43][44][45] The theory of stochastic processes is considered to be an important contribution to mathematics[46] and it continues to be an active topic of research for both theoretical reasons and applications.[47][48][49]
  - *Wikipedia*



ARITHMETIC RANDOM WALK DEFINITION



S1 = S0 + mu + sigma x z1

S2 = S1 + mu + sigma x z2

S2 = S0 + 2*mu + sigma*(z1 + z2)

Sj = S0 + j*mu + sigma*(z1 + z2 + ... zj)

E[Sj|S0] = S0 + j*mu

Var[Sj|S0] = sigma^2 * var(z1 + z2 +...+ zj)

= j*sigma^2

Cov(Sj,Sk|S0) =
sigma^2*min(j,k)

- The stochastic process has a given starting point, e.g. S_0, the stock price at time 0
- The increment in the process is a normal random variable with mean μ and standard deviation σ
- $S_{t+1} = S_t + \mu + \sigma \cancel{z_1}$ z(t+1)

  independent values of z over time
  result is normally distributed
- The parameter μ is called the *drift* of the process

  sigma = volatility
  sigma * z1 = innovation, shock, residual, unexpected part

DISTRIBUTION OF ARITHMETIC RANDOM WALK OUTCOMES

- How can we find the distributions of future values of a process?
- We have $S_{t+1} = S_t + \mu + \sigma z_1$
- What is $S_{t+2} =$ ________________________________?
- Then $S_{t+j} = S_t + j\mu + \sigma \sum_{i=1}^{j} z_{t+i}$
- Since the z valued are iid standard normal, the sum in S_{t+j} is also normally distributed
  - What is the mean of S_{t+j} ? Its variance ?
  - What is the covariance of S_{t+j} with S_{t+k} ?

ARITHMÉTIC RANDOM WALK

- $S_{t+1} = S_t + \mu + \sigma z_1$
- $S_{t+j} = S_t + j\mu + \sigma \sum_{i=1}^{j} z_{t+i}$
- $E[S_{t+j}|S_t] = S_t + j\mu$
- $Var[S_{t+j}|S_t] = j\sigma^2$
- $Cov[S_{t+j}, S_{t+m}|S_t] = \sigma^2 \min(j, m)$

Linearly growing perpetuity?

Used for spreads; also a foundation for other models

ex. ABM can become negative

don't want to use for stocks

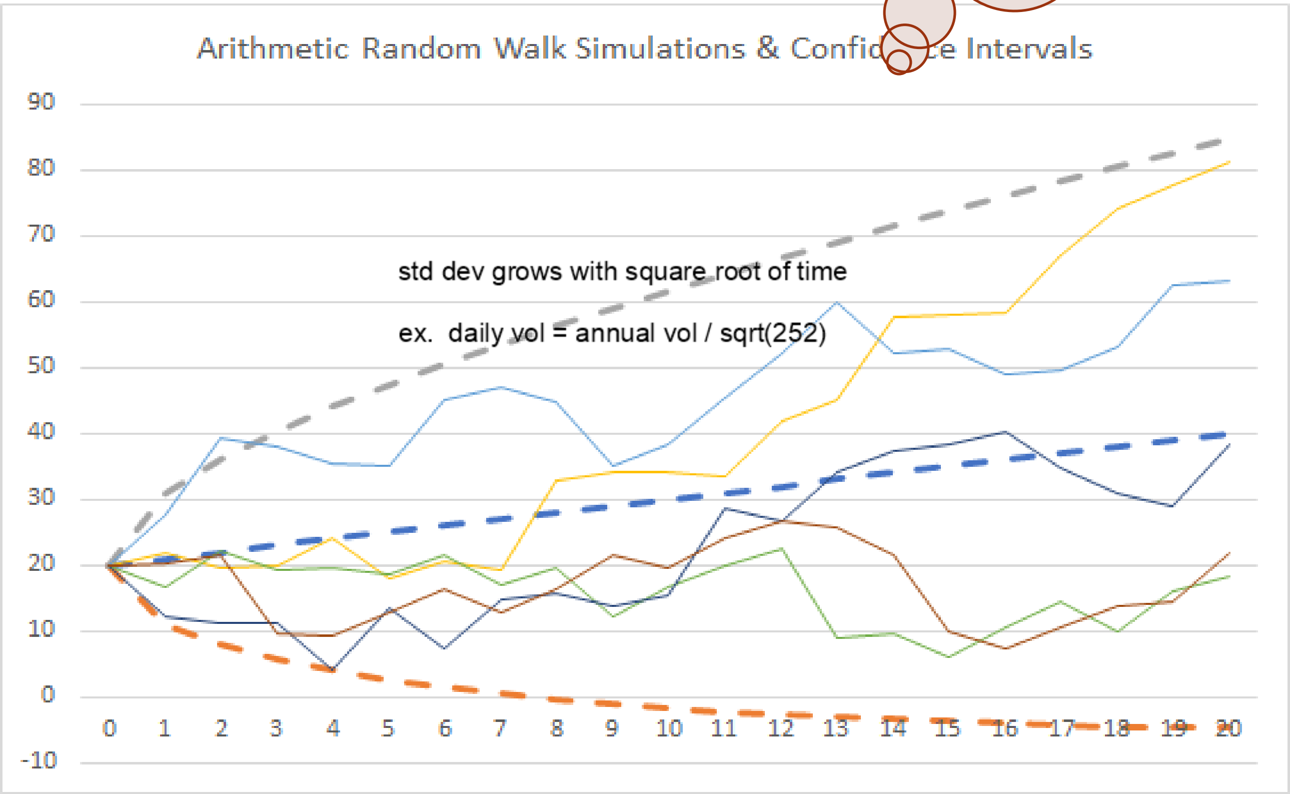
(stocks have limited liability)

Oil futures spread = Brent - WTI

Brent = north sea crude oil

WTI = West Texas Intermediate Light Sweet Crude

Pairs trading -



COVARIANCE MATRIX FOR ARW OUTCOMES

- $Cov[S_{t+j}, S_{t+m}|S_t] = \sigma^2 \min(j, m)$
- What is its Cholesky decomposition j∈{1..m}?

Covariance matrix in yellow

	1	2	3	...	m
1	1	1	1	...	1
2	1	2	2	...	2
3	1	2	3	...	3
...	...	...	...	...	...
m	1	2	3	...	m

Cholesky decomposition

- What can you use the Cholesky decomposition for?

  simuate stochastic process as a vector instead of a sequence of random numbers

CLASS EXERCISE

- Generate a 24-month arithmetic random walk for *the natural log of* Amazon stock prices with these parameters in Excel (α and σ are annual). Show at least six simulations on a graph.
  - S=1800, α=20%, σ=25%
- Also show the mean and confidence intervals

A STOCHASTIC PROCESS WITH NORMALLY DISTRIBUTED RETURNS

- Helpful to compute *log* returns first
  - the difference in logs relies on strictly positive prices

drift vol

  - $lnS_{t+1} = lnS_t + \left(\alpha - \frac{1}{2}\sigma^2\right) + \sigma z_{t+1}$

St+j can never reach zero or negative numbersw

  - Compare this to the Arithmetic random walk equation:
    - $S_{t+1} = S_t + \mu + \sigma z_1$

exponentiate

- Now we can compute $S_{t+1} = S_t e^{\left(\alpha - \frac{1}{2}\sigma^2\right) + z_{t+1}\sigma}$

- Then
  - $E[S_{t+1}] = S_t e^{\left(\alpha - \frac{1}{2}\sigma^2\right)} E[e^{\sigma z_{t+1}}] = S_t e^{\left(\alpha - \frac{1}{2}\sigma^2\right)} e^{\left(\frac{1}{2}\sigma^2\right)} = S_t e^{\alpha}$ exponential growth

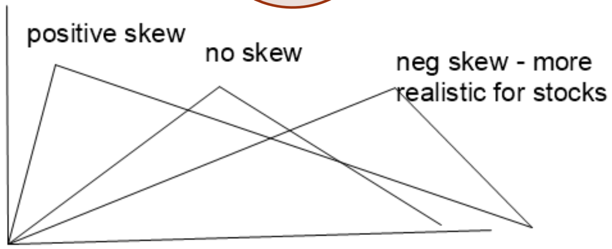
Recall that if X is normal (μ,σ) then $E[e^X]=\exp(\mu+\sigma^2/2)$

GEOMETRIC RANDOM WALK

- $S_{t+j} = S_t e^{\left(\alpha - \frac{1}{2}\sigma^2\right)j + z_j \sigma \sqrt{j}}$ can be used to generate future values for continuous positive values of j
- $E\left[S_{t+j} \middle| S_t\right] = S_t e^{\alpha j}$
- … exactly as we wanted, with constant expected exponential growth of α.
- Therefore we can express the moments of *ln*(S_{t+j}) to be consistent with this outcome:
- *Log-moments of the Geometric Random Walk*
- $E\left[ln\left(S_{t+j}\right) \middle| S_t\right] = ln(S_t) + \left(\alpha - \frac{\sigma^2}{2}\right) j$
- $Var\left[ln\left(S_{t+j}\right) \middle| S_t\right] = \sigma^2 j$
- $Cov\left[ln\left(S_{t+j}\right), ln(S_{t+m}) \middle| S_t\right] = \sigma^2 \, min(j, m)$

Exponentially growing perpetuity?

Used for asset prices



Issues with asset price modeling:

- Zero value unattainable
- Positively skewed future price distributions
- Lognormal future price distributions

CLASS EXERCISE

- Generate a 24-month geometric random walk for Amazon stock prices with these parameters in Excel (α and σ are annual). Show at least six simulations on a graph.
  - $S=1800$, $\alpha=20\%$, $\sigma=25\%$

- Simulate the value of Amazon stock in two years

- Estimate the mean value and verify it is correct

MEAN REVERSION

Mean-reverting perpetuity?

"forgetfulness"

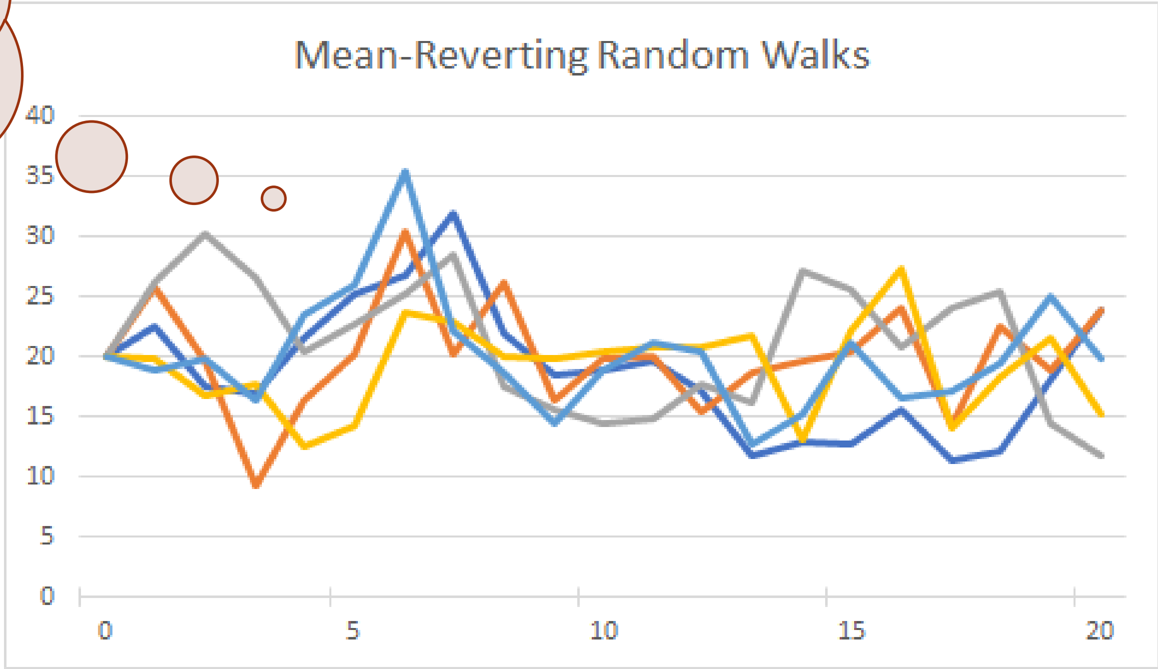
Used for interest rates, commodity prices, some exchange rates

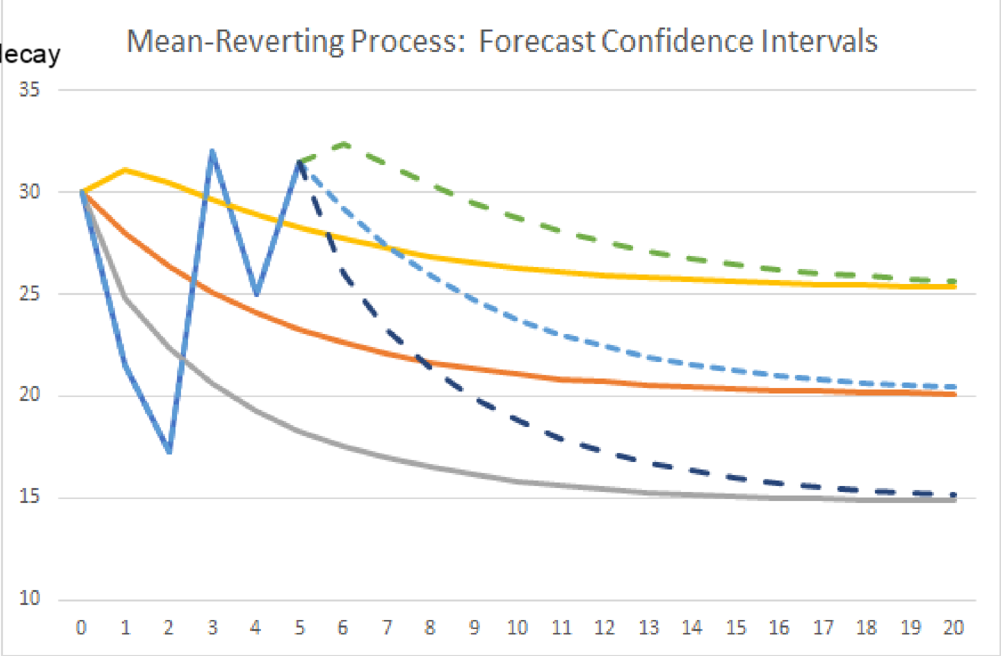
speed of mean reversion

long-run mean

exponential decay

- $\boxed{S_{t+1} = S_t + \kappa(\mu - S_t)} + \sigma z_1$
- $S_{t+j} = \mu + (S_t - \mu)(1-\kappa)^j + \sigma \sum_{i=0}^{j-1}(1-\kappa)^i z_{t+j-i}$
- $E[S_{t+j}|S_t] = \mu + (S_t - \mu)(1-\kappa)^j$
- $Var[S_{t+j}|S_t] = \sigma^2\left(\frac{1-(1-\kappa)^{2j}}{\kappa(2-\kappa)}\right)$
- $Cov[S_{t+j}, S_{t+m}|S_t] = \sigma^2\left(\frac{1-(1-\kappa)^{2j}}{\kappa(2-\kappa)}\right)(1-\kappa)^{m-j}\ for\ m \geq j$
- What happens when κ=0? no longer mean-reverting

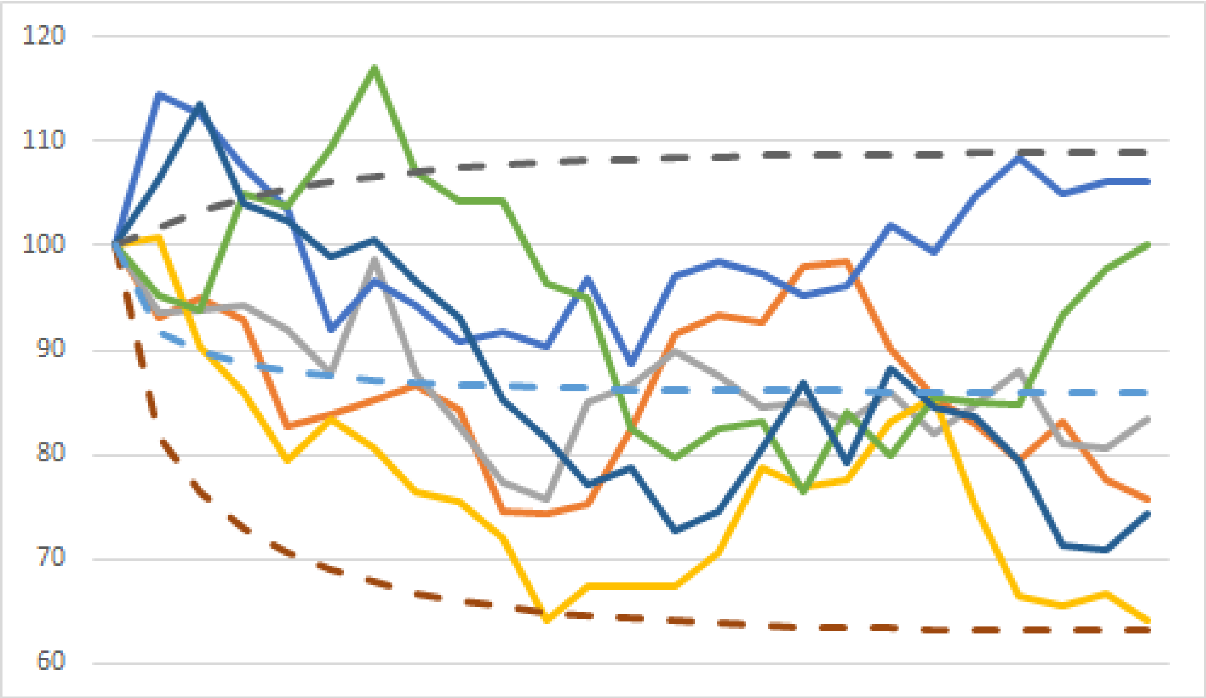




EXERCISE

- Simulate a mean-reverting process for 24 months that has initial value 100, long run mean 80, speed of mean reversion 10% per month, volatility per month of \$5

- Add the two standard deviation confidence intervals



GENERATING SIMULATED PRICES WITHOUT PATHS

Mean-reverting Cholesky decomposition

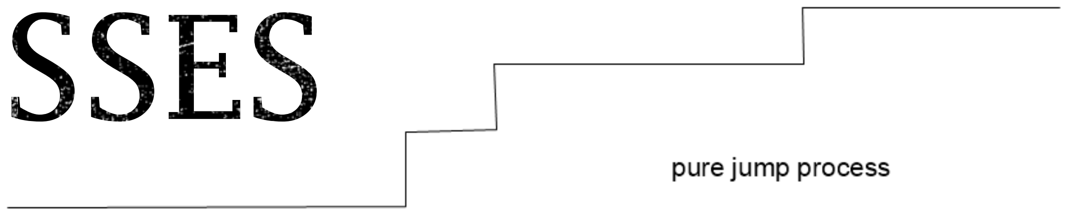
Covariance matrix (σ=1, κ=20%)

1	0.8	0.64	0.512	0.4096
0.8	1.64	1.312	1.0496	0.83968
0.64	1.312	2.0496	1.63968	1.311744
0.512	1.0496	1.63968	2.311744	1.849395
0.4096	0.83968	1.311744	1.849395	2.479516
			sum	30.20826

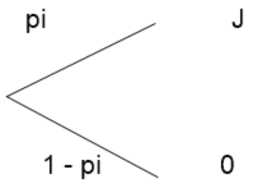
Cholesky decomposition (σ=1, κ=20%)

1	0	0	0	0
0.8	1	0	0	0
0.64	0.8	1	0	0
0.512	0.64	0.8	1	0
0.4096	0.512	0.64	0.8	1

JUMP PROCESSES



- Can be added to any process
- At any time, there is a probability π that an amount J will be added to the process
  - J may be a function of the underlying process
  - J may be a random variable itself
  - The value of π may follow a stochastic process itself



ex. Simulate future power prices --- jump corresponds to loss of power generation facility

- Class exercise
  - Model a mean-reverting process (24 prices) with these parameters and show the graphs
  - Starting value = 100, Long run mean = 80
  - Speed = 20%, Volatility = 5, Jump size = 50, Jump probability = 10%
  - What can you conclude?

PARAMETER EST. GEOMETRIC RANDOM WALK

- Calculate log returns from historical data
- Calculate standard deviation of log returns
- **Adjust the drift by adding half the variance**
- Scale as needed
  - Use the T rule for the mean (this is a mean percentage return)
  - Using the $\sqrt{T}$ rule for the volatility (*volatility* is a percentage s.d.)

12 months

not percentage change but log change

hist price data

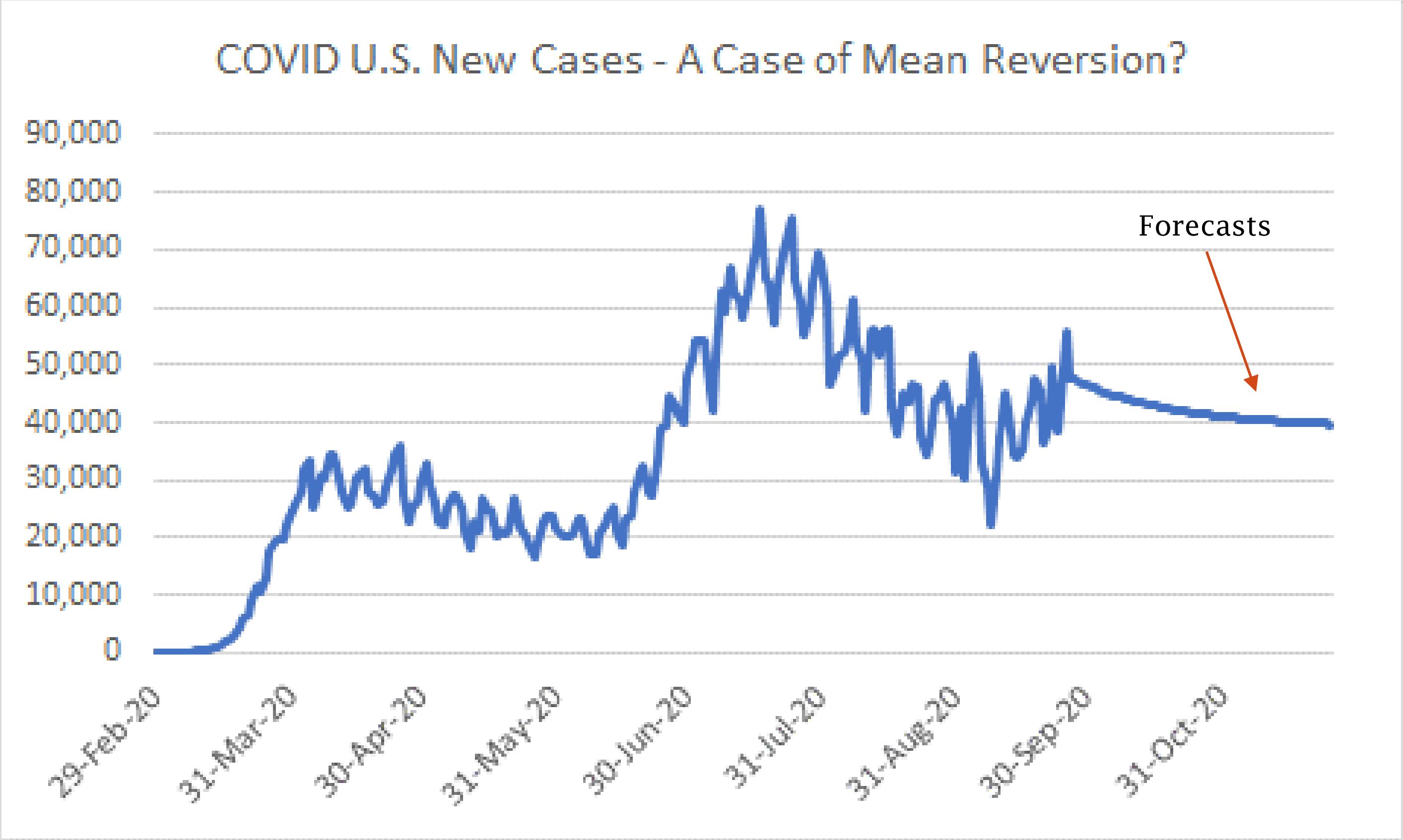
Prices	Log of ratio
20	
22	0.09531018
19	-0.14660347
15	-0.23638878
18	0.18232156
19	0.05406722
21	0.10008346
20	-0.04879016
18	-0.10536052
19	0.05406722
22	0.14660347
21	-0.04652002
23	0.09097178

		annual
Average	0.011647	x 12
Stdev	0.127688	x sqrt(12)

Estimated geometric parameters

Drift	0.237587
Volatility	0.442325

PARAMETER EST. MEAN REVERSION



dependent (Y)

- Basic MR equation
- $\Delta X = \kappa(\mu - X)\Delta t + \sigma\varepsilon$
- $\Delta X = \kappa\mu - \kappa X + \sigma\varepsilon$
- Regression
  - Constant = $\kappa\mu$
  - Slope = $-\kappa$
  - Half-life = $-\ln(2)/\ln(1-\kappa)$
- Run an autoregression
  - Test significance of κ
  - Solve for μ
    - μ = 38,972 new cases/day
    - κ = 4.02% per day
    - t = 2.26

File VL06 – Mean Reversion Test.xslx

Mean reverting if kappa = 0

COMBINATIONS OF CORRELATED PROCESSES

- Class discussion:
- How would you simulate the value of oil and the value of an oil company jointly?
- How would you determine the correlation to use?

Oil - Mean-reversion   oil1 = oil0 + kappa*(mu-oil0) + sigma*z(oil)

Oil stocks - Geometric random walk S1 = S0*exp(alpha - 0.5*sigma^2 + sigma(S)*z(S))

Therefore correlate z(oil) and z(S) in the usual way

What if we want to estimate the correlation to use?

Find between two series:

a. Residuals from the MR regression
b. Convert to logs, subtract average from each outcome

CONTINUOUS ANALOGUES

- Geometric Brownian motion
  - $dS = \alpha S dt + \sigma S dz$
- Arithmetic Brownian motion
  - $dX = \alpha dt + \sigma dz$
- Ornstein-Uhlenbeck mean-reverting process
  - $dX = \kappa(\mu - X)dt + \sigma dz$
- Wiener process
  - dz (dz_t) is a normally distributed infinitesimal, independent for any t pair
  - $E[dz_t] = 0$, $Var[dz_t] = dt$
  - $dW_t = \alpha dt + \sigma dz$; can be correlated in the usual method for normal distributions
  - $df(W_t, t) = \left(\alpha f_W + \frac{1}{2}\sigma^2 f_{WW} + f_t\right) dt + \sigma f_W dz$ *(Itô's lemma)*

OPTIMIZATION WITH STOCHASTIC PROCESSES

Assume

The price of a tree follows an arithmetic random walk with a given drift and standard deviation

The risk-free interest rate is constant

From now until time T in the future, we can “harvest” the tree and collect its value once at any time we choose

Once harvested, the tree is gone

Questions

What does the optimal harvest policy look like?

- We will assume it is when the tree reaches a given optimized constant value in this example

What is the optimal harvesting price?

What is the value assuming optimal harvesting?

HOW TO OPTIMIZE?

- Optimization and simulation do not fit well together
- Best to use fixed random numbers (or fixed seeds) to value each asset under different random scenarios
- Recommended steps:
  1. Decide the form of the optimization rule
  2. Establish "cases" which represent sets of fixed random numbers. Number the cases.
  3. Value the asset assuming arbitrary optimization parameters for a single case.
  4. Revalue the asset for a range of discrete parameter values for every case.
  5. Choose the parameter value that produces the highest valuation.

TREE ILLUSTRATION

Optimization of a simulation

Cutting down the ABM "tree", finite case

Parameters		*Monthly*
Starting height	100	
Annual drift	6	0.50
Annual std dev	3	0.87
Risk-free interest rate	5.00%	0.42%
Maximum time	5	60

Fixed random number case #	7	
Arbitrary cutting height	110	*Single case*
- type "1" for random	0	
PV for one case (not optimized)	102.09	

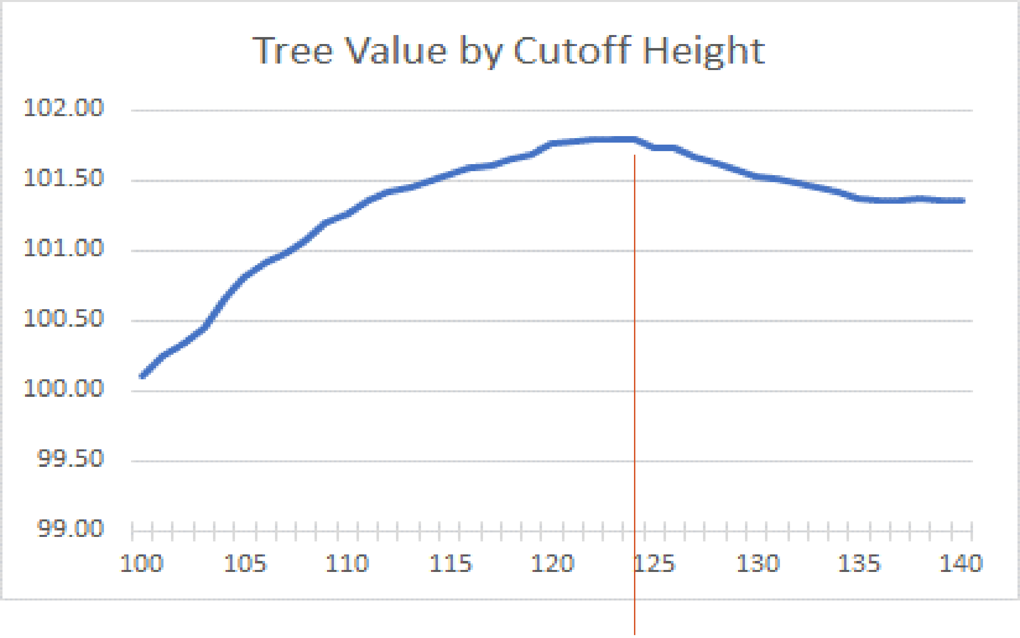
Reference file VL06: Cutting down an ABM tree.xlsx

Analysis for one case

Month	z	Height	CF	DF
1	0.16	100.64	0	0.9959
2	-0.02	101.12	0.00	0.9917
3	-0.84	100.89	0.00	0.9876
4	-0.84	100.66	0.00	0.9835
5	0.54	101.62	0.00	0.9794
6	1.19	103.16	0.00	0.9754
7	-0.42	103.29	0.00	0.9713
8	-0.70	103.18	0.00	0.9673
9	-0.83	102.96	0.00	0.9633
10	-0.98	102.62	0.00	0.9593
11	-1.14	102.13	0.00	0.9553
12	0.84	103.36	0.00	0.9513
13	1.75	105.38	0.00	0.9474
14	-0.06	105.83	0.00	0.9434
15	0.24	106.53	0.00	0.9395
16	-0.04	107.00	0.00	0.9356
17	0.43	107.87	0.00	0.9318
18	0.93	109.18	0.00	0.9279
19	-0.51	109.24	0.00	0.9240
20	1.39	110.95	110.95	0.9202
21	-0.20	111.27	0.00	0.9164
22	0.20	111.95	0.00	0.9126

GENERATING CASES FOR EACH CUTTING POLICY

	MaxSelect	0	0	0	0	0	0	0	0	0
	Valuation Statistics									
	Average	100.11	100.25	100.34	100.45	100.65	100.82	100.91	100.98	101.07
	MaxErr	0.07	0.09	0.12	0.14	0.15	0.16	0.18	0.19	0.20
	Valuation by case and cutting height									
	Value		1		*Cutting height policy*					
	102.09	100	101	102	103	104	105	106	107	108
	1	99.65	100.65	98.34	99.18	99.43	92.71	93.16	93.79	93.79
	2	99.81	100.40	101.27	101.38	101.36	102.62	102.62	99.76	100.06
All	3	100.29	100.29	101.36	101.36	101.43	98.48	99.36	99.36	100.10
Cases	4	99.38	99.57	99.57	100.75	100.75	100.13	100.11	101.00	99.61
	5	100.51	101.05	102.35	102.35	101.28	101.28	99.62	99.87	101.93
1	6	99.23	99.02	99.75	100.08	99.51	100.91	100.91	101.59	99.81
to	7	100.22	100.28	100.61	100.61	99.83	99.83	100.09	100.51	101.31
1000	8	99.79	100.08	101.41	101.41	100.93	102.25	102.25	101.39	101.64
	9	99.67	100.13	100.25	101.01	100.05	96.37	97.14	98.10	98.11
	10	99.23	101.28	101.28	101.03	97.47	97.56	97.67	98.35	98.35



Optimized values (within column range)

Cutting height	124
Valuation	101.79
Value times interest rate	6.2

REVIEW QUESTIONS

- (Answer these questions as if you were in an interview. Explain it to someone who doesn't know.)

1. What is a stochastic process?

2. What's the difference between stochastic processes and time series?

3. What are stochastic processes used for in finance?

4. Why do interest rates have different stochastic models than asset pries?

5. Why do commodity prices follow mean reversion?

REVIEW QUESTIONS B: ANSWER NORMALLY

6. Which types of stochastic processes are we focused on in this presentation?

7. Mathematically, how would you simulate an arithmetic random walk?

8. What is the mean and variance of S_{t+j} in an arithmetic random walk if S_t is known?

9. What is the covariance between S_{t+j} and S_{t+k} if S_t is known (for the ARW)?

REVIEW QUESTIONS C

10. What is the Cholesky decomposition of the covariance matrix in #9?

11. Generate a 24-month arithmetic random walk for *the natural log of* Amazon stock prices with these parameters in Excel (α and σ are annual). Show at least six simulations on a graph.

    S=1800, α=20%, σ=25%

12. Show the mean and confidence intervals

REVIEW QUESTIONS D

13. To generate a lognormal random walk with drift α and volatility σ, what log process do we use to generate the stochastic process for log prices in the future?

14. What formula generates future values of a geometric random walk?

REVIEW QUESTIONS E

15. Simulate a mean-reverting process for 24 months that has initial value 100, long run mean 80, speed of mean reversion 10% per month, volatility per month of $5

16. Add the two standard deviation confidence intervals

17. Add a jump

REVIEW QUESTIONS F

18. Given a set of daily stock prices, how would you estimate α and σ for a geometric random walk?

19. How would you convert the α and σ for monthly and annual simulations?

20. How do you test for statistically significant mean reversion?

REVIEW QUESTIONS G

21. How would you simulate the value of oil and the value of an oil company share jointly?

22. How would you estimate the parameters of the last question from historical prices?

- Note to prof: Optimization example

FORMULAS IN THIS PRESENTATION A

1. $Var(\mathrm{E}[y]) = s_e^2 X_h'(X'X)^{-1}X_h$
2. $S_{t+1} = S_t + \mu + \sigma z_1$
3. $S_{t+j} = S_t + j\mu + \sigma \sum_{i=1}^{j} z_{t+i}$
4. $E[S_{t+j}|S_t] = S_t + j\mu$
5. $Var[S_{t+j}|S_t] = j\sigma^2$
6. $Cov[S_{t+j}, S_{t+m}|S_t] = \sigma^2 \min(j, m)$
7. $lnS_{t+1} = lnS_t + \left(\alpha - \frac{1}{2}\sigma^2\right) + \sigma z_{t+1}$
8. $S_{t+j} = S_t e^{\left(\alpha - \frac{1}{2}\sigma^2\right)j + z_{t+1}\sigma\sqrt{j}}$

FORMULAS IN THIS PRESENTATION B

9. $S_{t+1} = S_t + \kappa(\mu - S_t) + \sigma z_1$
10. $S_{t+j} = \mu + (S_t - \mu)(1-\kappa)^j + \sigma \sum_{i=0}^{j-1} (1-\kappa)^i \, z_{t+j-i}$
11. $E[S_{t+j}|S_t] = \mu + (S_t - \mu)(1-\kappa)^j$
12. $Var[S_{t+j}|S_t] = \sigma^2 \left(\frac{1-(1-\kappa)^{2j}}{\kappa(2-\kappa)}\right)$
13. $Cov[S_{t+j}, S_{t+m}|S_t] = \sigma^2 \left(\frac{1-(1-\kappa)^{2j}}{\kappa(2-\kappa)}\right)(1-\kappa)^{m-j} \; for \; m \geq j$
14. $-\ln(2)/(1-\kappa)$
15. $\Delta X = \kappa\mu - \kappa X + \sigma\varepsilon$
16. $dS = \alpha S dt + \sigma S dz$
17. $dX = \alpha dt + \sigma dz$
18. $dX = \kappa(\mu - X)dt + \sigma dz$