

AAEA Pre-Conference Teaching Workshop

Simulation for Risk Analysis

James W. Richardson, Keith D. Schumann, and Paul A. Feldman
Texas A&M University

Allan W. Gray
Purdue University

Subject

The demand for agribusiness graduates trained in risk analysis and risk management is increasing faster than Agricultural Economics Departments can produce qualified students. A few departments have courses that touch on simulation and even fewer offer complete courses on simulation for risk analysis and management. Agricultural Economics Departments need to improve student training in the area of simulation, risk analysis, and risk management or agribusiness may turn to MBA schools to fill their needs in this area.

Objective

The objective of this pre-conference is to help participants gain an understanding and proficiency with stochastic simulation for risk analysis. A primary goal is to have participants become comfortable with teaching the concepts of stochastic simulation at the advanced undergraduate or graduate level. This subject material will be of interest to agricultural economists who want to teach simulation. The material will also be of interest to researchers and extension specialists, USDA researchers and other agricultural professionals who want to use simulation in their work.

Motivation

Over the past several years there has been an increased interest in risk analysis by businesses due to a perception of increased market risks. Increased price volatility and globalization of markets have contributed to the risk faced by business. The “freedom to farm” features of the farm program have also contributed to a perception of increased risk for agribusiness. One result of agribusinesses’ perception of increased risk has been an increased demand for students trained in risk analysis and management.

Courses that teach risk analysis and risk management are not widely available in Departments of Agricultural Economics. Many students graduate without an understanding of how to analyze risk, evaluate alternative risk management strategies, or how to choose the “best” risky strategy. Simulation remains the preferred methodology for analyzing risk and evaluating alternative risk management strategies. In the past, large computers and expert programming ability were required to develop risk analysis simulation models. However, the wide spread availability of micro computers and our graduate’s familiarity with Excel makes it possible for businesses of all sizes to benefit from advances in the field of simulation for risk analysis.

Dr. James Richardson at Texas A&M University presently teaches two graduate courses in simulation for risk analysis, one at the Masters level and another at the Ph.D. level. Dr. Allan Gray at Purdue University also teaches two graduate course on the subject, one in the Master's of Science program and one in the Executive MBA in Food and Agribusiness Program. At both universities, the courses have been extremely well received by the students, and in many cases students have received jobs based specifically on their training in these courses.

The courses use Excel as the vehicle for teaching students how to program complex risk analysis simulation models. An Excel Add-In for simulation, Simetar©: Simulation for Excel to Analyze Risk, has been developed at A&M to facilitate development and use of simulation models in Excel. The simulation Add-In is easier to use than @Risk and offers a range of functions for parameter estimation and data analysis, such as: multiple regression, hypothesis testing, time series analysis, stochastic dominance, certainty equivalents, and numerous graphics capabilities for data presentation. Simetar© can be installed on micro computers running Excel 97 or 2000, regardless of the computers's speed or memory.

Intended Audience

The workshop will focus on helping graduate and undergraduate teachers gain confidence in teaching simulation. Researchers and extension specialists, USDA researchers and other agriculture professionals involved in risk analysis would benefit from the material, as well. Participants will be encouraged to bring their computers, load the demonstration simulation programs, and follow along with the presentations on: data analysis, multiple regression estimation, time series analysis, forecasting, distribution parameter estimation, stochastic simulation, scenario and sensitivity analysis, hypothesis testing for validation, and choosing the "best" strategy using stochastic dominance, dynamic certainty equivalents and probabilistic target graphs. Experience with Microsoft Excel will be necessary for the participants to benefit fully from the workshop.

A book on simulation for risk analysis in Ag Economics will be provided to each participant. A copy of Simetar© and more than 75 demonstration simulation programs for teaching and research purposes also will be provided. Enrollment should be limited to 35 participants.

Planned Follow-up

The presenters will provide help to teachers interested in starting a simulation course or expanding on the topic in their existing courses. Participants interested in using simulation in their work will be supported as well via e-mail. Course syllabi, homeworks, example programs, and overheads will be shared with prospective instructors. As additional follow up, an Organized Symposium on the subject of "simulation in our graduate and undergraduate programs" will be proposed by the organizers of this pre-conference for 2003 AAEA Annual Meetings.

Presentations and Format

The workshop will run from 9 am to 5 pm with registration starting at 8:30 am. The presentation format will involve four 1 ½ hour blocks with presentation and lab time being split about 50/50. The four presenters will help participants during the lab times to gain proficiency in developing and applying simulation models for risk analysis in a hands on setting using the participants computers.

The four 1 ½ hour blocks will be organized as follows:

- Introduction to simulation modeling in Excel and distribution parameter estimation, multiple regression, and forecasting.
- Simulation of multivariate distributions, sensitivity analysis, and hypothesis testing for validation.
- Scenario analysis with simulation and selection of the “best” risky strategy for risk management.
- Use of simulation for project management and analysis under risk.

Speakers and Organizers

Dr. James W. Richardson – Professor and TAES Faculty Fellow, Department of Agricultural Economics, Texas A&M University, College Station, TX 77843, jwrichardson@tamu.edu, 979-845-5913, fax 979-845-3140.

Mr. Keith D. Schumann – Extension Risk Management Specialist, Department of Agricultural Economics, Texas A&M University, College Station, TX 77843, k-schumann@tamu.edu, 979-845-5913, fax 979-845-3140.

Mr. Paul A. Feldman – Extension Associate, Department of Agricultural Economics, Texas A&M University, College Station, TX 77843, paulf@tamu.edu, 979-845-5913, fax 979-845-3140.

Dr. Allan W. Gray – Assistant Professor, Department of Agricultural Economics, Purdue University, West Lafayette, IN 47907, gray@agecon.purdue.edu, 765-494-4323, fax 765-496-1224.

Time Required

- 8:30 am to 5 pm

Preliminary Budget

- Speaker fees will be covered by their institutions. Speakers will provide their own audio-visual projectors, etc.
- Course materials @ \$35 per person.
- Morning and afternoon refreshments and a sandwich bar for the noon meal @ \$20 per person.
- Additional electrical outlets will be required so each person has an outlet guessed at \$500 for the room.
- Estimated total cost for 35 participants is:
 $\$2,425 = (\$55 / \text{person} * 35) + \$500/\text{meeting room expense}$

**Agenda for
AAEA Post-Conference Workshop on
Simulation for Risk Analysis**

- 8:00-9:30 Computer setup and software installation testing
- 9:30-10:00 Break and refreshments
- 10:00-11:00 Parameter estimation in Excel using Simetar. Lab includes:
- Data analysis for statistics, trend, hypothesis testing
 - Multiple regression and parameter estimation
- 11:00-12:00 Simulation modeling in Excel using Simetar. Lab includes:
- Simulation of random values for various univariate distributions
- 12:00-12:45 Lunch Break
- 12:45-2:00 Simulation of multivariate probability distributions, sensitivity analysis, and hypothesis testing for validation of simulation models. Lab includes:
- Multivariate Normal and Non-Normal Distributions
 - Testing correlation of simulated variables
- 2:00-3:00 Scenario analysis and selecting the “best” risky strategy. Lab includes:
- Scenario analysis
 - Stochastic dominance with respect to a function,
 - Dynamic Certainty Equivalents,
 - Stoplight charts, fan graphs, CDF graphs, and probability charts
- 3:00-3:15 Break and refreshments
- 3:15-4:30 Business analysis in a risk setting using simulation. Lab will give you a chance to experiment on one or two of the following topics
- Project management
 - Portfolio analysis
 - Marketing options analysis
 - Business plan analysis
 - Crop insurance
 - Farm simulation
 - Bootstrap simulation of econometric models
 - Deterministic and probabilistic forecasting
 - Inventory management
- 4:30-5:00 Experiences teaching simulation in Ag Economics Departments

AAEA Post Conference Workshop on Simulation for Risk Analysis

James Richardson

Keith Schumann

Paul Feldman

Allan Gray

Andrew Falwell

July 31, 2002

What is Simetar?

- An Excel Add-In designed to address most all statistical and econometric functions required to develop and apply Monte Carlo simulation models
- Developed at A&M to make my teaching and research jobs easier
 - Under development for two years plus
 - Replaces @Risk for simulation
 - Better than Excel for regression
 - Provides dynamic tools for analyzing simulation results

Do the Housekeeping Now

- What is it with the registration fees?
 - \$25 for lunch and refreshments – so enjoy
 - \$25 for meeting room
 - \$10 for electricity
 - \$15 for AAEA
 - Nothing for postage or the presenters
- Now lets get to work

Part 1: Parameter Estimation

- Data Analysis to obtain estimates of the deterministic and stochastic components of random variables
- Using Multiple Regression to estimate parameters for stochastic variables
- We will use the Excel workbook named:
Data Analysis and Regression.XLS

What are the Parameters?

- A random variable can be fully defined by two parameters
 - Deterministic component or parameter
 - Mean or forecast of the mean
 - Stochastic component or parameter
 - Standard deviation of actual residuals about the forecast
- Normal (mean, standard deviation)
- Empirical (mean, sorted deviates, probability of deviates)

Parameter Estimation for Univariate Distributions

- A random variable X has two components:
 - Deterministic component
 - Stochastic component
- The goal of econometric modeling and forecasting is to minimize the unexplained portion or the stochastic component
- Econometric models and forecasting models estimate parameters for the Deterministic Component
- Simulation Modeling estimates the parameters for the stochastic component

Let the random variable Y be a function of other variables

$$Y = a + b_1X_1 + b_2X_2 + e$$

- Use econometric methods to estimate parameters for deterministic component

$$\hat{Y} = \hat{a} + \hat{b}_1X_1 + \hat{b}_2X_2$$

- Stochastic component is estimated as

$$\hat{e}_i = Y_i - \hat{Y}_i$$

Parameters to define and simulate the stochastic component depend on the Dist. for \hat{e}

- If \hat{e} distributed Normal

We calculate mean of \hat{e} and std. dev. of \hat{e}
or $\tilde{e} \sim N(\bar{e}, \sigma_e)$

- If \hat{e} distribution Uniform

The parameters are min \hat{e} and max \hat{e}
or $\tilde{e} \sim U(\text{Min } \hat{e}, \text{Max } \hat{e})$

- If \hat{e} is distributed Empirical

the parameters are $S_i, P(S_i)$

recall S_i is the sorted \hat{e}_i values

$P(S_i)$ is the probabilities of the \hat{e} 's from 0 to 1

Data Analysis and Parameter Estimation Using Simetar

- Simple Statistics for a data series
- Trend Regression for one or more variables
- Estimate a Demand model using the data
- Refit the model without a lagged demand
- Correct errors in the data
 - Last 4 values for income should be:
24.0 24.1 24.3 24.5
- Examine the parameters for simulation
 - Forecast mean
 - Standard Deviation

Parameter Estimation

- Empirical Distribution
 - Parameters are sorted data from trend or mean and probabilities of observing the values
- Discrete Uniform Distribution
 - Parameters are sorted values
- Bernoulli Distribution
 - Probability of an event occurring, as rain or mechanical failure

Part 2: Simulating Univariate Probability Distributions

- Simulate random values from different types of distributions
 - Uniform, Normal, Empirical, Discrete
Uniform, Bernoulli
- Simulate a regression equation to generate a probabilistic forecast

Uniform Distribution

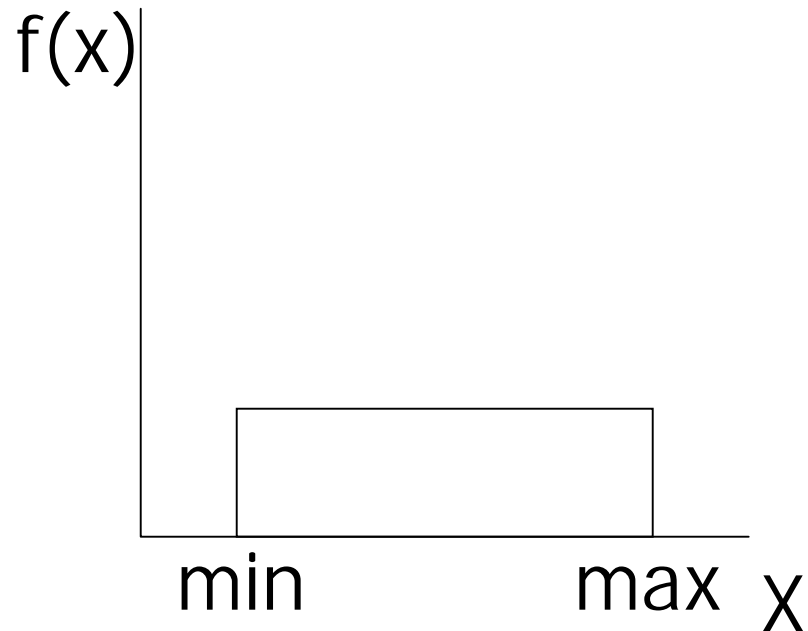
- =UNIFORM(min,max)
- The uniform distribution is a continuous distribution where each value of X has an equal probability of being observed.
- The parameters for the uniform are the minimum and maximum values and the domain includes all real #'s.
- The mean and variance of this distribution are:

$$\bar{X} = \frac{\text{min} + \text{max}}{2}$$

$$\sigma^2 = \frac{(\text{max} - \text{min})^2}{12}$$

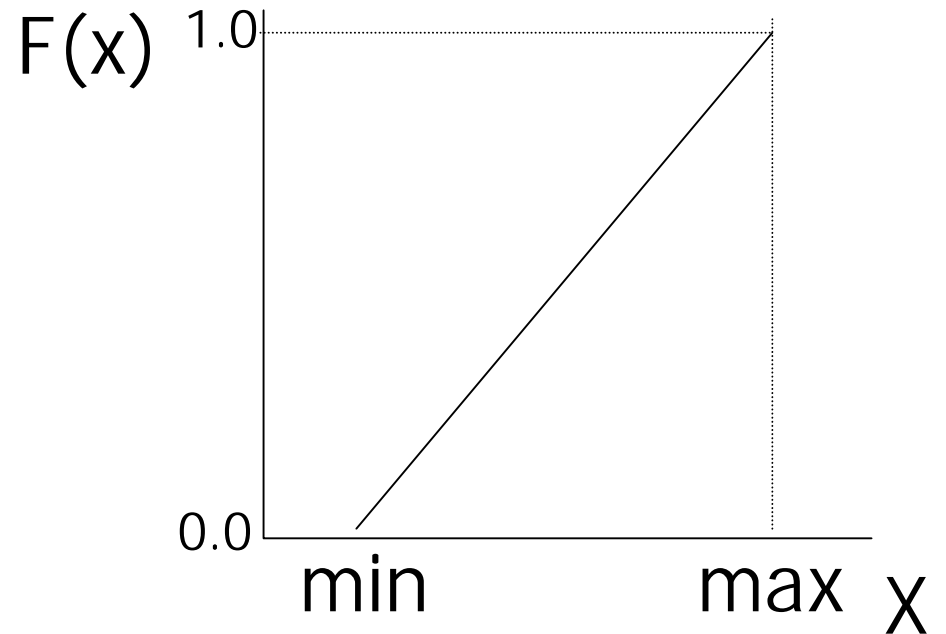
PDF and CDF for a Uniform Dist.

Probability Density Function



$$f(X) = \frac{1}{\max - \min}$$

Cumulative Density Function



$$F(X) = \frac{x - \min}{\max - \min}$$

Using the Uniform Distribution

- **Use the uniform distribution when each observation between the minimum and maximum values has an equal chance of occurring.**
- **You may use this distribution when you have no idea what type of distribution to use.**

Uniform Standard Deviate

- **=UNIFORM(0,1)**
- **Special case of the uniform distribution**
- **Building block for all random number generation using the Inverse Transformation method for simulation**
- **Can use to simulate any uniform distribution as follows:**

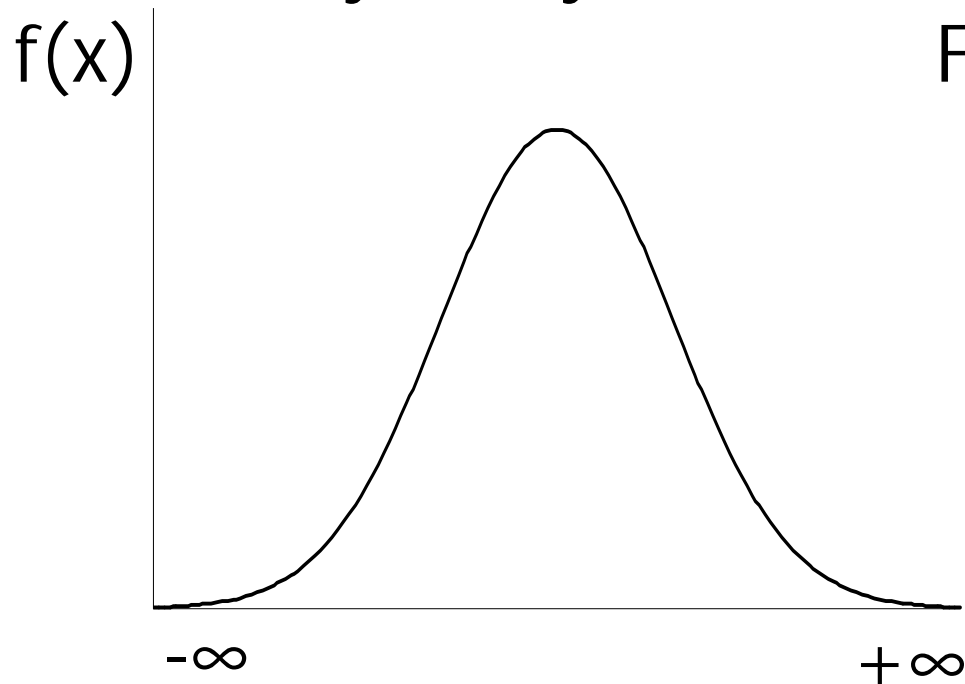
$$\text{Min} + (\text{Max}-\text{Min}) * \text{Uniform}(0,1)$$

Normal Distribution

- **=NORM(mean, standard deviation)**
- **The normal distribution is a continuous distribution that produces a bell shaped distribution with set probabilities.**
- **The parameters for the normal are the mean and standard deviation and the STD must be > 0 .**
- **The function reaches to +/- infinity.**
- **Most of us have memorized several probabilities of the normal distribution:**
 - **66% of observation within +/- 1σ of the mean**
 - **95% of observation within +/- 2σ of the mean**
 - **50% of observations lie above and below the mean.**

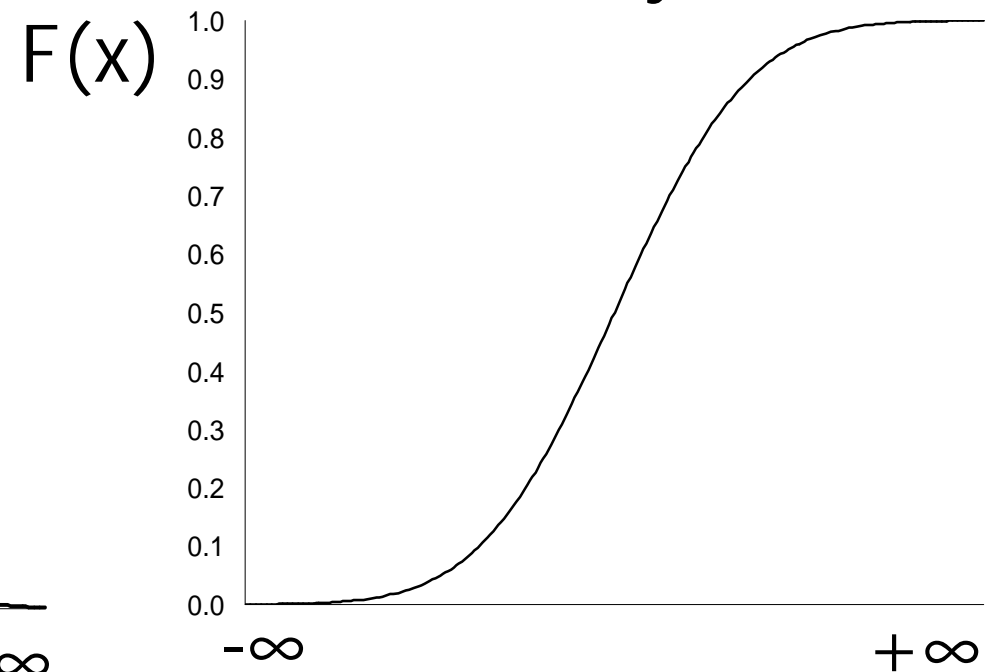
PDF and CDF for a Normal Dist.

Probability Density Function



$$f(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}$$

Cumulative Density Function



$$F(X) = \text{No closed form}$$

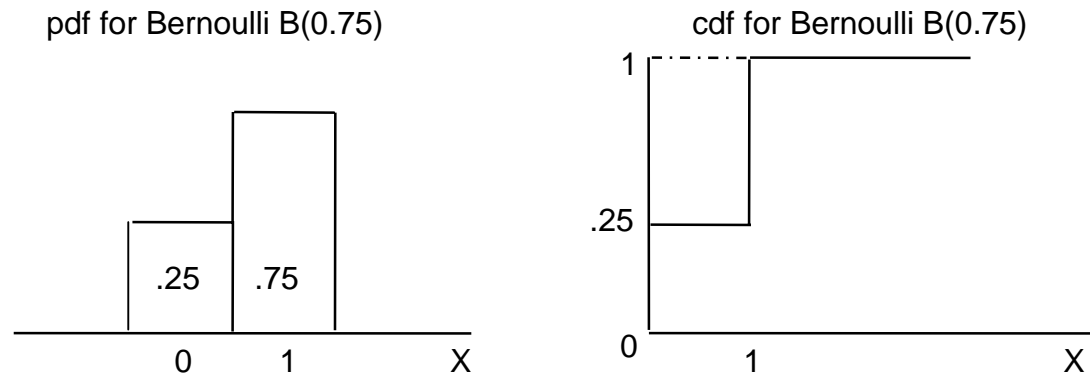
Using the Normal Distribution

- **Might use the normal distribution if the random variable is the error term on a regression equation or the data have been tested and are indeed normally distributed.**
- **Use the normal distribution if you have lots of observations and have tested for normality.**
- **Watch for irrelevant values from a normal distribution (negative yields, prices, etc.)**

Standard Normal Deviate

- **=Norm(0,1)**
- **Special case of the normal distribution**
- **Returns the z-score for a standard normal distribution allowing you to simulate any normal distribution**
- **The SND can be used as follows:**
- **Mean + Standard deviation*NORM(0,1)**

Bernoulli Distribution Parameter is probability of true



PDF and CDF for a Bernoulli Distribution.

- Parameter is $B(P)$
P is the probability that X is 1 or true (or occurs), so by definition $1-P$ is the probability that X is false (or does not occur).
- Simulated as:
 - = BERNOLLI (P)
 - = BERNOLLI (A1)

Bernoulli Distribution (continued)

- Use Bernoulli in a Conditional Distribution as these two examples demonstrate:
 - It rains 20% of the time during the season being analyzed and if it rains the amount of rain is distributed $U(0.10, 0.90)$
A2 cell = Bernoulli (0.20)
A3 cell = UNIFORM (0.1, 0.9)
A4 cell = IF (A2. eq. 1, A3, 0)
 - Probability of mechanical failure is 5%, cost of repair is 10,000, 20,000 or 30,000
A4 cell = BERNOLLI (0.05)
A5 cell = RANDSORT (10,000, 20,000, 30,000)
A6 cell = IF (A4. eq. 1, A5, 0)

Discrete Uniform

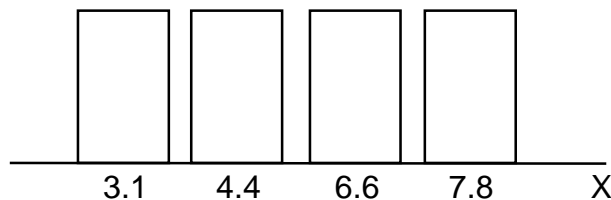
- **The discrete uniform distribution is used for situations where only a few possible values can occur. Each with an equal probability.**
- **This function might be used for things such as, discrete labor hiring, number of head of cattle on a truck, etc.**

Discrete Uniform Distribution

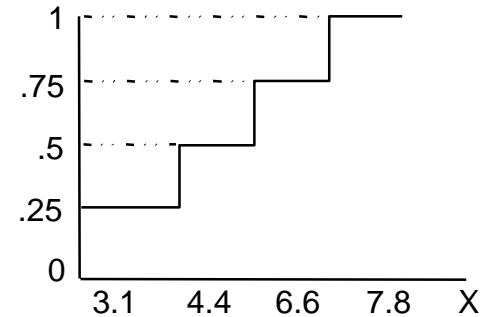
-Parameters are DU ($x_1, x_2, x_3,$
values

... x_N) based on historical

pdf for DU(3.1, 4.4, 6.6, 7.8)



cdf for DU(3.1, 4.4, 6.6, 7.8)



PDF and CDF for a Discrete Uniform Distribution.

- Discrete Uniform means that each observed value of X_i , has an equal probability of being observed

Row	A	B	C
1	10	= RANDSORT (A1:A5)	
2	12		
3	20	=DEMPIRICAL (A1:A5)	
4	15		
5	13		

Discrete Uniform Distribution (continued)

- Simulate this two ways as:
 - Discrete empirical with four observed values

Sorted Values S_i	Implied Probabilities of Values or $P(S_i)$
3.1	.25
4.4	.25
6.6	.25
7.8	.25

= DEMPIRICAL (S_i)

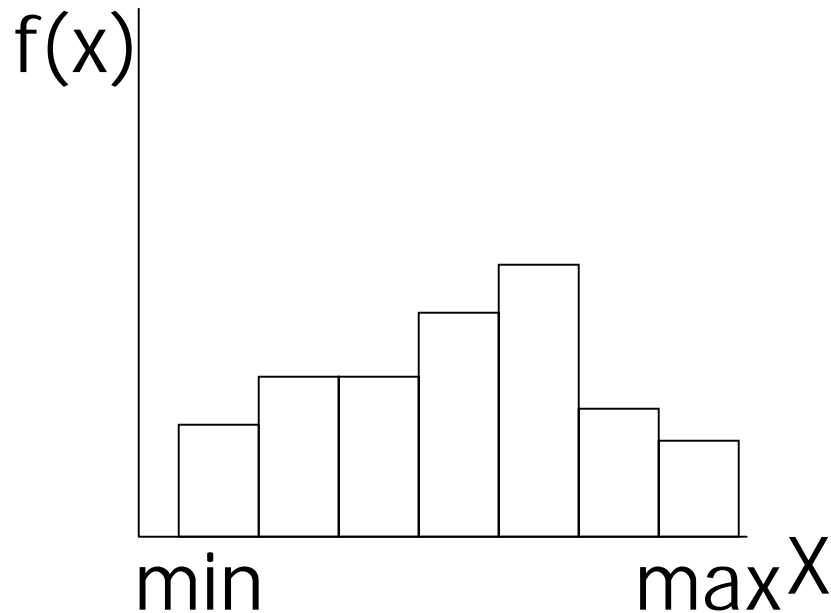
- DU can also be simulated as a random draw without replacement

= RANDSORT (S_i)

This function returns one of the S_i values each iteration. Each S_i has an equal chance of being selected.

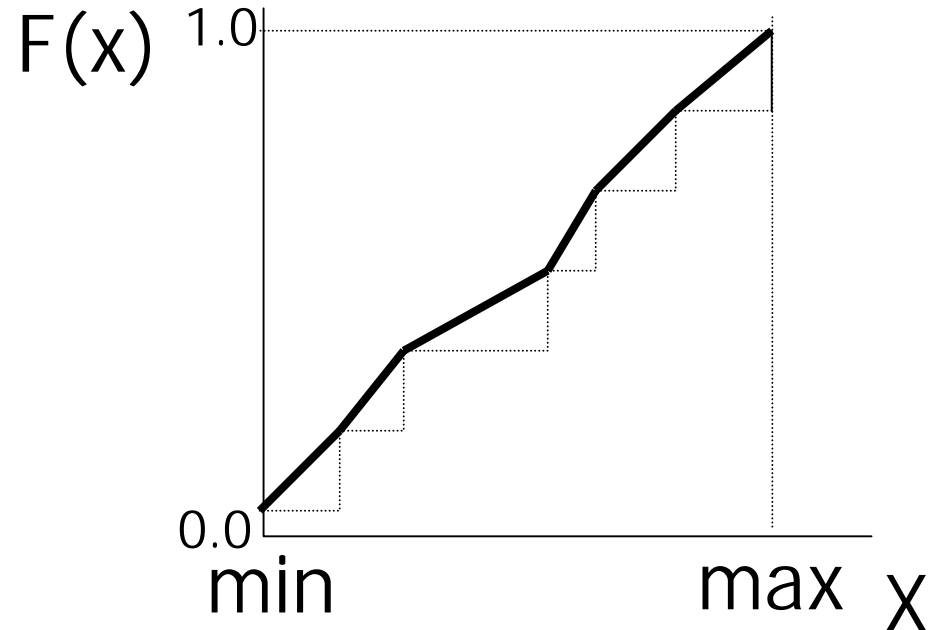
PDF and CDF for an Empirical Dist.

Probability Density Function



$$f(X) = \frac{p_{i+1} - p_i}{x_{i+1} - x_i}$$

Cumulative Density Function



$$F(X) = \frac{x - x_i}{x_{i+1} - x_i} (p_{i+1} - p_i) + p_i$$

$$x_i \leq x \leq x_{i+n} \quad i \in \{0, \dots, n+1\}$$

$$x_0 = \min, x_{n+1} = \max, p_0 = 0, p_{n+1} = 1$$

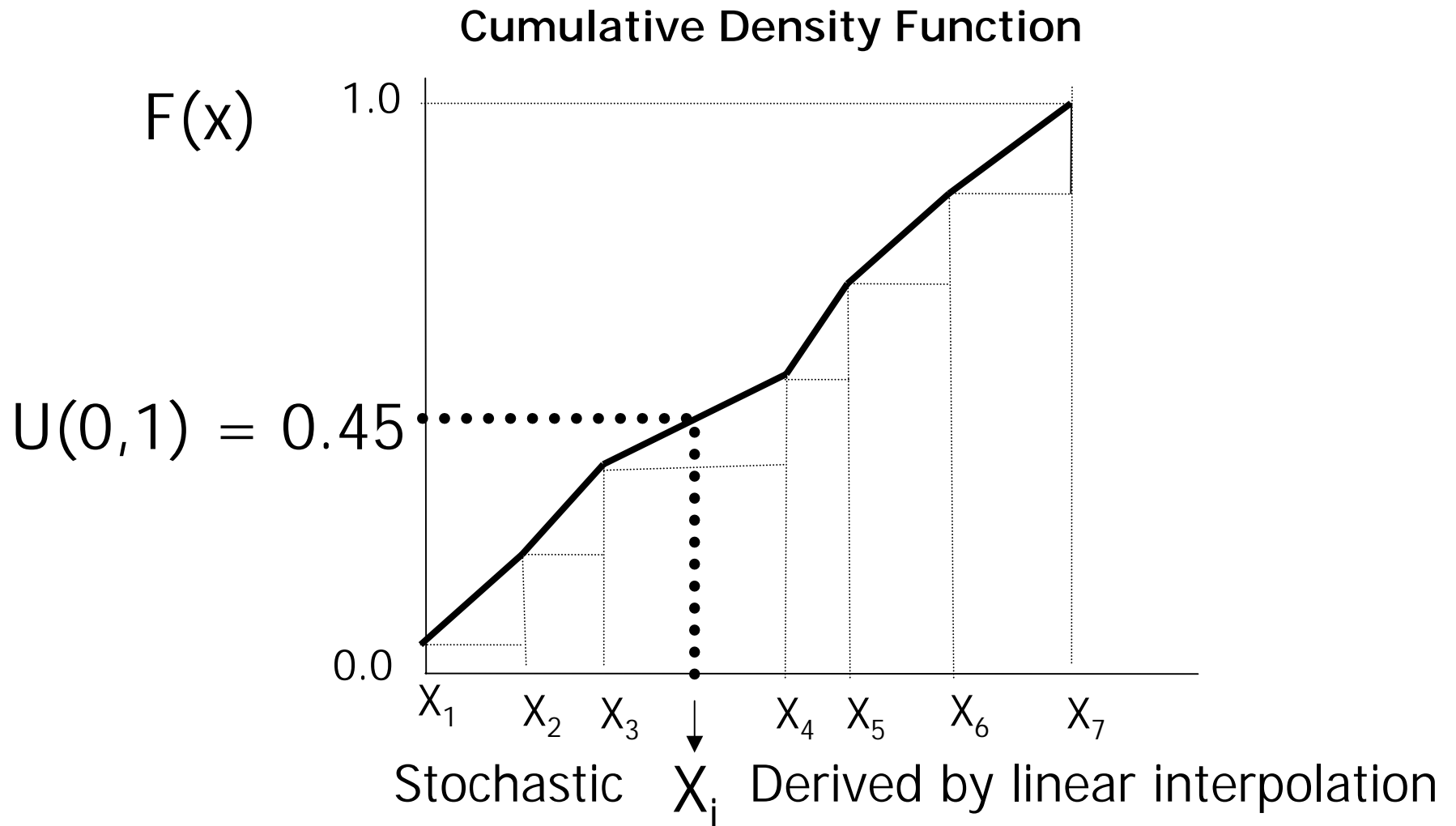
Using the Empirical Distribution

- **The empirical distribution should be used if the random value can take on a continuous value, you have limited observations for the variable, and/or you cannot estimate parameters for the true PDF.**
- **Simulating crop yields as an empirical distribution when you have only 10 historical values is a good example.**
- **In this situation we know:**
 - **Yield can be any positive value**
 - **We don't have enough observations to test for normality.**
 - **We know the 10 yields were observed with a probability of 1/10, or one each year.**

More on the Empirical Distribution

- **An empirical distribution is defined totally by the observations for the data.**
- **The basic steps for simulating an empirical distribution are:**
 - **Sort the individual data points from low to high**
 - **Assign a cumulative probability to each data point (usually equal probability)**
 - **Assume the distribution is continuous so you can interpolate between the points.**
 - **Use the inverse transform to simulate the distribution.**

More on the Empirical Distribution



Empirical Distribution for a Variable X with 7 Observations

<u>Obs</u>	<u>Sorted X_i or S_i</u>	<u>$P(S_i)$</u>	
1	2.09	0	
2	2.10	.1	
3	4.20	.3	The $P(S_i)$ values are uniformly distributed from 0 to 1.
4	7.30	.5	
5	11.2	.7	
6	14.1	.9	
7	14.11	1.00	
Mean =	7.78		

- The S_i values can be actual numbers or deviates from mean or trend.
- The parameters to simulate the X variable above, expressed as deviations from the mean are:

<u>Obs</u>	<u>S_i as fract \bar{X}</u>	<u>$P(S_i)$</u>	for mean = 7.78
1	-0.7313	0	
2	-0.7300	.1	
3	-0.4601	.3	
4	-0.0617	.5	
5	0.4396	.7	
6	1.8123	.9	
7	1.8136	1.0	

- The parameters are estimated directly by Simetar©.

Three ways to simulate the Empirical distribution

- **If S_i expressed as actual values**

$$= \text{EMPIRICAL}(S_i, P(S_i))$$

$$= \text{EMPIRICAL}(A1:A7, B1:B7, \text{USD})$$

- **If S_i is expressed as deviates from mean**

$$= \text{Mean} * [1 + \text{EMPIRICAL}(S_i, P(S_i))]$$

$$= A8 * (1 + \text{EMPIRICAL}(A1:A7, B1:B7, \text{USD}))$$

- **If S_i is expressed as deviations from trend or from a structural model, use the residuals as fractions or:**

$$S_i = \hat{e}_i / \hat{Y}_i$$

$$\text{where } \hat{Y}_i = \hat{a} + \hat{b} T_i$$

$$\text{and } \hat{e}_i = Y_i - \hat{Y}_i$$

In Excel it is simulated as:

$$= (\hat{a} + \hat{b}T) * (1 + \text{EMPIRICAL}(S_i, P(S_i), \text{USD}))$$

Part 3: Simulating Multivariate Probability Distributions

- Multivariate Normal or MVN
- Multivariate Empirical or MVE
- Multivariate Mixed or MVM

Parameters for a MV Distribution

- Deterministic Component

\hat{X}_{ij} which is the mean or the predicted value for period i to simulate all j random variables

- Stochastic Component

\hat{e}_j which are the residuals from the predicted values for each of j random variables

- Multivariate Component

$$\rho = \begin{vmatrix} 1.0 & \rho_{12} & \rho_{13} & \rho_{14} \\ & 1.0 & \rho_{23} & \rho_{24} \\ & & 1.0 & \rho_{34} \\ & & & 1.0 \end{vmatrix}$$

The correlation matrix here is for 4 variables, but in the general case it is a $j \times j$ matrix

3 Variable Case for a Multivariate Normal (MVN)

- The 3 variables are yields for cotton, wheat and sorghum
- Assume the best estimates of their predicted or Deterministic Components (parameters) are found in Step I

$$\hat{C}_i = \hat{a} + \hat{b}_1 T_i + \hat{b}_2 C_{i-1}$$

$$\hat{W}_i = \hat{a} + \hat{b}_1 W_{i-1}$$

$$\hat{S}_i = \hat{a} + \hat{b}_1 T_i$$

- The Stochastic Components in Step II for a MVN distribution are:

$$\hat{e}_{ci} = C_i - \hat{C}_i \quad \text{MVN parameter is } \hat{\sigma}_c$$

$$\hat{e}_{wi} = W_i - \hat{W}_i \quad \text{MVN parameter is } \hat{\sigma}_w$$

$$\hat{e}_{si} = S_i - \hat{S}_i \quad \text{MVN parameter is } \hat{\sigma}_s$$

- The Multivariate Component in Step III is:

$$\begin{vmatrix} 1 & \rho_{\hat{e}_c, \hat{e}_w} & \rho_{\hat{e}_c, \hat{e}_s} \\ & 1 & \rho_{\hat{e}_w, \hat{e}_s} \\ & & 1 \end{vmatrix}$$

Simulation of a MVN Distribution

	A	B	C	D
1	<u>Variable</u>	<u>Mean</u>	<u>Std. Dev.</u>	
2	Cotton yield	\hat{C}	$\hat{\sigma}_C$	
3	Wheat yield	\hat{W}	$\hat{\sigma}_W$	
4	Sorghum yield	\hat{S}	$\hat{\sigma}_S$	
5				
6	Correlation Matrix of Stochastic Components			
7		Cotton	Wheat	Sorghum
8	Cotton	1.0	$\rho_{C,W}$	$\rho_{C,S}$
9	Wheat		1.0	$\rho_{W,S}$
10	Sorghum			1.0
11				
12				
13	Correlated SND's			
14	=CSND (B8:D10)			
15				
16				
17	Correlated yields			
18	Cotton	=B2 + C2 * A14		
19	Wheat	=B3 + C3 * A15		
20	Sorghum	=B4 + C4 * A16		

Note this is an array function
Press Ctrl Shift Enter

– See DemoMulti-VariateNormal.XLS and Lecture 15 Multivariate Distrib.XLS

3 Variable Case for a Non-Normal Multivariate Distribution (MVE)

- Step I Deterministic Component come from econometric analyses

$$\hat{C}_i = \hat{a} + \hat{b} T_i + \hat{b}_2 C_{i-1}$$

$$\hat{W}_i = \bar{W}$$

$$\hat{S}_i = \hat{a} + \hat{b} T_i$$

- Step II Stochastic Component is the unexplained portion of each variable

$$\hat{e}_{ci} = C_i - \hat{C}_i$$

$$\hat{e}_{wi} = W_i - \hat{W}_i$$

$$\hat{e}_{si} = S_i - \hat{S}_i$$

- Step III convert the stochastic component into the Empirical Distributions parameters S_{ij} for each variable j

$$S_{ic} = \text{Sorted} (\hat{e}_{ci} / \hat{C}_i)$$

$$S_{iw} = \text{Sorted} (\hat{e}_{wi} / \hat{W}_i)$$

$$S_{is} = \text{Sorted} (\hat{e}_{si} / \hat{S}_i)$$

- Step IV Multivariate Component calculated using unsorted residuals from Step II

$$\begin{vmatrix} 1 & \rho_{\hat{e}_c, \hat{e}_w} & \rho_{\hat{e}_c, \hat{e}_s} \\ & 1 & \rho_{\hat{e}_w, \hat{e}_s} \\ & & 1 \end{vmatrix}$$

- The MVE for j variables is defined by:

\hat{X}_j = deterministic component

$S_{\hat{e}ji}$ = stochastic component

$P(S_j)$ = probabilities of $S_{\hat{e}ji}$

ρ_{jj} = multivariate component

Simulation of a MVE Distribution

	A	B	C	D
1		<u>Cotton Yield</u>	<u>Wheat Yield</u>	<u>Sorghum Yield</u>
2	Mean	195	30	40
3	<u>P(S_i)</u>	<u>S_c</u>	<u>S_w</u>	<u>S_s</u>
4	0.0	-.11	-.21	-.08
5	0.3	-.1	-.2	-.07
6	0.5	0	-.1	0
7	0.7	.2	0	.30
8	0.9	.3	.4	.38
9	1.0	.31	.41	.39
10	Correlation Matrix			
11	Cotton	1.0	-.6	.4
12	Wheat		1.0	-.3
13	Sorghum			1.0
14				
15	Correlated USD's			
16	=CUSD (B11:D13)			
17				
18				
19	Correlated Yields			
20	= B2 * (1.00 + EMPIRICAL (B4:B9, A4:A9, A16))			
21	= C2 * (1.00 + EMPIRICAL (C4:C9, A4:A9, A17))			
22	= D2 * (1.00 + EMPIRICAL (D4:D9, A4:A9, A18))			
23				

← Note this is an array function

Validation of a MVN or MVE Distribution

- Simulate the model and specify the random variables as the output variables
- Check the simulated values for the output variables
 - T test and F test between simulated and historical series for each variable
 - Test the correlation matrix used as input against the implicit correlation matrix for the simulated output using t tests
 - Use the Simetar hypothesis tests for correlation

Part 4: Selecting the Best Scenario

- Scenarios are “re-runs” of the model using the same random values but different assumed values for decision variables
- Simulation generates 100+ iterations worth of outputs for each scenario. Now which is best?

Scenario Analysis

- The model is held constant but change one or more control variables
 - Number of scenarios is held constant
 - Random number Seed is held constant
 - Use =Scenario() functions to increment the control variables

- Example is to change acres and fixed costs in the model

$$\tilde{Y} = 150 + 30 * \text{SND}$$

$$\tilde{P} = 2 + 0.4 * \text{SND}$$

$$\text{Cost} = \text{FC} + 150 * \text{Acre} + \text{Acre} * 0.25 * \tilde{Y}$$

$$\text{Receipts} = \text{Acre} * \tilde{Y} * \tilde{P}$$

$$\text{Net Return} = \text{Receipts} - \text{Cost}$$

$$\text{Acre} = \text{Scenario} (100, 200, 300, 400)$$

$$\text{FC} = \text{Scenario} (10, 1000, 3000, 6000)$$

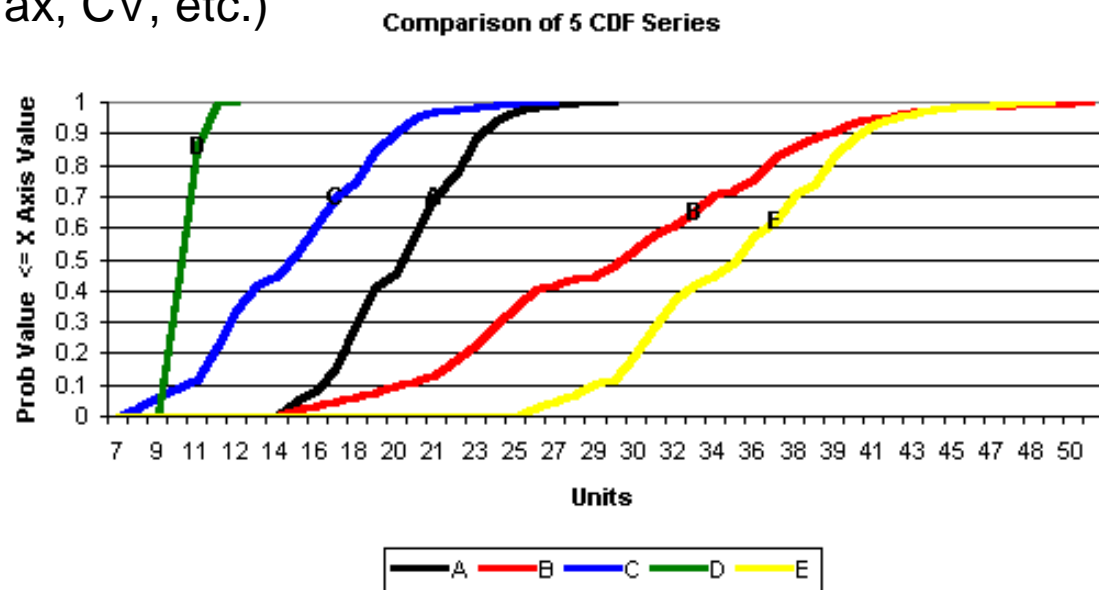
- See Demo Scenario.XLS

Example of a Scenario Table – 4 Scenarios on 2 Variables

	A	B	C	D	E	F
4						
5	VC/Acre	150			Conduct Scenario Analysis on FC and Acres	
6	VC/Yield	0.25			Acre	FC
7	Acre	100	=E11		100	10
8	Fixed Cost	10	=E12		200	1000
9	Yield Mean & Std. Dev.	150	30		300	3000
10	Price Mean & Std. Dev.	2	0.4		400	6000
11					= scenario (E7, E8, E9, E10)	
12					= scenario (F7, F8, F9, F10)	
13						
14	Random Yield	131.867		=norm(B9,C9)		
15	Random Price	2.295		=norm(B10,C10)		
16	Receipts	30,264		=B14*B15*B7		
17	Variable Costs	18,307		=B8+B7*B5+B6*B7*B14		
18	Net Returns is the KOV	11,957		=B16-B17		
19						

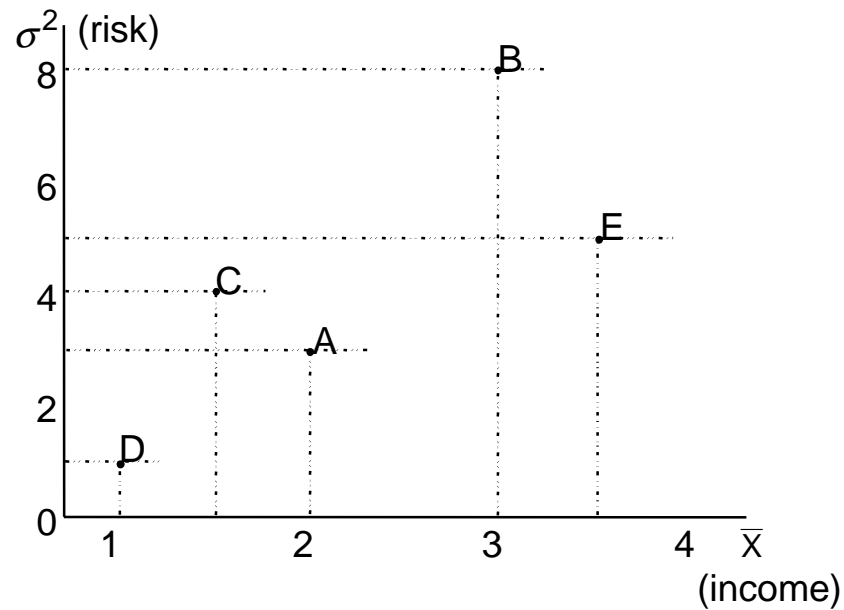
What's Best?

- After you simulate multiple scenarios you must help the managers pick the best one
- Role of the analyst is to present consequences not recommendations
- Economists are always looking for new ways to rank scenarios
- Simulation results for scenario analyses can be displayed several ways, for example:
 - Statistics (\bar{X} , σ , min, max, CV, etc.)
 - CDF graphs



Rank scenarios based on statistics for key output variables

- **Mean only** – throw away the risk
 - **Standard deviations** – ignores the level of returns
 - **Mean Variance** – often difficult to read, does not work well for non-normal distributions. Select alternative in lower right quadrant.
- From diagram below A preferred to C, E preferred to B

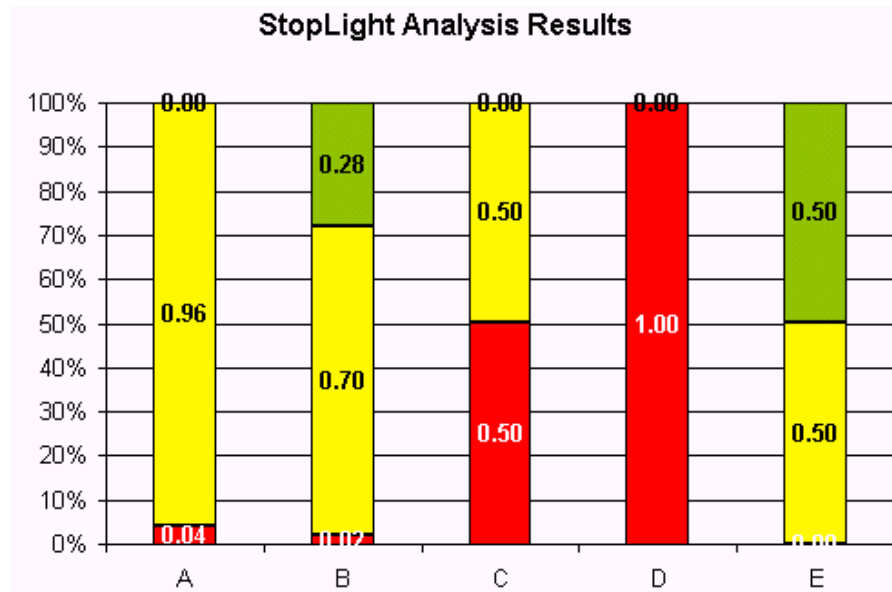


Rank Scenarios Methods (continued)

- **Worst case** – ignores the level of returns but has merit in that it avoids catastrophic losses, ignores overall risk. Bases decisions on only one observation that has a 1% chance.
- **Best case** – ignores the overall risk and downside potential risk. Looks at only one iteration that has a 1% chance.
- **Mini Max** – minimize the chance of a maximum error; select strategy with smallest minimum range between mean and minimum. Idea is to select strategy with lowest chance of a bad outcome. Ignores the level of returns and upside potential.
- **Relative Risk** – Coef. Variation (CV), pick the one that is lowest in absolute value. Easy to use, considers risk relative to the level of returns.
- **Probabilities of Target Values** – Calculate and report probability of achieving a preferred target and probability of failing to achieve a minimum target.

Rankings Based on Probabilities of Target Values.

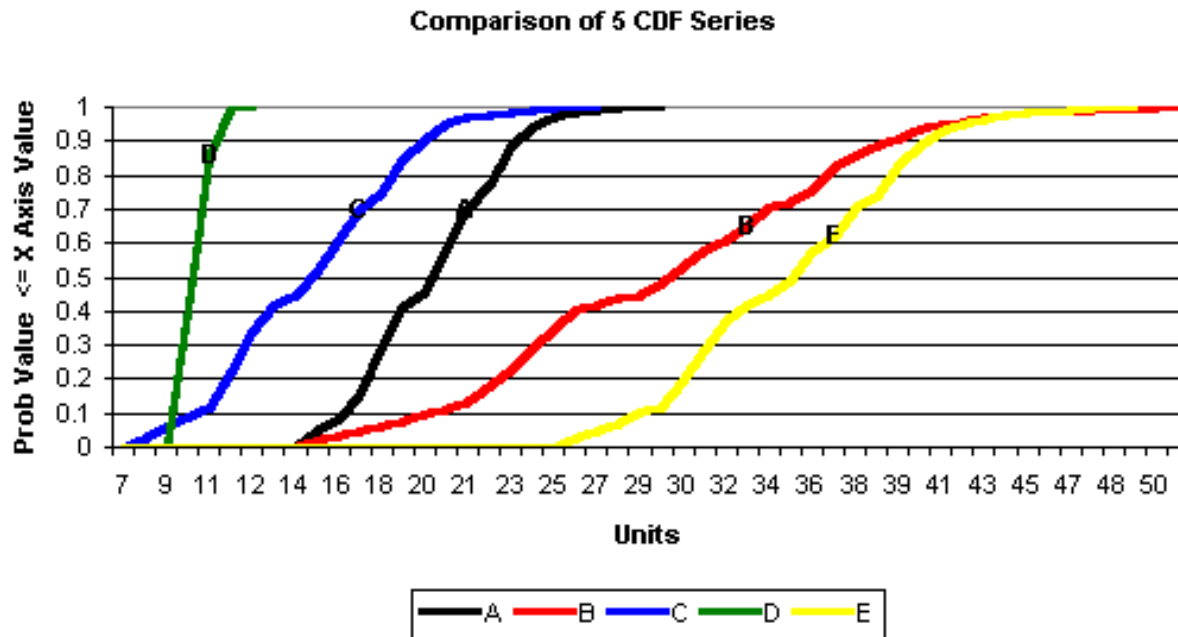
Rate of Return	Alternative Scenarios					
	A	B	C	D	E	
			(percent)			
Prob (X > 20%)	49.6	89.5	6.5	0	100	
Rank wrt P(X>20%)	3	2	4	5	1	
Prob (X>30%)	0	49.6	0	0	83.2	
Rank wrt P(X>30%)	3	2	4	5	1	
Prob (X>35%)	0	27.6	0	0	49.6	
Rank wrt P(X>35%)	3	2	4	5	1	



Full Distribution or CDF Visual Inspection

Table 10.7. Ranking Based on CDF Graphs.

<u>Rank by CDF Graph</u>	<u>Scenarios</u>
3	A
2	B
4	C
5	D
1	E



Stochastic Dominance

Considers the full range of the simulated distributions. Every point is used and compared one-to-one with every point of the other distributions.

- Draw back is it requires a feel for decision makers RAC or risk aversion coefficient
 - RAC < 0 risk lover
 - RAC > 0 risk averse
 - RAC = 0 risk neutral

- The larger the RAC the more risk averse

- Compares scenarios in pair wise fashion for two risky scenarios A vs. B
 - 1st degree A preferred if it is always preferred (never cross)
 - 2nd degree A preferred if the sum of (A > B) differences exceeds the sum of (B > A) differences
 - 3rd degree A preferred if weighted utility of A exceeds that of B

Results of Stochastic Dominance Analysis

@ 2000 Texas Agricultural Experiment Station

Lower Risk Aversion Ranking

Lower RAC **0**

Series Name Level of Preference

1 E	Most Preferred
2 B	2nd Most Preferred
3 A	3rd Most Preferred
4 C	4th Most Preferred
5 D	Least Preferred

Upper Risk Aversion Ranking

Upper RAC **1**

Series Name Level of Preference

1 E	Most Preferred
2 B	2nd Most Preferred
3 A	3rd Most Preferred
4 C	4th Most Preferred
5 D	Least Preferred

Summary Statistics

Series Name	Mean	Std Dev	CoefVar	Skewness	Minimum
1 A	19.90	2.90	14.59	0.26	14.33
2 B	29.73	7.74	26.03	0.26	14.87
3 C	14.86	3.87	26.03	0.26	7.44
4 D	9.97	0.55	5.47	0.08	9.08
5 E	34.83	4.84	13.89	0.26	25.55

Certainty Equivalents Under Alternative Utility Functions

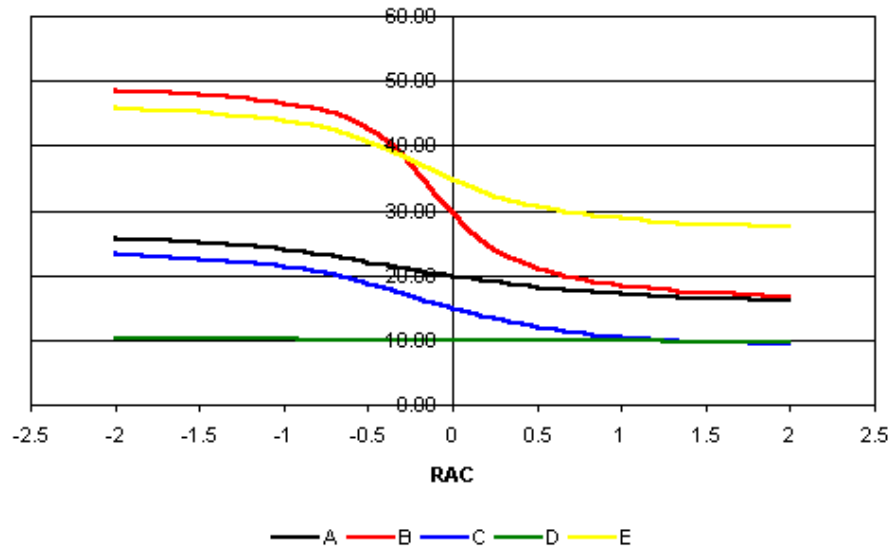
Series Name	CE Under Exponential Utility		CE Under Power Utility	
	Lower	Upper	Lower	Upper
3 A	19.90	17.12	19.90	18.01
2 B	29.73	18.51	29.73	21.09
4 C	14.86	10.55	14.86	10.54
5 D	9.97	9.83	9.97	9.83
1 E	34.83	28.86	34.83	31.82

Confidence Premiums Between Probability Distributions

Dominant Series Name	Lower Bound		Upper Bound	
	Absolute Amount	Percent of Dominant Mean	Absolute Amount	Percent of Dominant Mean
The Most Dominant Distribution				
1 E	5.10	14.65	10.36	17.68
2 B	14.93	42.87	11.74	51.73
3 A	19.97	57.33	18.31	69.17
4 C	24.86	71.37	19.03	86.11

Certainty Equivalents (CE)

Dynamic Certainty Equivalent Chart



- Specify lower and upper RACs and Simetar develops a table and two charts
- CE procedure is like having an infinite number of stochastic dominance runs in a graph
- Pick the highest strategy at any given RAC level. Specify a wide enough range of RACs so the distributions are spread out.
- Allows one to observe Breakeven RACs

Part 5: Business Analysis in a Risk Setting

- Twelve different program examples are available
- Pick one and experiment with it
- We are here to assist you

Part 6: Experiences Teaching Simulation

- Purdue Experience
 - Classroom
 - Agribusiness Center
- Texas A&M Experience
 - Classroom
 - Research and Extension application – can farmers understand CDFs?
- Discussion

The End

- Thank you for attending this Post-Conference Workshop
- Policy on Simetar.







Simulation for Risk Analysis

Instructors: James W. Richardson and Allan W. Gray
Texas A&M University and Purdue University



July 31, 2002
Overall
EVALUATION

PART I

Overall, how satisfied are you with your experience in this program?
(circle one)

Dis-Satisfied					Very Satisfied
1	2	3	4	5	

Class Size: 17
Responses: 17

Mean 4.53 Std. Dev. 0.51

What single thing was **most beneficial** to you and why?

- The empirical distribution component using data to develop a distribution and then simulating multivariate distributions
- Probably the discussion of 'quasi' random variables. It was not something about which I'd heard and I found it fascinating.
- Everything was excellent
- The ease of use of the statistical applications
- Introduction to Simetar Software! This software provides a great deal of flexibility and incorporates functions that have helped increase my efficiency in current research.
- Software, it is run on Excel and provides so easy way of simulation. This is amazing!
- "The exposure to Simetar software and the broad range of application and research that can be done with it
- The presence of the students to help some of us out was also beneficial. "
- get to know members of our profession who also teaches risk analysis
- While I had to leave the meeting early and could not actually attend the workshop, I received the book and software in advance. I have had a chance to work through much of the material in the book and play with the software. The book is clear and the software seems easy and straightforward to use. I am very

happy that I paid for both even though I could not attend the workshop.
Thanks.

- I couldn't pick a "single" thing the whole experience was great.
- stochastic simulation from empirical distribution because I am using it for my PhD dissertation.
- Explanation of mechanics of simulation process.
- Learning to make correlated yield/price draws that preserve historic spatial correlation. I have wanted to do this for some time but the tools I had were not near as useful as Simetar.
- The examples that we discussed in the session and that because it showed me what and when to use each of the batten in the program
- Bootstrap. It is much easier than other software.

What single **improvement** would you suggest?

- Development of a function key for the array command instead of the three keys needed now.
- More time!
- Two day workshop
- Everything was well done, I wish more time was allocated to the workshop
- None yet.
- Go slower :-)
- I was already exposed to simetar so it was ok for me but I saw many struggling. I think we were very optimistic in designing this training. We wanted to cover a lot of materials on the same day. We should only focus on what are the things simetar can do? How it can do? We will have to assume that participants know what is mean, median, mode, Cv, distributions, risk efficiency criteria, time series and so on.
- Narrow scope and focus on one example.
- I would greatly benefit from a longer course. This would allow more examples and using my own data and asking instruction on my specific cases.

PART II

Rate each of the following sessions for:

4 Introduction to Simulation Modeling

Richardson

Dis-Satisfied					Very Satisfied
1	2	3	4	5	

Class Size: 17

Responses: 14

Mean 4.57

Std. Dev.

0.51

Good job

- I cannot comment on the actual presentations, but the related material in the book and examples are very clear to me. You all have done a great job.
- Good job.

5 Simulation of Random Variables

Richardson

Dis-Satisfied					Very Satisfied
1	2	3	4	5	

Class Size: 17

Responses: 14

Mean 4.43

Std. Dev.

0.65

- Good job
- Slow down in examples.

6 Parameter Estimation for Univariate Distributions

Richardson

Dis-Satisfied	1	2	3	4	Very Satisfied	5
---------------	---	---	---	---	----------------	---

Class Size: 17

Responses: 14

Mean	4.50	Std. Dev.	0.76
------	------	-----------	------

- Good job
- Good

7 Simulating Univariate Probability Distributions

Gray

Dis-Satisfied	1	2	3	4	Very Satisfied	5
---------------	---	---	---	---	----------------	---

Class Size: 17

Responses: 14

Mean	4.50	Std. Dev.	0.65
------	------	-----------	------

- Great job
- Did a good job.

8 Simulating Multivariate Probability Distributions

Richardson

Dis-Satisfied	1	2	3	4	Very Satisfied	5
---------------	---	---	---	---	----------------	---

Class Size: 17

Responses: 14

Mean	4.50	Std. Dev.	0.65
------	------	-----------	------

- Great job
- A little bit fast!

9 Selecting the Best Scenario

Gray

Dis-Satisfied				Very Satisfied
1	2	3	4	5
<hr/>				

Class Size: 17

Responses: 14

Mean 4.43 Std. Dev. 0.76

- This is an area we could have used more time . . . like another day.
- Time was the constraint, but It was well presented
- Interesting. I see applications in finance.

10 Experience Teaching Simulation

Gray and Richardson

Dis-Satisfied				Very Satisfied
1	2	3	4	5
<hr/>				

Class Size: 17

Responses: 14

Mean 4.43 Std. Dev. 0.65

- I don't teach, so it was not particular interesting for me.
- I like the passion of both Dr. Gray and Dr. Richardson
- Excellent.

PART III

Rate each of the following sections:

11 Effectiveness of James Richardson

Dis-Satisfied				Very Satisfied
1	2	3	4	5

Class Size: 17
Responses: 13

Mean	4.54	Std. Dev.	0.52
------	------	-----------	------

Did a good job.

12 Effectiveness of Allan Gray

Dis-Satisfied				Very Satisfied
1	2	3	4	5

Class Size: 17
Responses: 14

Mean	4.43	Std. Dev.	0.65
------	------	-----------	------

Effective.

13 Effectiveness of Assistants

Dis-Satisfied				Very Satisfied
1	2	3	4	5

Class Size: 17
Responses: 14

Mean	4.43	Std. Dev.	0.76
------	------	-----------	------

- Extremely helpful!
- Their presence was a wonderful idea
- they were effective in finding the problems of participants and communicating them to instructors.
- They were necessary and they all did a good job.

PART IV

Overall, how satisfied were you with the simulation software?

Dis-Satisfied	1	2	3	4	Very Satisfied	5
<hr/>						

Class Size: 17

Responses: 16

Mean 4.63 Std. Dev. 0.50

Was the information received informative and applicable to your classes or research?

14 Yes 1 No

Make any additional comments you would like:

- Good seminar, would like more time to look at other specific things that I am interested in.
- Excellent seminar - Drs. Richardson and Gray did a phenomenal job in this seminar. I would definitely be interested if they ever offered an advanced seminar on Simetar.
- "Would like to attend another training session of Risk simulation and look forward to information about that.
- I would like to thank Dr. Gray, Dr. Richardson and the Assistant for their dedication"
- I'm sorry but I don't recall details of the workshop to provide section-specific feedback. I really enjoyed the workshop. It was worth every penny. Despite its speed (which I appreciated), it was feasible, thanks to the number and quality of assistants you had.
- Thanks.
- I am hoping to use this in a class next year.
- when you simulate results appear in simdata sheet but any correction to the parent sheet is not "dynamically" linked to the simdata sheet. Perhaps I need to use macro. I wanted to learn it but I saw many participants struggling with the basics so decided not to ask.
- Excellent workshop. We need another one. Perhaps with more of a focus. Select application and carry it all way through. Overall, enjoyed very much and gained useful information.
- I really learned a lot in the class. Thanks for offering it and making such useful and affordable software.