Formulating SP\
Stochastic Programming\
Scenario Planning
Models in
What's*Best*!

www.lindo.com

December 2011

LINDO SYSTEMS INC.

Modeling Uncertainty in General Optimization Problems

Is there a general way of incorporating probabilistic uncertainty into optimization problems?

Yes, goes by the name, “Stochastic Programming(SP)”.

Can also perhaps more suggestively think of it as
Scenario Planning(SP).

Basic idea is to use a finite number of scenarios, each with a specified probability.

May have a multi-period sequence of random events.

Why Use SP?

If uncertainty is a significant factor:

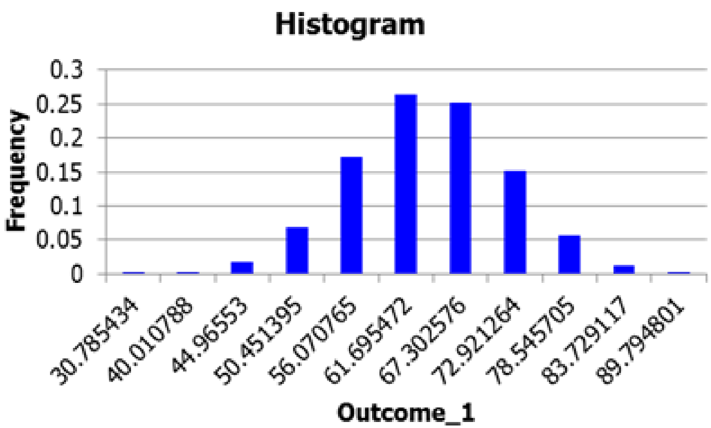
1) Simple deterministic analysis may suggest a solution far from optimal,
e.g., stocking to exactly meet expected demand may
miss the high profit of occasional really high demand.

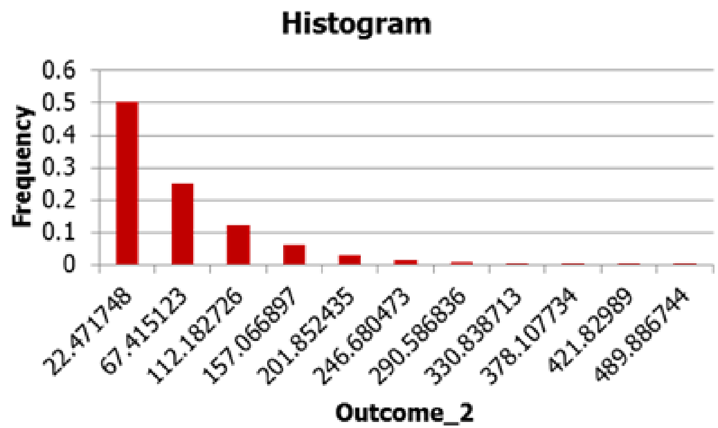
2) Simple scenario-by-scenario analysis, may miss the optimal solution,
e.g., the solution that is optimal when all scenarios are taken into account
may not be optimal for any single scenario.

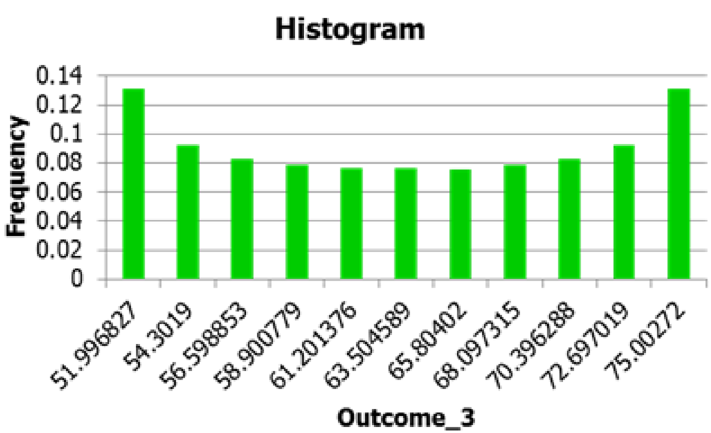
3) Simple expected value analysis, even if it takes into account uncertainty,
may miss the fact that we really care about the distribution of outcomes,
e.g., the low probability but catastrophic outcome.
SP optimization supplies you with the distribution of outcomes.
You may have two or more random variables with the same mean
and standard deviation, but dramatically different distributions…

Perhaps We Should Be Concerned About the Distribution…

Here are the histograms of three random variables, each with *Mean*= 64, *SD*= 8.







LINDO SYSTEMS INC.

Multi-Stage Decision Making Under Uncertainty

Stochastic programming, or Scenario Planning, or SP for short, is an approach for solving problems of multi-stage decision making under uncertainty. SP is designed to solve problems of the following form:

0) In stage 0 we make some decisions, taking into account that later,

1) At the beginning of stage 1, "Nature" makes a random decision,
1a) At the end of stage 1, having seen nature's decision, as well as our previous decisions, we make some decisions, taking into account that …

2) Somewhat later at the beginning in stage 2, "Nature" makes a random decision,

…

n) At the beginning of stage *n*, "Nature" makes a random decision, and
n.a) At the end of stage *n*, having seen all of nature's *n* previous decisions, as well as all our previous decisions, we make a decision,

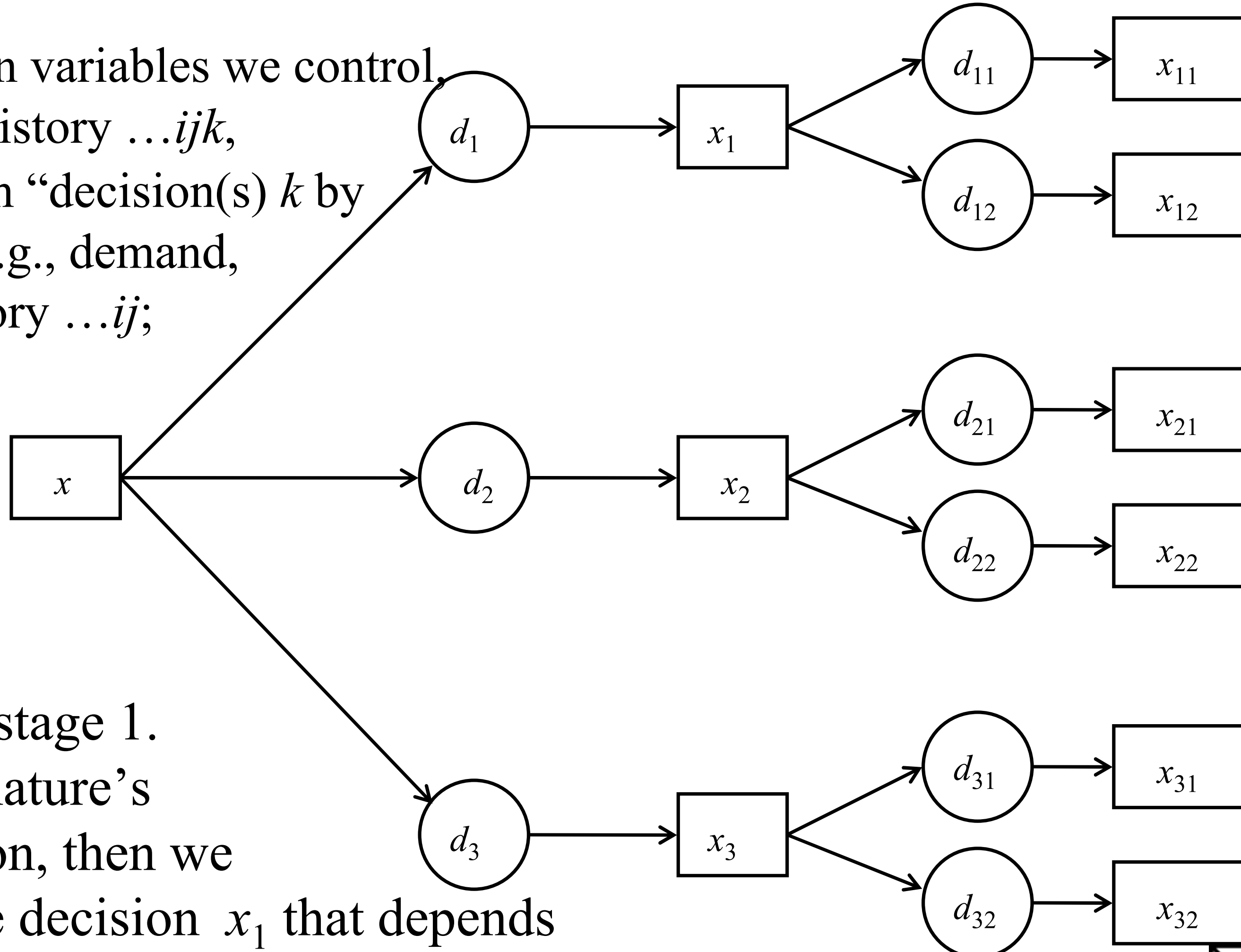
If there are only a finite number of outcomes(which is true computationally) for nature at each stage, then it may be helpful to visualize the process by a tree.

Viewed as a Tree…

Notation:

$x_{\ldots ijk}$ = decision variables we control, given history …*ijk*,

$d_{\ldots ijk}$ = random "decision(s) *k* by nature", e.g., demand, given history …*ij*;



For this tree:
3 possible outcomes in stage 1.
Once we see nature's stage 1 decision, then we make a unique decision x_1 that depends upon nature's decision, etc.

Applications of SP, Some Examples

+ Financial portfolio planning over multiple periods for insurance and other financial companies, in the face of uncertain prices, interest rates, exchange rates, and bankruptcies,

+ Capacity and Production planning in the face of uncertain future demands and prices,

+ Fuel purchasing when facing uncertain future fuel demand and prices,

+ Optimal exploration planning for petroleum companies,

+ Foundry metal blending in the face of uncertain input scrap qualities,

+ Fleet assignment: vehicle type to route assignment in the face of uncertain route demand,

+ Electricity generator unit commitment in the face of uncertain demand,

+ Hydro management and flood control in the face of uncertain rainfall,

+ Optimal time to exercise for options in the face of uncertain prices,

+ Product planning in the face of future technology uncertainty,

+ Revenue management in the hospitality and transport industries.

Some generic but common two stage (0 and 1), examples:

Example 1: Capacity Planning (Multi-dimensional Newsvendor)

Stage 0, decisions:

x_i = capacity installed of type i; made before seeing demand,

Stage 1 beginning, random events observed:

d_{sj} = demand for product type j in scenario s, for $s = 1, 2, \ldots, ns$,

Stage 1 end:

y_{sij} = amount shipped from i to j if scenario is s;

Model:

Max = $-\Sigma_i \; c_i * x_i + \Sigma_s \Sigma_i \Sigma_j \, r_{ij} * y_{sij} / ns$; ! Assumes all scenarios equally likely;

For each scenario s and source i: ! Capacity constraints;

$\Sigma_j \, y_{sij} \leq x_i$;

For each scenario s and demand type j: ! Demand constraints;

$\Sigma_i \, y_{sij} \leq d_{sj}$;

2) Portfolio planning.

Stage 0, decisions:

x_i = amount invested in instrument i;

Stage 1 beginning, observe random outcomes:

r_{si} = return on investment in instrument i in scenario s, for $s = 1, 2, \ldots, ns$,

Stage 1 end:

y_s = return of portfolio if scenario is s,

u_s, d_s = deviation up, down of return from target;

Model:

$\Sigma_i\, x_i \le 1$; ! Compute Budget constraint;

For each scenario s :

$y_s = \Sigma_i\, r_{si} * x_i$; ! Compute scenario return;

$u_s - d_s = y_s - target$; ! Compute deviations from target;

$\Sigma_s\, y_s\, /ns \geq= target$; ! Expected return achieves target, all scenarios equally likely;

$\text{Min} = \Sigma_s\, d_s\, /ns$; ! Min downside risk;

SP Applications More Specifically

Plant configuration decisions, e.g., General Motors

Had too much capacity.

Needed to close or refocus an unknown number of plants.

Investment Portfolios at Insurance Companies,

e.g., Yasuda-Kasai in Japan.

Had been using Markowitz "mean-variance" portfolio optimization.

Markowitz assumes risks have a Normal distribution(symmetric)

Actual risks were too non-symmetric (This is insurance)

Multi-Stage Tree Structures in Practice…

General Motors used a 5 period, (but 2 stage) model:

Periods 1-4: The next 4 years,

Period 5: Year 5 and out to infinity modeled using present values.

Plant reconfiguration decisions were made only at beginning of year 1. No reconfiguration decisions thereafter.

General Motors historically made three forecasts, with associated probabilities, for each year, into the future.

Stage	Branches	Represents
1	3^5 = 243	Next 4 years + infinity

Total number of full scenarios = 243.

LINDO SYSTEMS INC.

GM SP Model, Special Features

+ Downside risk

+ Unsatisfied demand for a product transfers to other products according to a substitution matrix. One dozen products.

+ Infinite final period.

Key parameters:

c_{pv} = cost per unit to produce vehicle *v* in plant *p* (only possible if plant is open),

τ_{vw} = fraction of unsatisfied demand for vehicle *v* that transfers to vehicle *w*, (from surveys),

$CAP_{p\sigma}$ = capacity of plant *p* in configuration σ,

Key variables:

x_{spv} = number of units of vehicle *v* produced in plant *p* in scenario *s*.

GM Model: Inventory Balance Constraint

The key constraints in words are:

For each scenario *s*

For each vehicle *v:*

$$Production_{vs} + Unsat_{sv} = Demand_{sv} + Transfer\_in_{sv};$$

For each vehicle *v* and *w*:

$$Transfer\_from\_to_{svw} \leq \tau_{vw} * Unsat_{sv};$$

For each plant *p* and configuration *σ:*

$$Total\_production_{sp} \leq CAP_{p\sigma} * y_{p\sigma}$$

Downside Risk in GM Model

$$penalty_s \geq threshold - profit_s ;$$

Expected downside risk constraint:

$$\sum_s Prob_s penalty_s \leq tolerance;$$

Both threshold and tolerance are parameters.

Gas Purchasing at Peoples Gas as an SP Problem

General Features:

Two stages,

Stage 0, make purchase and storage decisions,

Stage 1: Ten scenarios, corresponding to ten previous representative weather patterns, scaled up to today.

Each scenario has 365 periods.

Storage costs are nonlinear, first units are easy to pump in, last units require much energy to pump in.

First units withdrawn can be withdrawn rapidly, last units can be withdrawn only slowly.

Contracts have daily min and max and total over all days.

Doing SP in either What's*Best*! or LINGO

Essential Steps:

1) Write a standard deterministic model (the core model) as if the random variables were variables or parameters.

2) Identify the random variables, and decision variables, and their staging.

3) Provide the distributions describing the random variables, [Why separate (2) and (3) ?]

4) Specify manner of sampling from the distributions, (mainly the sample size), and

5) List the variables for which we want a (What'sBest! only) scenario by scenario report or a histogram.

How is SP Information Stored in the SpreadSheet?

All information about the SP features is stored explicitly/openly on the spreadsheet.

1) Core model is a regular deterministic What'sBest! or LINGO model. You can plug in regular numbers in a random cell to check results.

2) Staging information is stored in
Decisions: WBSP_VAR(stage, cell_list) and
Random variables: WBSP_RAND(stage, cell_list);

3) Distribution specification is stored in
WBSP_DIST_distn(table, cell_list);
where distn specifies the distribution, e.g., NORMAL cell.

4) Sample size for each stage is stored in
WBSP_STSC(table);

5) Cells to be reported are listed in
WBSP_REP(cell_list) or WBSP_HIST(bins, cell);

LINDO SYSTEMS INC.

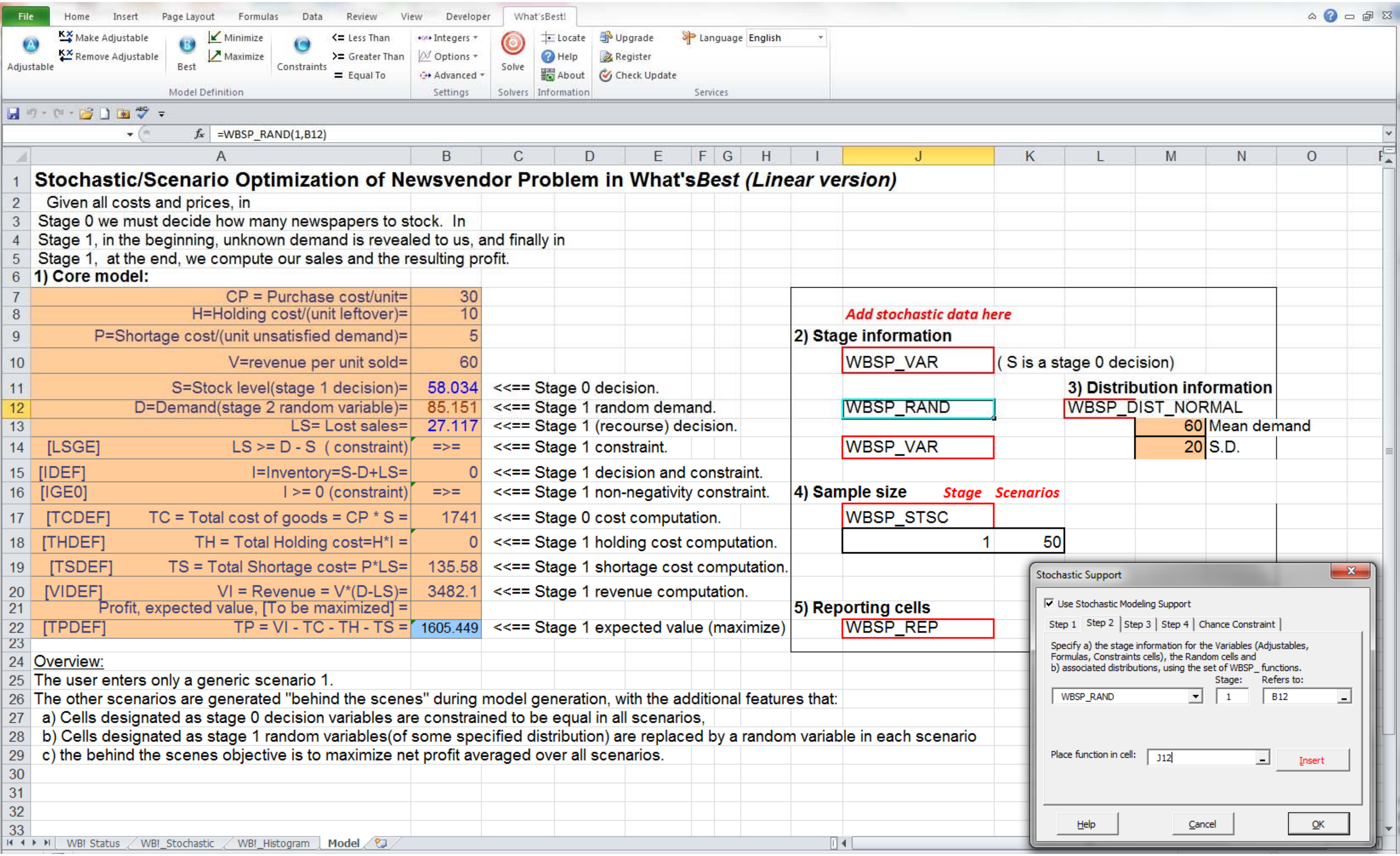
Core Comments

The “Core Model” is a completely valid Excel model.

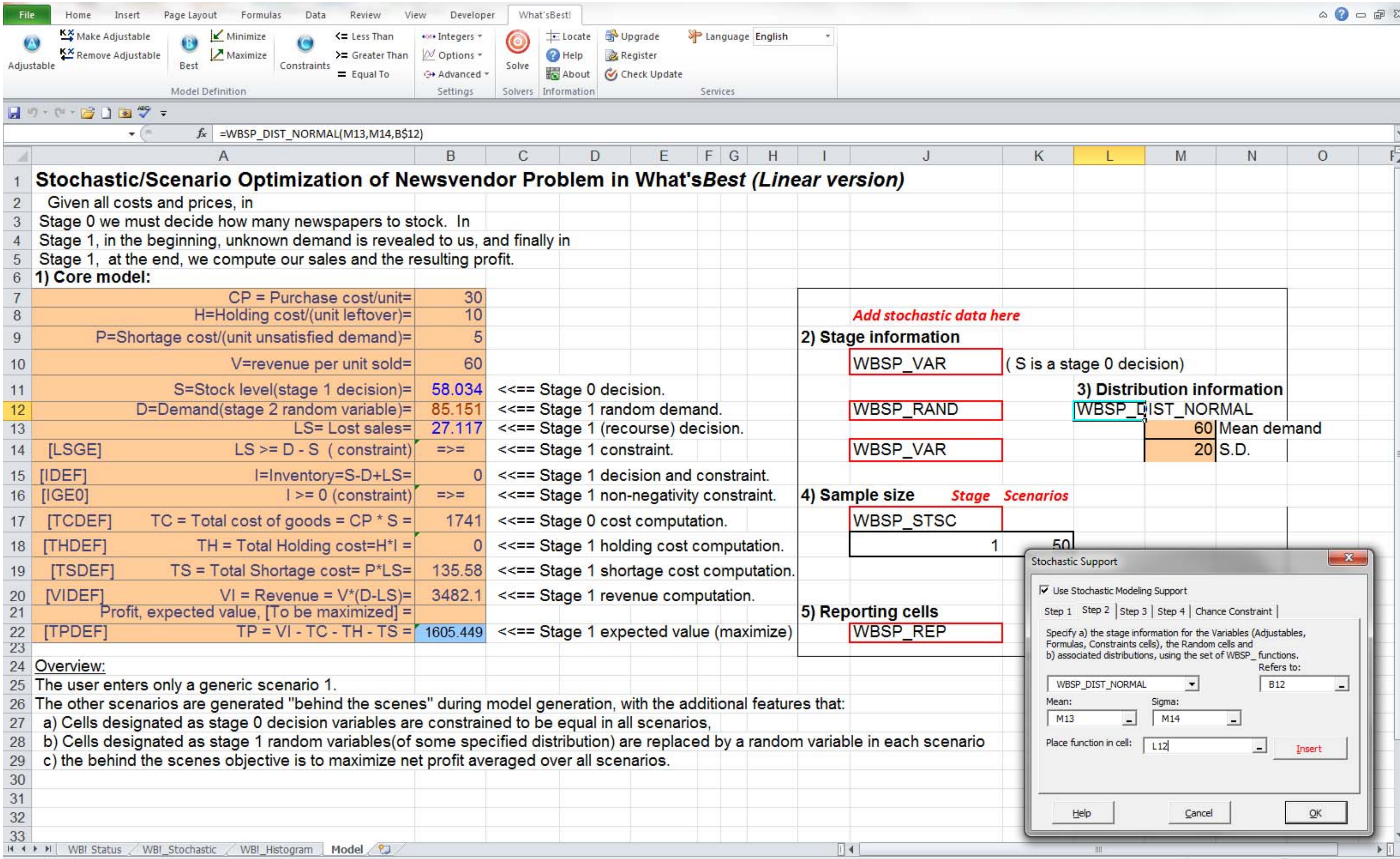
If you are doing neither simple optimization nor SP, you can do complete “What-If” analysis with it as a valid deterministic model.

If you have not turned on SP, you can do simple optimization with it like any deterministic What’sBest model.

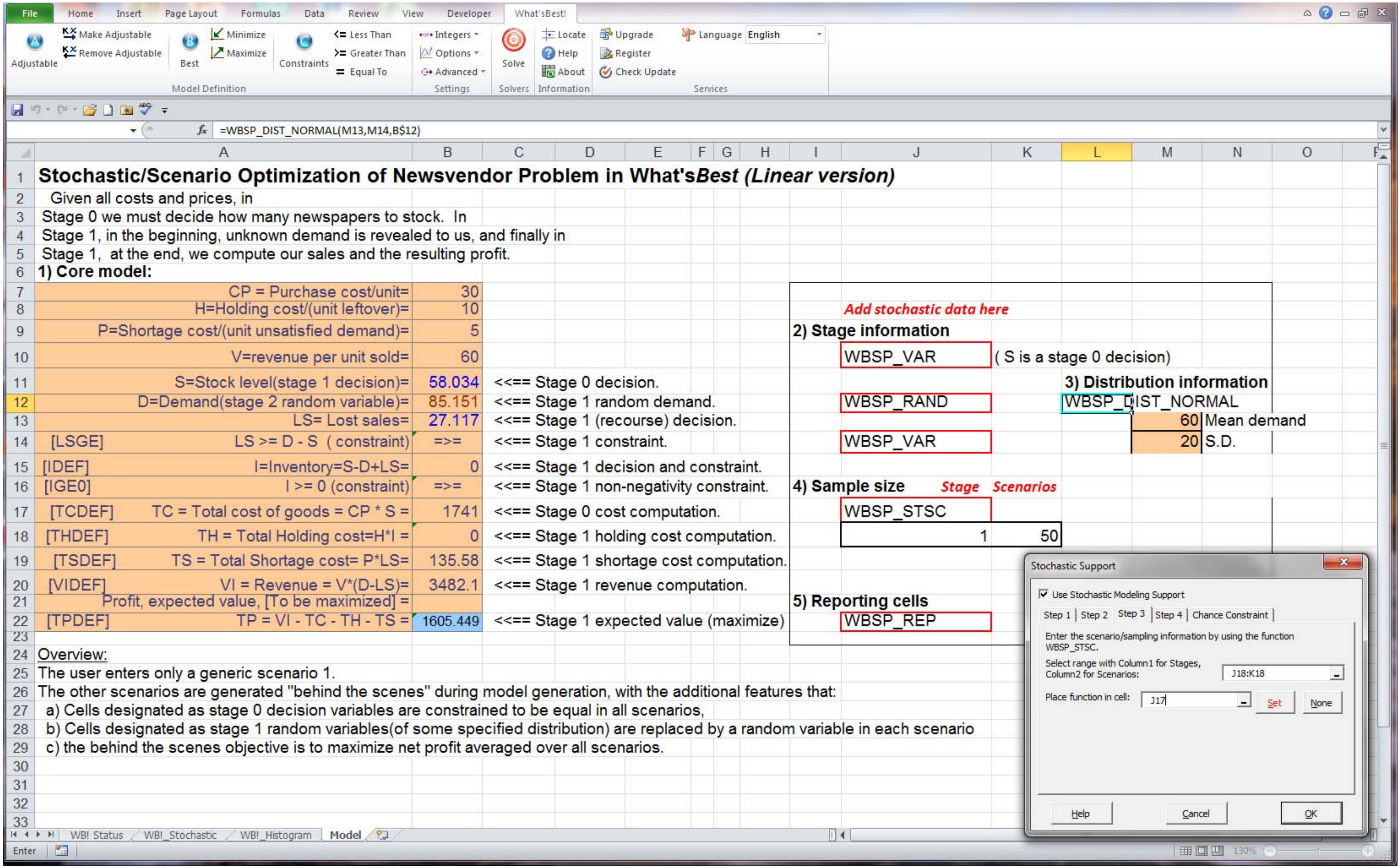
Input via a Dialog Box, Newsvendor, Steps 1, 2, Staging



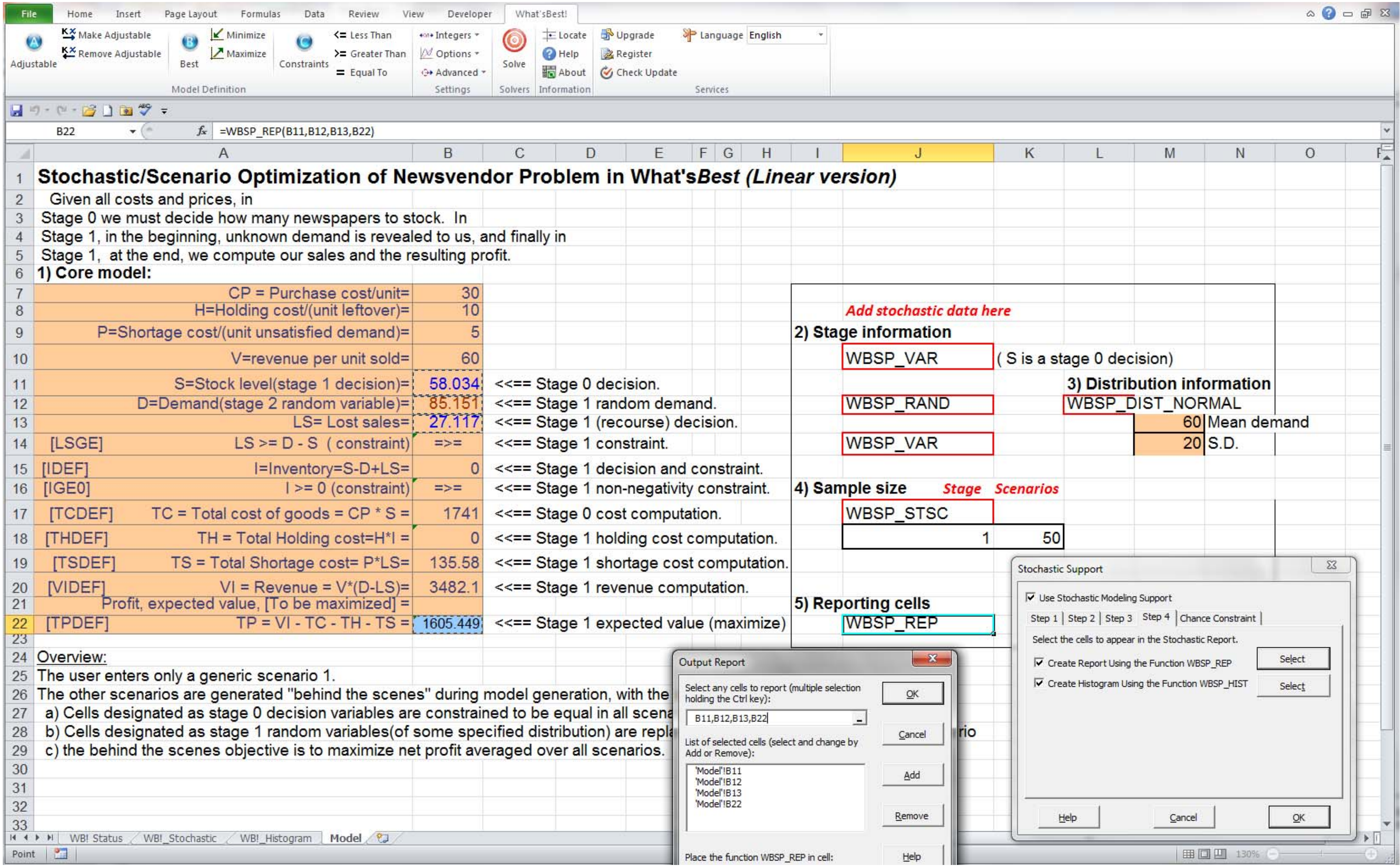
Input via a Dialog Box, Newsvendor, Step 3, Distribution



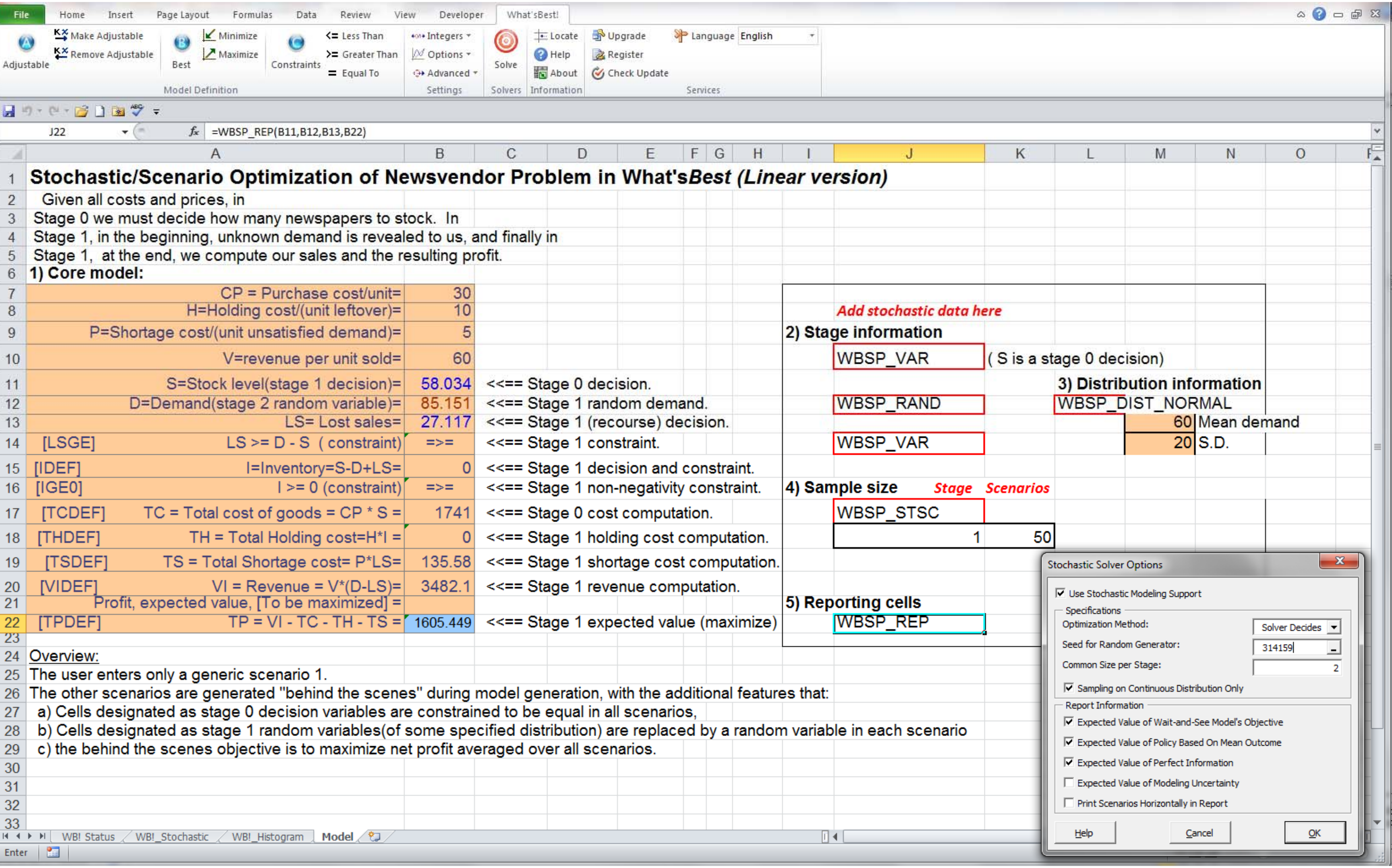
Input via a Dialog Box, Newsvendor, Step 4, Sample Size



Input via a Dialog Box, Newsvendor, Step 5 Reporting



Input via a Dialog Box, Setting Various Options



J22 =WBSP_REP(B11,B12,B13,B22)

Stochastic/Scenario Optimization of Newsvendor Problem in What'sBest (Linear version)

Given all costs and prices, in
Stage 0 we must decide how many newspapers to stock. In
Stage 1, in the beginning, unknown demand is revealed to us, and finally in
Stage 1, at the end, we compute our sales and the resulting profit.

1) Core model:

Name	Definition	Value	Comment
	CP = Purchase cost/unit=	30	
	H=Holding cost/(unit leftover)=	10	
	P=Shortage cost/(unit unsatisfied demand)=	5	
	V=revenue per unit sold=	60	
	S=Stock level(stage 1 decision)=	58.034	<<== Stage 0 decision.
	D=Demand(stage 2 random variable)=	85.151	<<== Stage 1 random demand.
	LS= Lost sales=	27.117	<<== Stage 1 (recourse) decision.
[LSGE]	LS >= D - S (constraint)	=>=	<<== Stage 1 constraint.
[IDEF]	I=Inventory=S-D+LS=	0	<<== Stage 1 decision and constraint.
[IGE0]	I >= 0 (constraint)	=>=	<<== Stage 1 non-negativity constraint.
[TCDEF]	TC = Total cost of goods = CP * S =	1741	<<== Stage 0 cost computation.
[THDEF]	TH = Total Holding cost=H*I =	0	<<== Stage 1 holding cost computation.
[TSDEF]	TS = Total Shortage cost= P*LS=	135.58	<<== Stage 1 shortage cost computation.
[VIDEF]	VI = Revenue = V*(D-LS)=	3482.1	<<== Stage 1 revenue computation.
	Profit, expected value, [To be maximized] =		
[TPDEF]	TP = VI - TC - TH - TS =	1605.449	<<== Stage 1 expected value (maximize)

Add stochastic data here

2) Stage information

WBSP_VAR (S is a stage 0 decision)
WBSP_RAND
WBSP_VAR

3) Distribution information

WBSP_DIST_NORMAL
60 Mean demand
20 S.D.

4) Sample size *Stage* *Scenarios*

WBSP_STSC
1 50

5) Reporting cells

WBSP_REP

Overview:
The user enters only a generic scenario 1.
The other scenarios are generated "behind the scenes" during model generation, with the additional features that:
a) Cells designated as stage 0 decision variables are constrained to be equal in all scenarios,
b) Cells designated as stage 1 random variables(of some specified distribution) are replaced by a random variable in each scenario
c) the behind the scenes objective is to maximize net profit averaged over all scenarios.

Input via a Dialog Box, Setting Various Options

Setting Retention:

Any settings made with a dialog box are retained when the workbook is saved. The same settings will be there when the workbook is next re-opened.

Settings such as Adjustable cells, constraints can be found by clicking on:

Add-Ins | WB! | Locate

Standard Scenario Report, One Line/Scenario

SP_NewsBoyN.xls [Compatibility Mode] - Microsoft Excel

STAGES	2
NODES	1001
SCENARIOS	1000
EXPECTED VALUE	1.803336e+003
Expected Value of Perfect Information Lower Bound	5.963933e+002

REPORTING CELLS

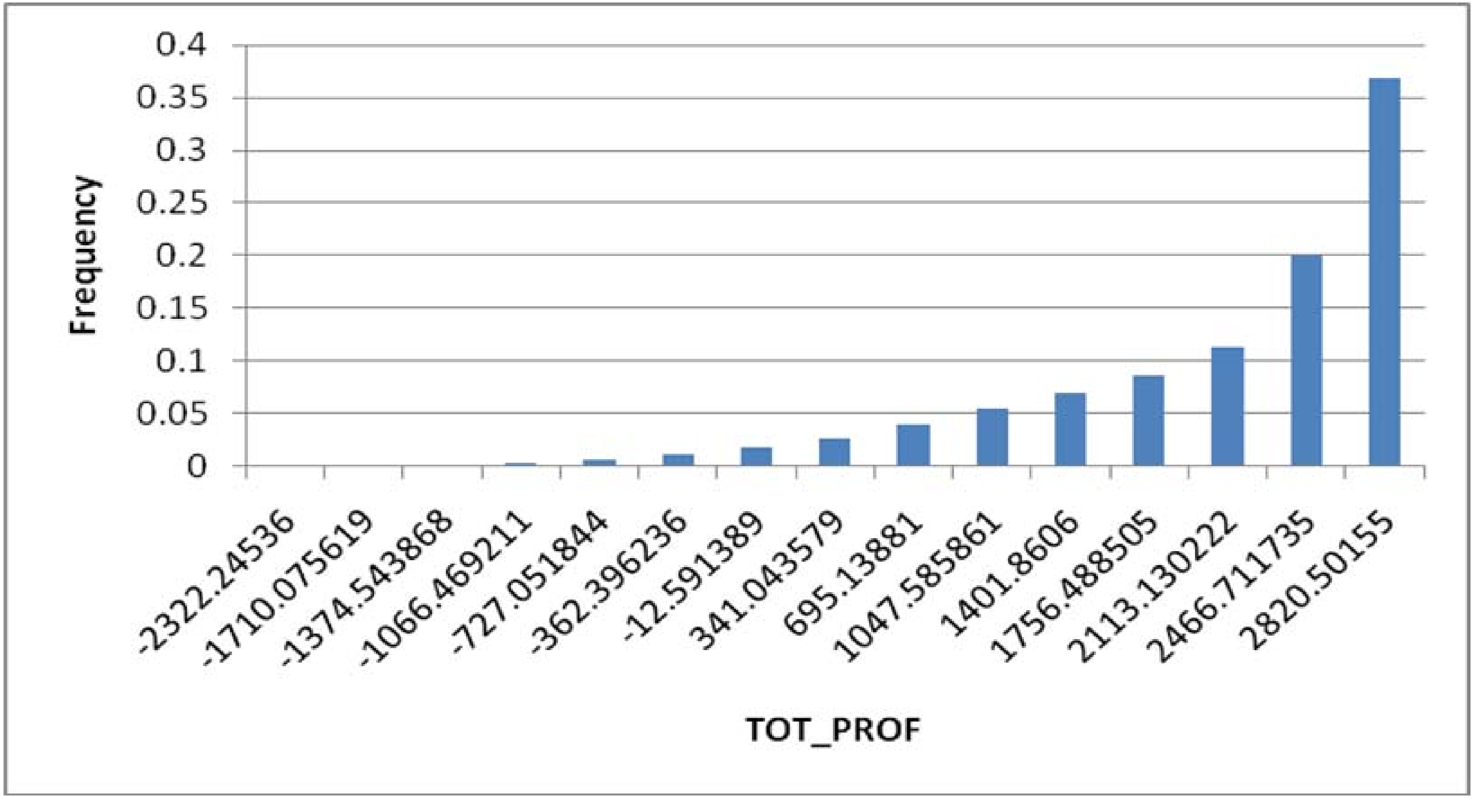
SCENARIO	Model!B11 S STAGE 0	Model!B12 D STAGE 1	Model!B13 LOSTSALES STAGE 1	Model!B22 TOT_PROF STAGE 1
– 1–	78.305639	104.552061	26.246422	2217.937053
– 2–	78.305639	82.832201	4.526562	2326.536352
– 3–	78.305639	56.621004	0	831.244757
– 4–	78.305639	72.035854	0	1910.284251
– 5–	78.305639	74.494865	0	2082.415009
– 6–	78.305639	98.879033	20.573394	2246.302194
– 7–	78.305639	88.135404	9.829765	2300.020336
– 8–	78.305639	87.484145	9.178506	2303.276633
– 9–	78.305639	96.468666	18.163027	2258.354027
– 10–	78.305639	68.556681	0	1666.742126
– 11–	78.305639	118.232267	39.926629	2149.53602
– 12–	78.305639	100.651738	22.346099	2237.438667
– 13–	78.305639	88.638484	10.332845	2297.504937
– 14–	78.305639	64.005719	0	1348.174764
– 15–	78.305639	92.479653	14.174015	2278.29909
– 16–	78.305639	59.770935	0	1051.739899
– 17–	78.305639	112.291194	33.985555	2179.241386
– 18–	78.305639	72.715838	0	1957.883143
– 19–	78.305639	111.322138	33.016499	2184.086669
– 20–	78.305639	74.572581	0	2087.855151
– 21–	78.305639	39.855992	0	–342.306132
– 22–	78.305639	93.293771	14.988132	2274.228504
– 23–	78.305639	78.964977	0.659339	2345.87247
– 24–	78.305639	52.418108	0	537.042035
– 25–	78.305639	59.609509	0	1040.440057
– 26–	78.305639	53.327633	0	600.708757
– 27–	78.305639	100.794988	22.48935	2236.722415
– 28–	78.305639	77.521508	0	2294.280039
– 29–	78.305639	70.892475	0	1830.247714
– 30–	78.305639	60.581691	0	1108.492814
– 31–	78.305639	48.289092	0	248.010883
– 32–	78.305639	104.998436	26.692797	2215.705178
– 33–	78.305639	82.888439	4.5828	2326.255163
– 34–	78.305639	95.214398	16.90876	2264.625366
– 35–	78.305639	71.40176	0	1865.897664
– 36–	78.305639	97.308513	19.002875	2254.15479

What does the distribution of Total Profit look like?

WB! Status | WB!_Stochastic | Model | WB!_Hist

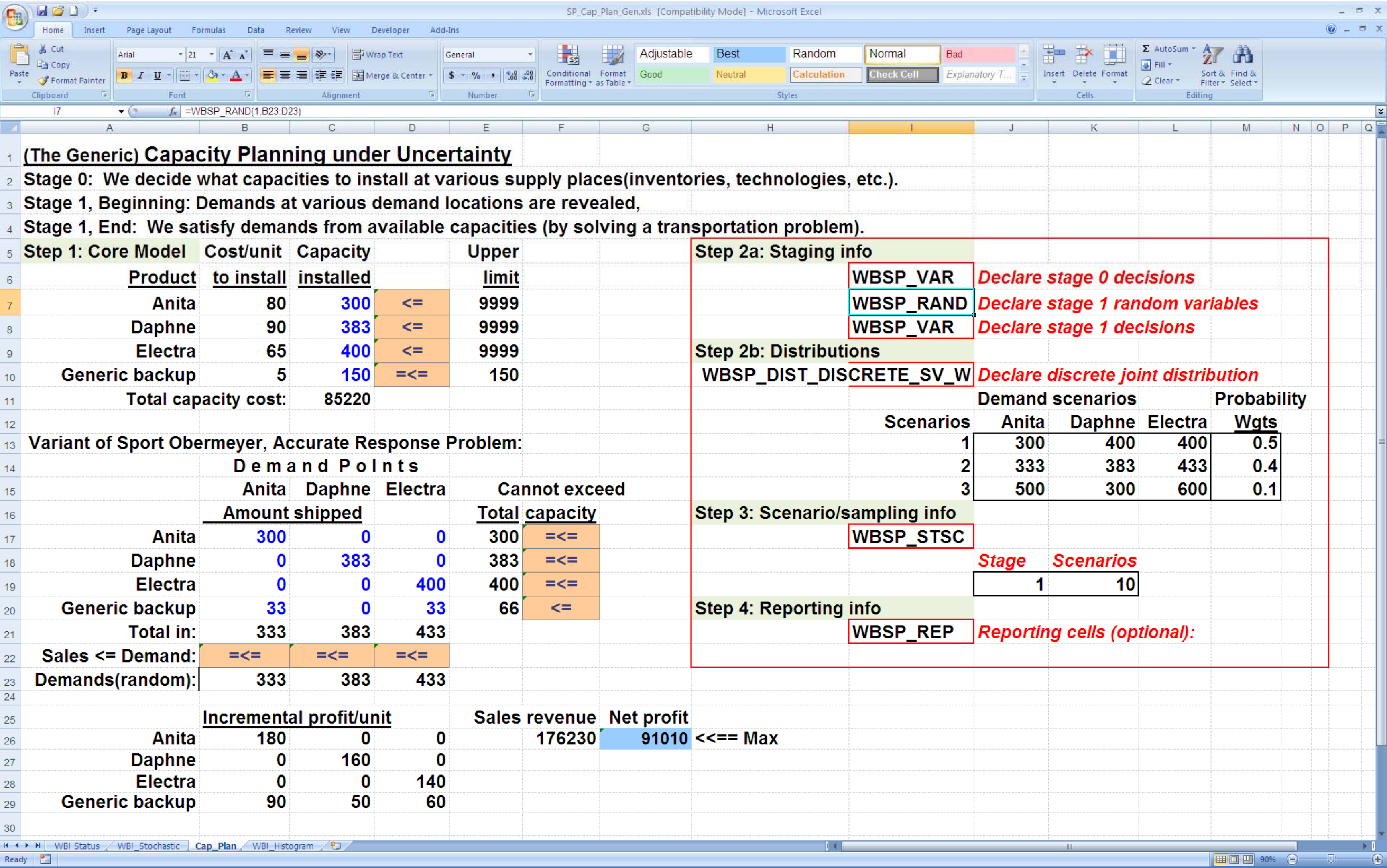
Ready

Newsvendor with Normal Demand



Even though the driving random variable, Demand, has a symmetric distribution, why is the output, Profit, so skewed?

The Generic Capacity Planning Under Uncertainty Model



SP_Cap_Plan_Gen.xls [Compatibility Mode] - Microsoft Excel

I7 =WBSP_RAND(1,B23:D23)

(The Generic) Capacity Planning under Uncertainty

Stage 0: We decide what capacities to install at various supply places(inventories, technologies, etc.).

Stage 1, Beginning: Demands at various demand locations are revealed,

Stage 1, End: We satisfy demands from available capacities (by solving a transportation problem).

Step 1: Core Model	Cost/unit	Capacity		Upper
Product	to install	installed		limit
Anita	80	300	<=	9999
Daphne	90	383	<=	9999
Electra	65	400	<=	9999
Generic backup	5	150	=<=	150
Total capacity cost:	85220			

Variant of Sport Obermeyer, Accurate Response Problem:

	Demand Points			Cannot exceed	
	Anita	Daphne	Electra		
	Amount shipped			Total	capacity
Anita	300	0	0	300	=<=
Daphne	0	383	0	383	=<=
Electra	0	0	400	400	=<=
Generic backup	33	0	33	66	<=
Total in:	333	383	433		
Sales <= Demand:	=<=	=<=	=<=		
Demands(random):	333	383	433		

	Incremental profit/unit			Sales revenue	Net profit	
Anita	180	0	0	176230	91010	<<== Max
Daphne	0	160	0			
Electra	0	0	140			
Generic backup	90	50	60			

Step 2a: Staging info

WBSP_VAR *Declare stage 0 decisions*

WBSP_RAND *Declare stage 1 random variables*

WBSP_VAR *Declare stage 1 decisions*

Step 2b: Distributions

WBSP_DIST_DISCRETE_SV_W *Declare discrete joint distribution*

	Demand scenarios			Probability
Scenarios	Anita	Daphne	Electra	Wgts
1	300	400	400	0.5
2	333	383	433	0.4
3	500	300	600	0.1

Step 3: Scenario/sampling info

WBSP_STSC

Stage	*Scenarios*
1	10

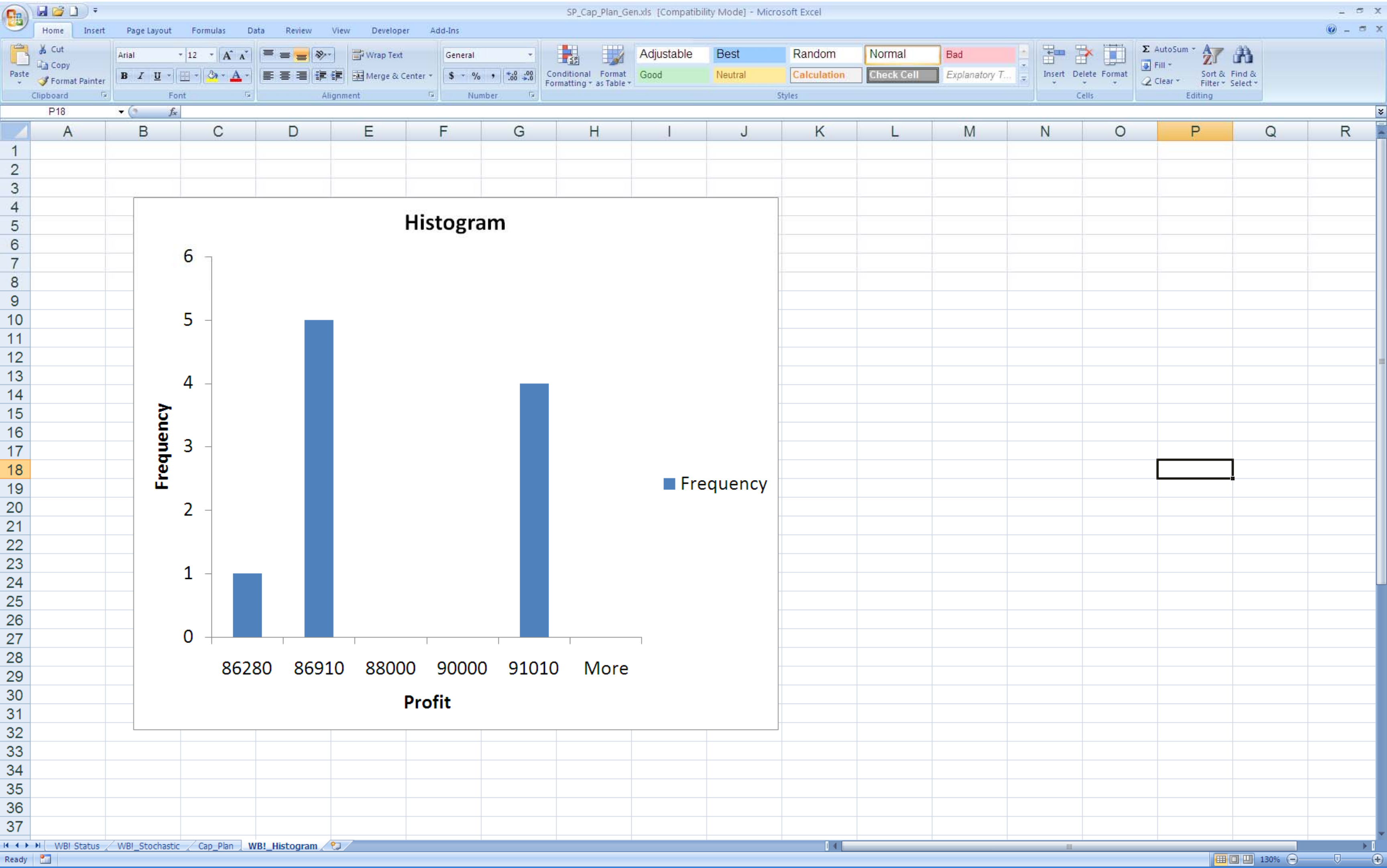
Step 4: Reporting info

WBSP_REP *Reporting cells (optional):*

WB! Status | WB!_Stochastic | Cap_Plan | WB!_Histogram

Ready 90%

Capacity Planning Under Uncertainty, Scenario Profit



Capacity Planning, Scenario by Scenario Report

SP_Cap_Plan_Gen.xls [Compatibility Mode] - Microsoft Excel

STOCHASTIC INFORMATION:

RANDOMS	3
STAGES	2
NODES	4
SCENARIOS	3

Expected Value (EV)	8.848700e+004
Expected Value of Wait-and-See (EVWS)	9.263400e+004
Expected Value using Expected Value Policy (EVEVP)	8.595240e+004
Expected Value of Perfect Information (=|EVWS-EV|)	4.147000e+003
Expected Value of Modeling Uncertainty (=|EV-EVEVP|)	2.534600e+003

REPORTING CELLS

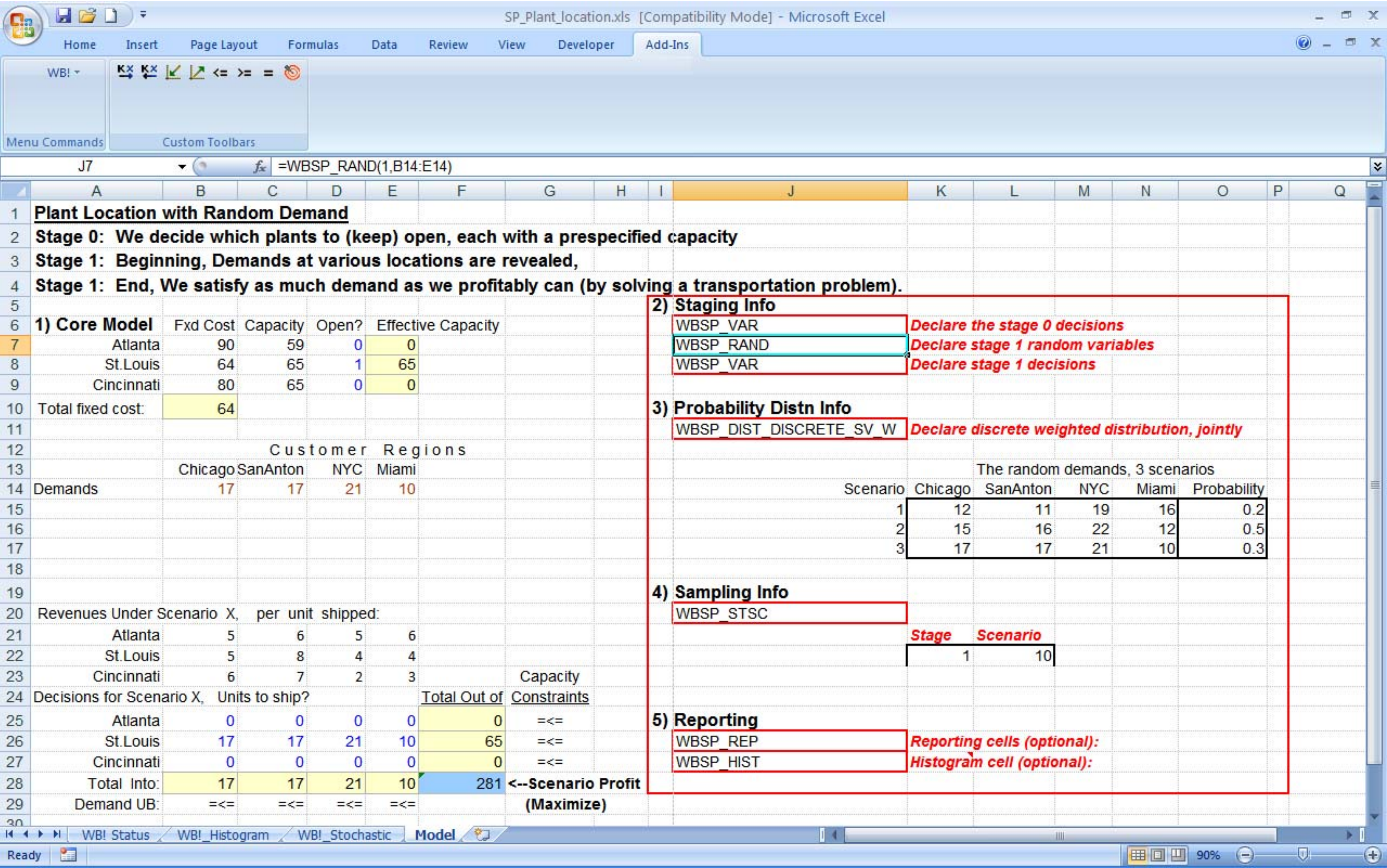
SCENARIO	PROBABILITY	Cap_Plan!C7	Cap_Plan!C8	Cap_Plan!C9	Cap_Plan!C10	Cap_Plan!B23	Cap_Plan!B20	Cap_Plan!C23	Cap_Plan!C20	Cap_Plan!D23	Cap_Plan!D20	Cap_Plan!G26
		CAPANITA	CAPDAPHNE	CAPELECTRA	CAPGENERIC	DEMANITA	GENANITA	DEMDAPHNE	GENDAPHNE	DEMELECTRA	GENELECTRA	WBMAX
		STAGE 0	STAGE 0	STAGE 0	STAGE 0	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1
- 1-	0.5	300	383	400	150	300	0	400	17	400	0	86910
- 2-	0.4	300	383	400	150	333	33	383	0	433	33	91010
- 3-	0.1	300	383	400	150	500	150	300	0	600	0	86280

End of Report

WB! Status | WB!_Stochastic | Cap_Plan

Ready

Plant Location with Random Demand



SP_Plant_location.xls [Compatibility Mode] - Microsoft Excel

J7 =WBSP_RAND(1,B14:E14)

Plant Location with Random Demand

Stage 0: We decide which plants to (keep) open, each with a prespecified capacity

Stage 1: Beginning, Demands at various locations are revealed,

Stage 1: End, We satisfy as much demand as we profitably can (by solving a transportation problem).

1) Core Model	Fxd Cost	Capacity	Open?	Effective Capacity
Atlanta	90	59	0	0
St.Louis	64	65	1	65
Cincinnati	80	65	0	0
Total fixed cost:	64			

Customer Regions

	Chicago	SanAnton	NYC	Miami
Demands	17	17	21	10

Revenues Under Scenario X, per unit shipped:

Atlanta	5	6	5	6
St.Louis	5	8	4	4
Cincinnati	6	7	2	3

Decisions for Scenario X, Units to ship?					Total Out of	Capacity Constraints
Atlanta	0	0	0	0	0	=<=
St.Louis	17	17	21	10	65	=<=
Cincinnati	0	0	0	0	0	=<=
Total Into:	17	17	21	10	281	**<--Scenario Profit**
Demand UB:	=<=	=<=	=<=	=<=		**(Maximize)**

2) Staging Info

WBSP_VAR	*Declare the stage 0 decisions*
WBSP_RAND	*Declare stage 1 random variables*
WBSP_VAR	*Declare stage 1 decisions*

3) Probability Distn Info

WBSP_DIST_DISCRETE_SV_W — *Declare discrete weighted distribution, jointly*

The random demands, 3 scenarios

Scenario	Chicago	SanAnton	NYC	Miami	Probability
1	12	11	19	16	0.2
2	15	16	22	12	0.5
3	17	17	21	10	0.3

4) Sampling Info

WBSP_STSC

Stage	*Scenario*
1	10

5) Reporting

WBSP_REP	*Reporting cells (optional):*
WBSP_HIST	*Histogram cell (optional):*

WB! Status | WB!_Histogram | WB!_Stochastic | **Model**

Plant Location with Random Demand, Output

The output tab,
WB!_Stochastic, contains two types of information:

1) Various expected values that measure the cost of uncertainty,
2) A scenario by scenario listing of selected variables so we can explicitly verify what happens in each possible scenario.

We may optionally also generate histograms in a WB!_Histogram tab.

Later, we will discuss the various expected values and the various costs of uncertainty.

Plant Location, Scenario Report

What'sBest!® 11.0.3.0 (Nov 03, 2011) - Library 7.0.1.390 - 64-bit - Stochastic Report -

Expected Value (EV) 2.6660E+02

Expected Value of Wait-and-See Model's Objective (EVWS) 2.6720E+02

Expected Value of Perfect Information (=|EVWS-EV|) 6.0000E-01

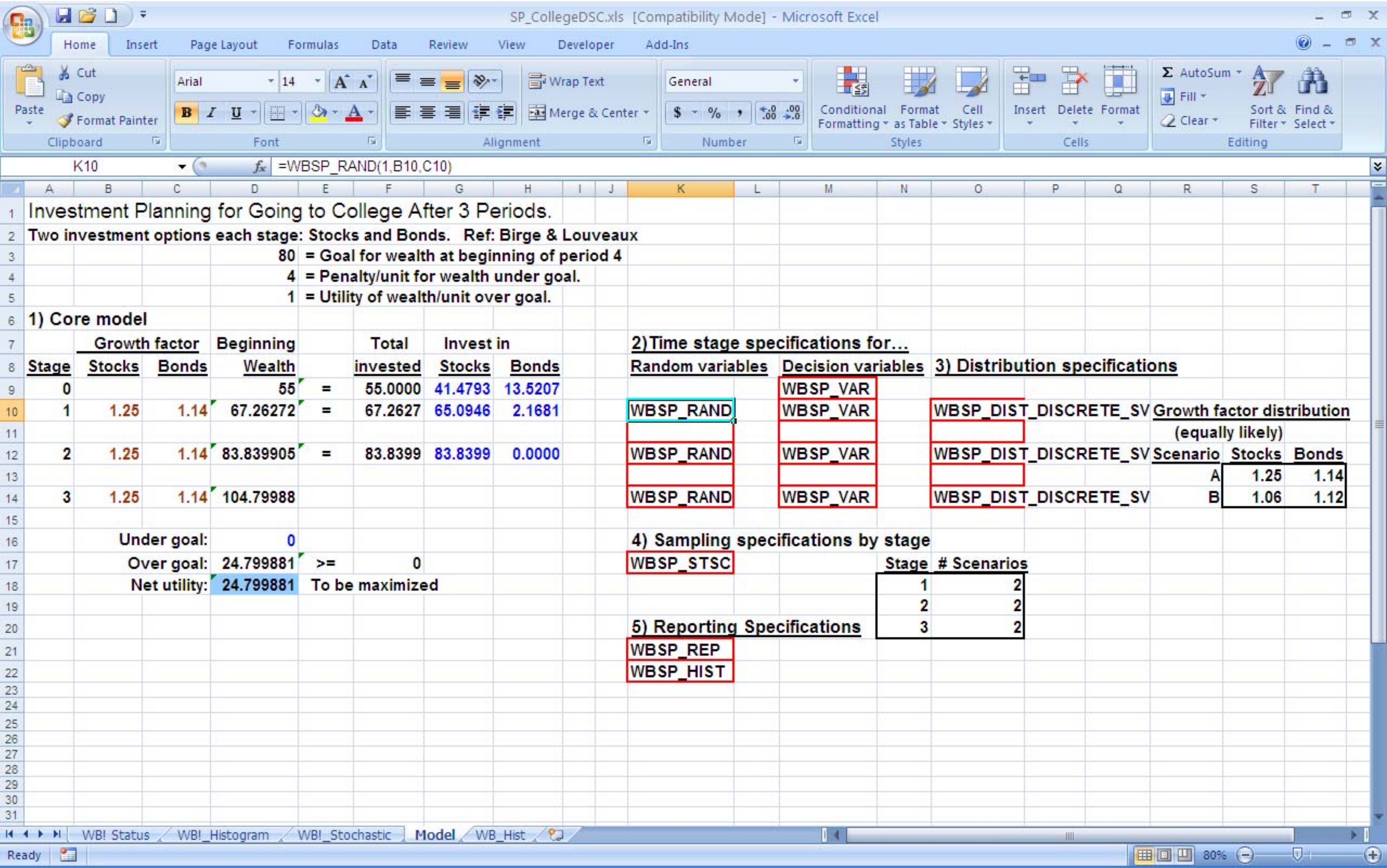
Expected Value of Modeling Uncertainty (=|EV-EVEM|) 0.0000E+00

REPORTING CELLS

SCENARIO	PROBABILITY	Model!B14 CHICAGO STAGE 1	Model!C14 SANANTON STAGE 1	Model!D14 NYC STAGE 1	Model!E14 MIAMI STAGE 1	Model!B26 STL_CHI STAGE 1	Model!C26 STL_SAN STAGE 1	Model!D26 STL_NYC STAGE 1	Model!E26 STL_MIA STAGE 1	Model!F28 TOTAL_PROFIT STAGE 1
– 1–	0.1	17	17	21	10	17	17	21	10	281
– 2–	0.1	15	16	22	12	15	16	22	12	275
– 3–	0.1	15	16	22	12	15	16	22	12	275
– 4–	0.1	15	16	22	12	15	16	22	12	275
– 5–	0.1	12	11	19	16	12	11	19	16	224
– 6–	0.1	17	17	21	10	17	17	21	10	281
– 7–	0.1	12	11	19	16	12	11	19	16	224
– 8–	0.1	15	16	22	12	15	16	22	12	275
– 9–	0.1	15	16	22	12	15	16	22	12	275
– 10–	0.1	17	17	21	10	17	17	21	10	281

End of Report

Multi-Stage Portfolio Model with Downside Risk



Multi-Stage Portfolio Model with Downside Risk

SPCollegeDSC.xls [Compatibility Mode] - Microsoft Excel

Home | Insert | Page Layout | Formulas | Data | Review | View | Developer | Add-Ins

Gas Interruption ▾ · WB! ▾ — Menu Commands · Custom Toolbars

A1 — f_x Investment Planning for Going to College After 3 Periods.

Investment Planning for Going to College After 3 Periods.

Two investment options each stage: Stocks and Bonds. Ref: Birge & Louveau

80 = Goal for wealth at beginning of period 4

4 = Penalty/unit for wealth under goal.

1 = Utility of wealth/unit over goal.

1) Core model

	Growth factor		Beginning		Total	Invest in	
Stage	Stocks	Bonds	Wealth		invested	Stocks	Bonds
0			55	=	55.0000	41.4793	13.5207
1	1.06	1.12	59.11124	=	59.1112	36.7432	22.3680
2	1.25	1.14	71.42857	=	71.4286	0.0000	71.4286
3	1.06	1.12	80				

Under goal:	0		
Over goal:	0	=>=	0
Net utility:	0	To be maximized	

2)Time stage specifications for…

Random variables	Decision variables	3) Distribution specifications
	WBSP_VAR	
WBSP_RAND	WBSP_VAR	WBSP_DIST_DISCRETE
WBSP_RAND	WBSP_VAR	WBSP_DIST_DISCRETE
WBSP_RAND	WBSP_VAR	WBSP_DIST_DISCRETE

Growth factor distribution

(equally likely)

Scenario	Stocks	Bonds
A	1.25	1.14
B	1.06	1.12

4) Sampling specifications by stage

WBSP_STSC

Stage	# Scenarios
1	2
2	2
3	2

5) Reporting Specifications

WBSP_REP

Stage Assumptions

1) Default stage assignment for a formula or constraint
= highest stage of any variable in the constraint
or RHS of the formula, except,

2) Objective function is automatically assigned to stage 0.

WB! Status | WB! Stochastic | Core model | Report

Ready — 100%

Multi-stage Portfolio: Solution and Policy

SP_CollegeDSC.xls [Compatibility Mode] - Microsoft Excel

H19 = Model!D12

```
-------------------------------------------------------
RANDOMS                                   6
STAGES                                    4
NODES                                    15
SCENARIOS                                 8

EXPECTED VALUE                                          -1.514085e+000
Expected Value of Perfect Information Lower Bound        1.201109e+001

     REPORTING CELLS
SCENARIO
```

		Model!F9 WEALTH0 STAGE 0	Model!G9 STOCKINVEST0 STAGE 0	Model!H9 BONDINVEST0 STAGE 0	Model!D10 WEALTH1 STAGE 1	Model!G10 STOCKINVEST1 STAGE 1	Model!H10 BONDINVEST1 STAGE 1	**Model!D12 WEALTH2 STAGE 2**	**Model!G12 STOCKINVEST2 STAGE 2**	**Model!H12 BONDINVEST2 STAGE 2**	Model!D14 WEALTH3 STAGE 3	Model!D16 UNDERGOAL STAGE 3	Model!D1 OVERGOA STAGE
-	1-	55	41.479272	13.520728	59.111244	36.743215	22.368029	**71.428571**	**0**	**71.428571**	80	0	
-	2-	55	41.479272	13.520728	59.111244	36.743215	22.368029	**71.428571**	**0**	**71.428571**	81.428571	0	1.42857
-	3-	55	41.479272	13.520728	59.111244	36.743215	22.368029	**64**	**64**	**0**	67.84	12.16	
-	4-	55	41.479272	13.520728	59.111244	36.743215	22.368029	**64**	**64**	**0**	80	0	
-	5-	55	41.479272	13.520728	67.26272	65.094582	2.168138	**83.839905**	**83.839905**	**0**	88.870299	0	8.87029
-	6-	55	41.479272	13.520728	67.26272	65.094582	2.168138	**83.839905**	**83.839905**	**0**	104.799881	0	24.79988
-	7-	55	41.479272	13.520728	67.26272	65.094582	2.168138	**71.428571**	**0**	**71.428571**	80	0	
-	8-	55	41.479272	13.520728	67.26272	65.094582	2.168138	**71.428571**	**0**	**71.428571**	81.428571	0	1.42857

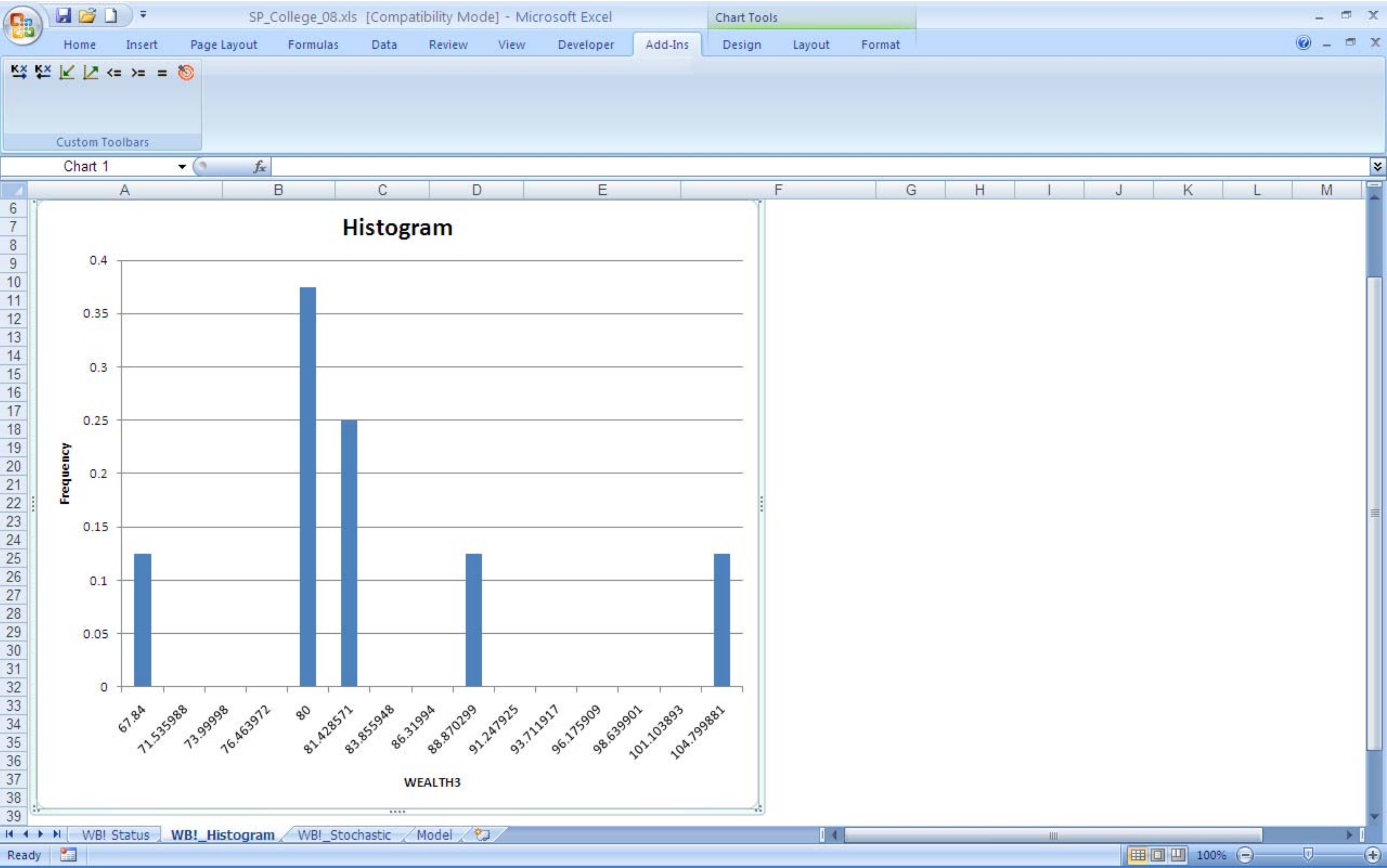
End of Report

Notice when we put all our money in stocks in stage 2….

WB! Status | WB!_Stochastic | Model | WB_Hist

Ready   Average: 48.44950783   Count: 33   Sum: 1162.788188   120%

Terminal Wealth Distribution: College/Retirement Planning



Yield Management: Bird in Hand vs. Future Bird in Bush

SP_Seat_Mana.xls [Compatibility Mode] - Microsoft Excel

Q12 =WBSP_RAND(4,B12)

Yield Management Decision:

Each period we decide how many of the remaining seats to sell, and how many to hold back to sell in the future for a higher price, if enough demand occurs.

1) Core model:

Seats(t) = Seats(t-1) - Sales(t); Sales(t) <= Demand(t);

Period	(Random) Demand		(Decision) Sales	Seats			Price per seat	Revenue
0				20				
1	10	=>=	10	10	>=	0	12	120
2	7	>=	5	5	>=	0	15	75
3	7	>=	1	4	>=	0	20	20
4	10	>=	1	3	>=	0	25	25

Max expected revenue: 240

Stochastic extension

2) Stage information

Specify stage of Accept decisions	Mark seat demands as random variables and give stage	3) Distribution information
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE

WBSP_VAR

Distribution of Possible demands

2
7
10

4) Sample size

WBSP_STSC

Stage	Scenario
1	3
2	3
3	3
4	3

5) Reporting Cells

WBSP_REP

WB! Status | WB! Stochastic | Core_Model

Yield Management: Report and Policy

SP_Seat_Man.xls [Compatibility Mode] - Microsoft Excel

STOCHASTIC INFORMATION:

RANDOMS	4
STAGES	5
NODES	23
SCENARIOS	8
EXPECTED VALUE	2.440000e+002

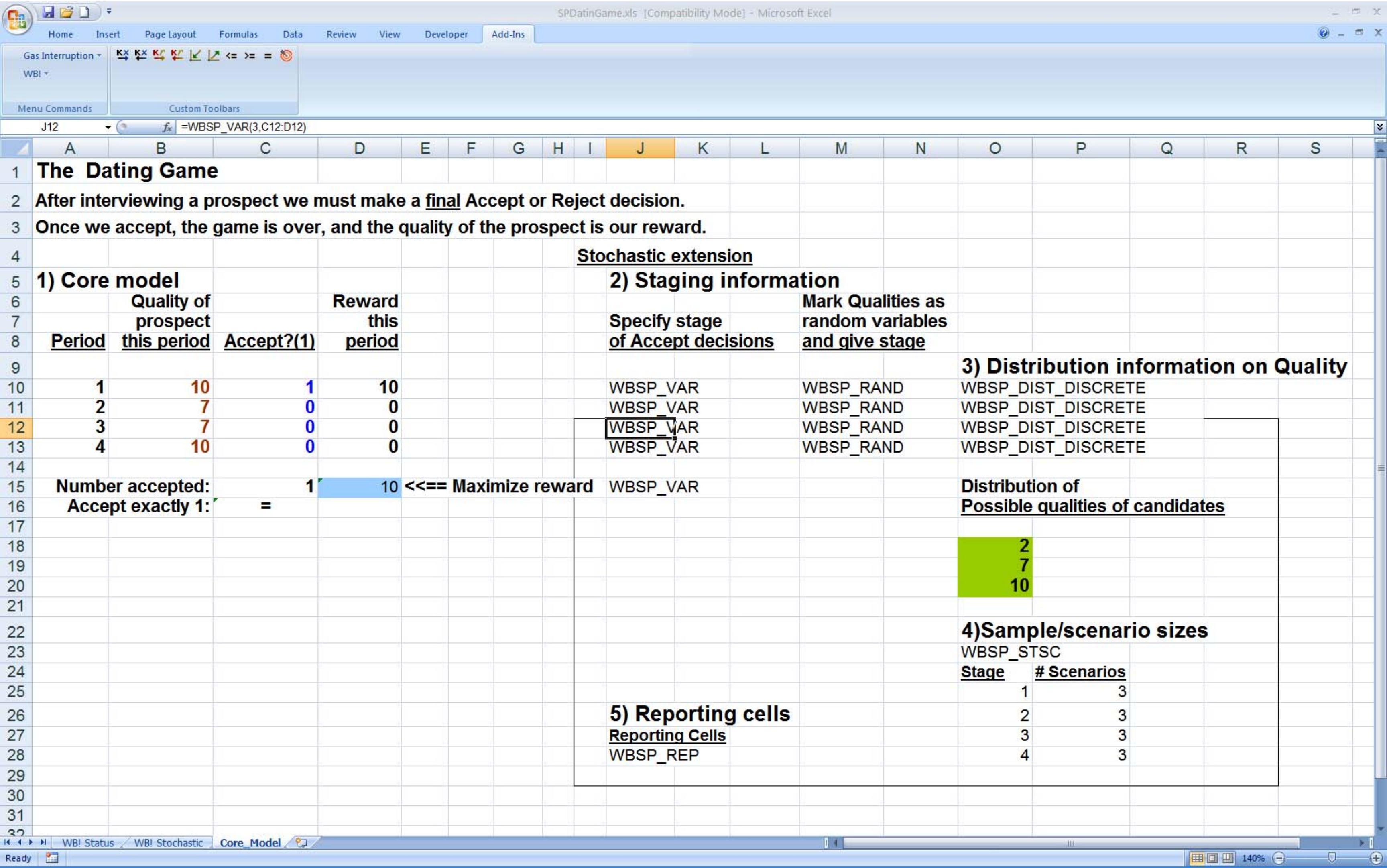
REPORTING CELLS

SCENARIO

	B9	D9	E9	B10	D10	E10	B11	D11	E11	B12	D12	E12
	DEMAND1	SALES1	SEATS1	DEMAND2	SALES2	SEATS2	DEMAND3	SALES3	SEATS3	DEMAND4	SALES4	SEATS4
	STAGE 1	STAGE 0	STAGE 0	STAGE 2	STAGE 0	STAGE 0	STAGE 3	STAGE 0	STAGE 0	STAGE 4	STAGE 0	STAGE 0
- 1-	7	2	18	2	2	16	10	7	9	2	2	7
- 2-	7	2	18	2	2	16	7	7	9	2	2	7
- 3-	7	2	18	10	2	16	10	7	9	2	2	7
- 4-	7	2	18	10	2	16	7	7	9	2	2	7
- 5-	2	2	18	2	2	16	10	7	9	2	2	7
- 6-	2	2	18	2	2	16	7	7	9	2	2	7
- 7-	2	2	18	10	2	16	10	7	9	2	2	7
- 8-	2	2	18	10	2	16	7	7	9	2	2	7

WB! Status | WB! Stochastic | Core_Model

Stopping Problem Example



Stopping Problem Solution and Policy

SPDatinGame.xls [Compatibility Mode] - Microsoft Excel

H31 = 10

REPORTING CELLS

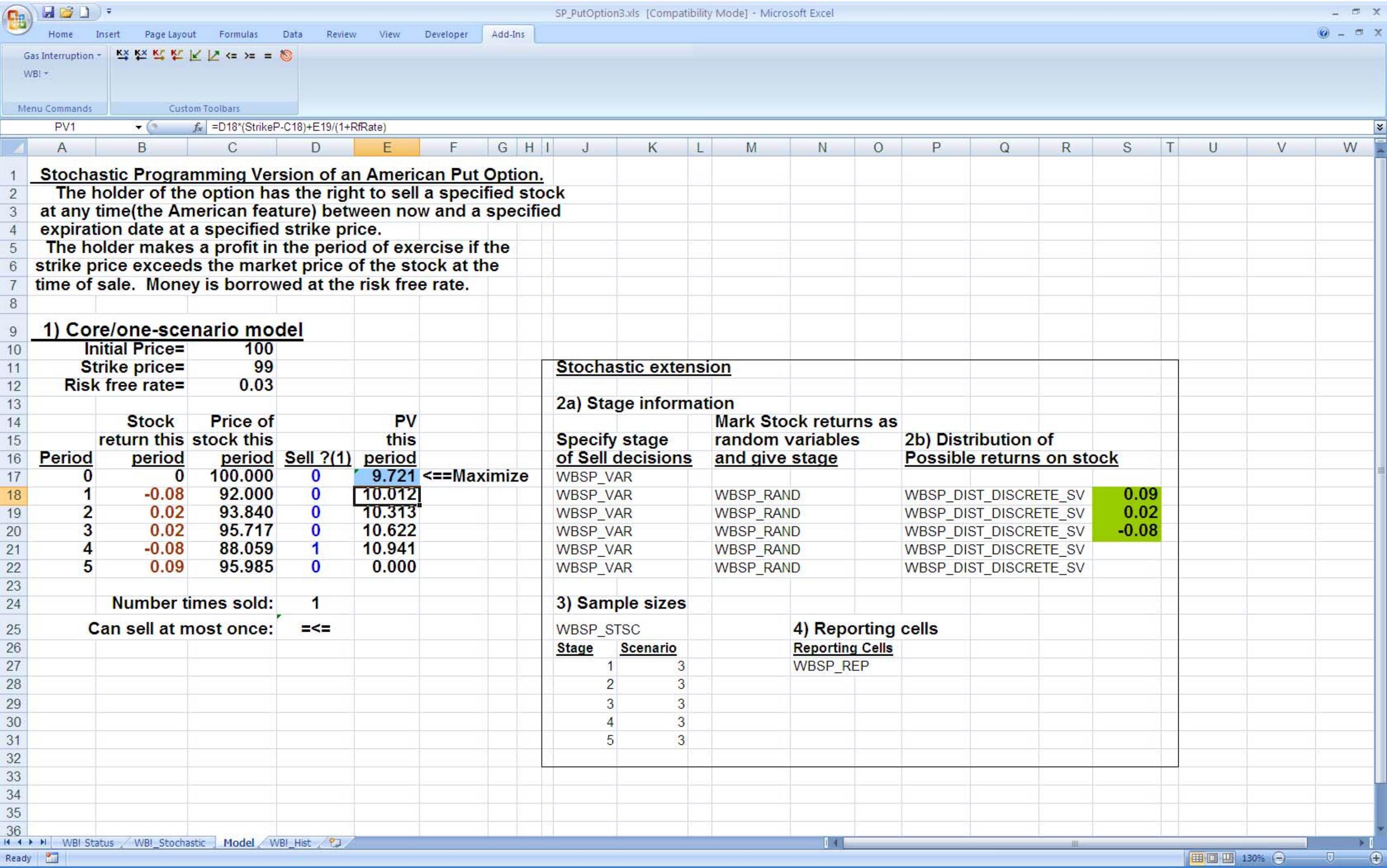
SCENARIO

	B10	C10	B11	C11	B12	C12	B13	C13
	QUALITY1	WBBINACCEPT	QUALITY2	WBBINACCEPT	QUALITY3	WBBINACCEPT	QUALITY4	WBBINACCEPT
	STAGE 1	STAGE 1	STAGE 2	STAGE 2	STAGE 3	STAGE 3	STAGE 4	STAGE 4
– 1–	**10**	**1**	7	0	7	0	10	0
– 28–	**2**	**0**	7	0	7	1	10	0
– 37–	2	0	**10**	**1**	7	0	10	0
– 46–	2	0	2	0	7	1	10	0
– 47–	2	0	2	0	7	1	7	0
– 48–	2	0	2	0	**7**	**1**	2	0
– 49–	2	0	2	0	**2**	**0**	**10**	**1**
– 50–	2	0	2	0	2	0	**7**	**1**
– 51–	2	0	2	0	2	0	**2**	**1**
– 52–	2	0	2	0	**10**	**1**	10	0
– 53–	2	0	2	0	10	1	7	0
– 54–	2	0	2	0	10	1	2	0
– 55–	**7**	**0**	**7**	**0**	7	1	10	0
– 56–	7	0	**7**	**0**	7	1	7	0
– 57–	7	0	7	0	7	1	2	0
– 58–	7	0	7	0	2	0	10	1
– 59–	7	0	7	0	2	0	7	1
– 60–	7	0	7	0	2	0	2	1
– 61–	7	0	7	0	10	1	10	0
– 62–	7	0	7	0	10	1	7	0
– 63–	7	0	7	0	10	1	2	0

WB! Status | WB! Stochastic | Core_Model

FALSE — Average: 3.666666667 Count: 6 Sum: 22 — 110%

Put-Option Formulated as an SP



Stochastic Programming Version of an American Put Option.

The holder of the option has the right to sell a specified stock at any time(the American feature) between now and a specified expiration date at a specified strike price.

The holder makes a profit in the period of exercise if the strike price exceeds the market price of the stock at the time of sale. Money is borrowed at the risk free rate.

1) Core/one-scenario model

Initial Price=	100
Strike price=	99
Risk free rate=	0.03

Period	Stock return this period	Price of stock this period	Sell ?(1)	PV this period	
0	0	100.000	0	9.721	<==Maximize
1	-0.08	92.000	0	10.012	
2	0.02	93.840	0	10.313	
3	0.02	95.717	0	10.622	
4	-0.08	88.059	1	10.941	
5	0.09	95.985	0	0.000	

Number times sold: 1

Can sell at most once: =<=

Stochastic extension

2a) Stage information

Specify stage of Sell decisions	Mark Stock returns as random variables and give stage	2b) Distribution of Possible returns on stock	
WBSP_VAR			
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE_SV	0.09
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE_SV	0.02
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE_SV	-0.08
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE_SV	
WBSP_VAR	WBSP_RAND	WBSP_DIST_DISCRETE_SV	

3) Sample sizes

WBSP_STSC

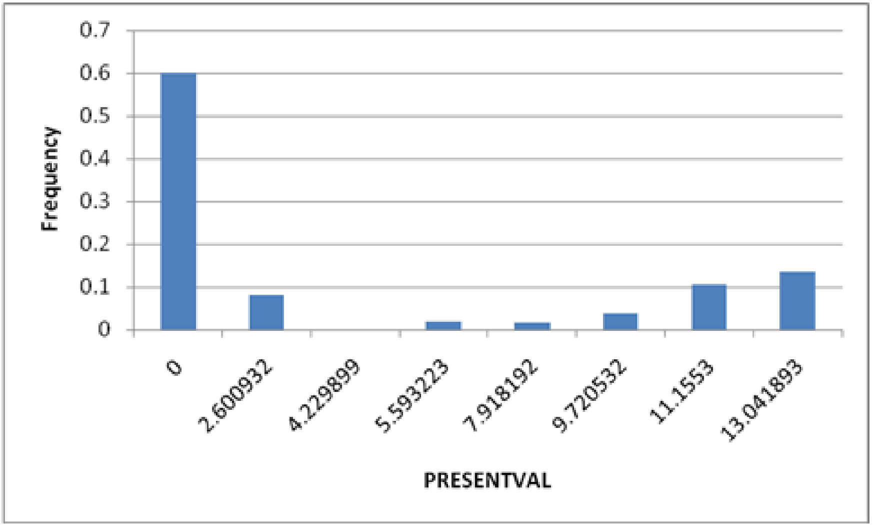
Stage	Scenario
1	3
2	3
3	3
4	3
5	3

4) Reporting cells

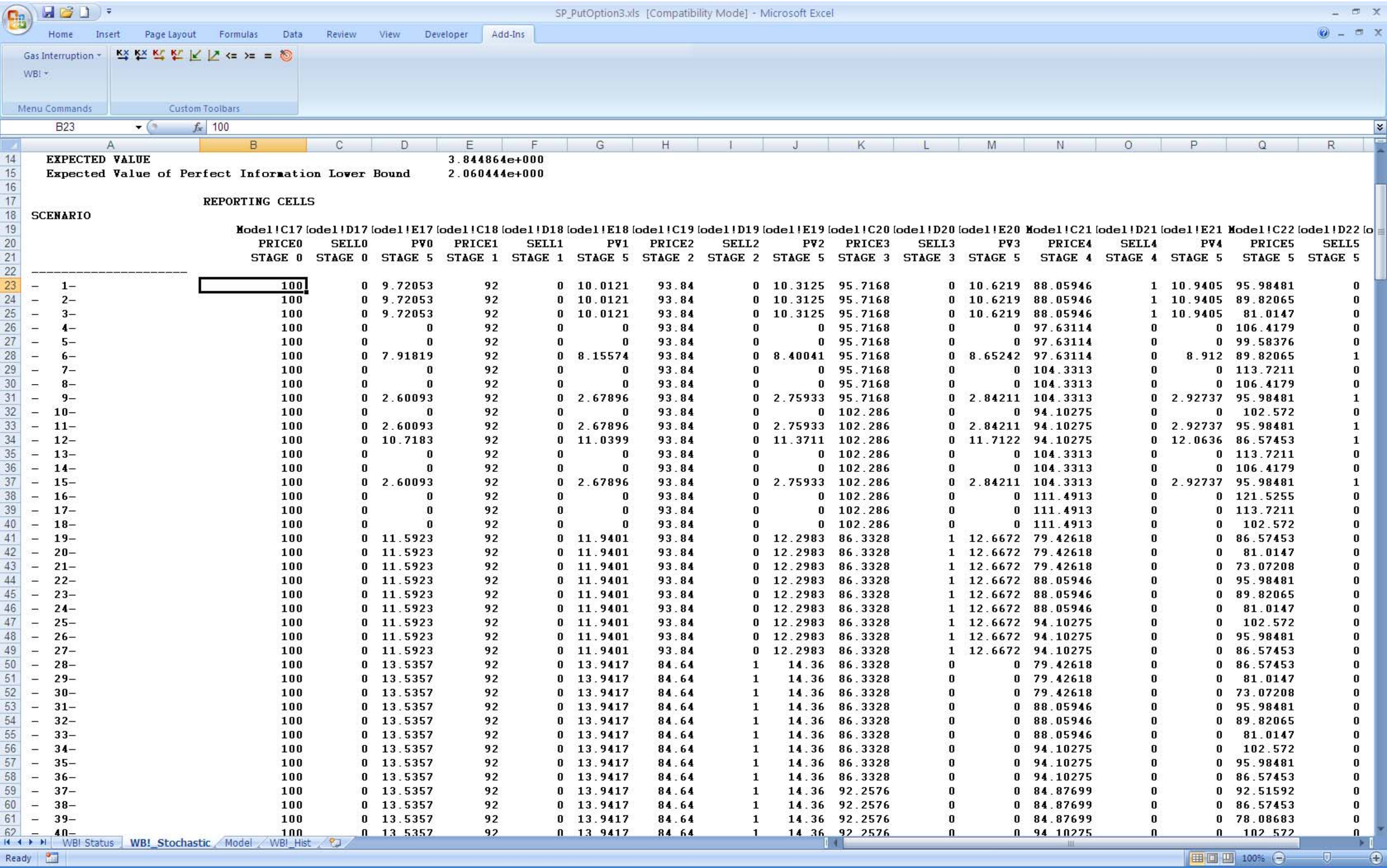
Reporting Cells

WBSP_REP

Put-Option, 60% of Time Does Not Pay Off



Put Option, Scenario Detail



SP_PutOption3.xls [Compatibility Mode] - Microsoft Excel

EXPECTED VALUE 3.844864e+000

Expected Value of Perfect Information Lower Bound 2.060444e+000

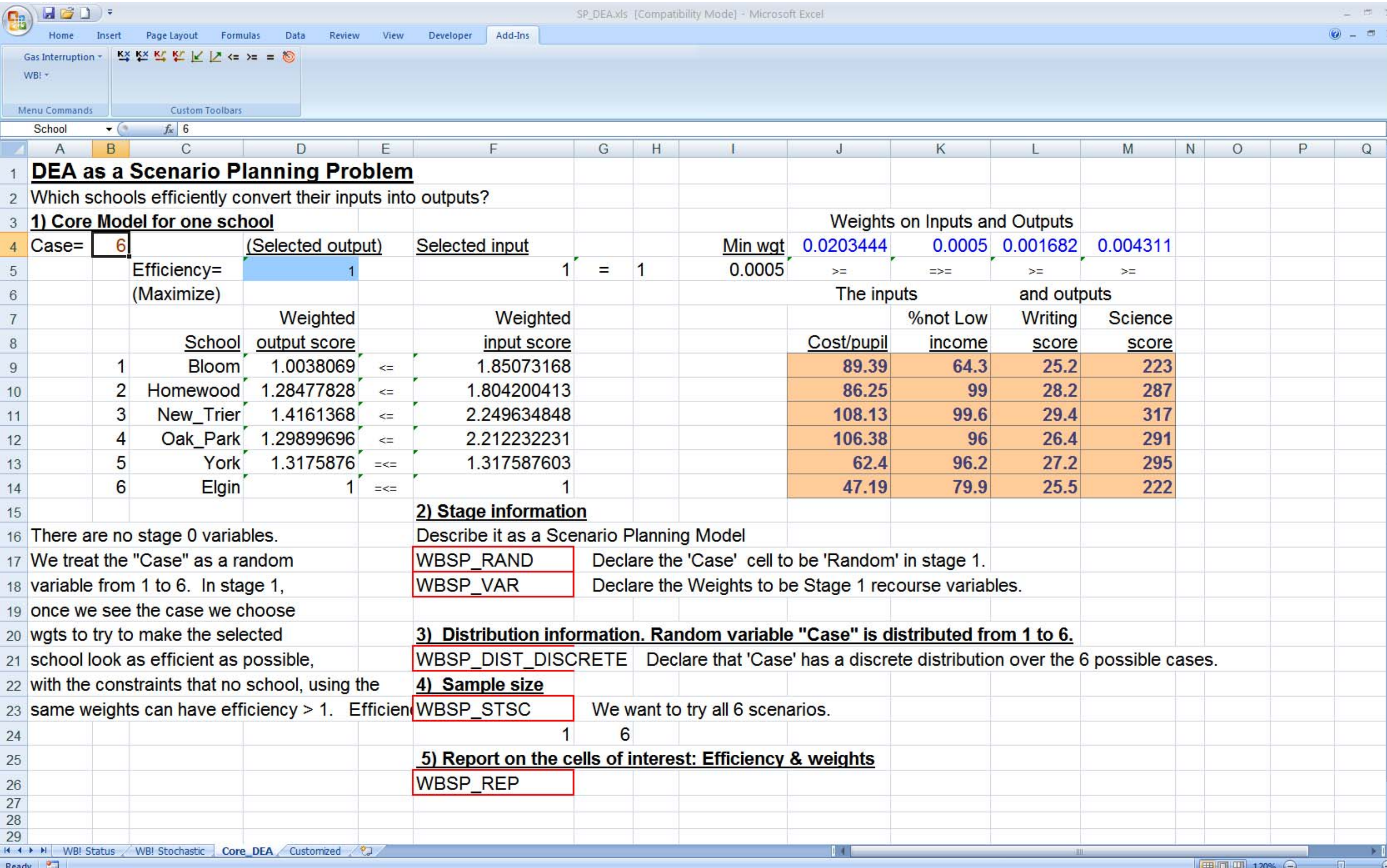
REPORTING CELLS

SCENARIO

	Model!C17	[odel!D17	[odel!E17	[odel!C18	[odel!D18	[odel!E18	[odel!C19	[odel!D19	[odel!E19	[odel!C20	[odel!D20	[odel!E20	Model!C21	[odel!D21	[odel!E21	Model!C22	[odel!D22
	PRICE0 STAGE 0	SELL0 STAGE 0	PV0 STAGE 5	PRICE1 STAGE 1	SELL1 STAGE 1	PV1 STAGE 5	PRICE2 STAGE 2	SELL2 STAGE 2	PV2 STAGE 5	PRICE3 STAGE 3	SELL3 STAGE 3	PV3 STAGE 5	PRICE4 STAGE 4	SELL4 STAGE 4	PV4 STAGE 5	PRICE5 STAGE 5	SELL5 STAGE 5
– 1–	100	0	9.72053	92	0	10.0121	93.84	0	10.3125	95.7168	0	10.6219	88.05946	1	10.9405	95.98481	0
– 2–	100	0	9.72053	92	0	10.0121	93.84	0	10.3125	95.7168	0	10.6219	88.05946	1	10.9405	89.82065	0
– 3–	100	0	9.72053	92	0	10.0121	93.84	0	10.3125	95.7168	0	10.6219	88.05946	1	10.9405	81.0147	0
– 4–	100	0	0	92	0	0	93.84	0	0	95.7168	0	0	97.63114	0	0	106.4179	0
– 5–	100	0	0	92	0	0	93.84	0	0	95.7168	0	0	97.63114	0	0	99.58376	0
– 6–	100	0	7.91819	92	0	8.15574	93.84	0	8.40041	95.7168	0	8.65242	97.63114	0	8.912	89.82065	1
– 7–	100	0	0	92	0	0	93.84	0	0	95.7168	0	0	104.3313	0	0	113.7211	0
– 8–	100	0	0	92	0	0	93.84	0	0	95.7168	0	0	104.3313	0	0	106.4179	0
– 9–	100	0	2.60093	92	0	2.67896	93.84	0	2.75933	95.7168	0	2.84211	104.3313	0	2.92737	95.98481	1
– 10–	100	0	0	92	0	0	93.84	0	0	102.286	0	0	94.10275	0	0	102.572	0
– 11–	100	0	2.60093	92	0	2.67896	93.84	0	2.75933	102.286	0	2.84211	94.10275	0	2.92737	95.98481	1
– 12–	100	0	10.7183	92	0	11.0399	93.84	0	11.3711	102.286	0	11.7122	94.10275	0	12.0636	86.57453	1
– 13–	100	0	0	92	0	0	93.84	0	0	102.286	0	0	104.3313	0	0	113.7211	0
– 14–	100	0	0	92	0	0	93.84	0	0	102.286	0	0	104.3313	0	0	106.4179	0
– 15–	100	0	2.60093	92	0	2.67896	93.84	0	2.75933	102.286	0	2.84211	104.3313	0	2.92737	95.98481	1
– 16–	100	0	0	92	0	0	93.84	0	0	102.286	0	0	111.4913	0	0	121.5255	0
– 17–	100	0	0	92	0	0	93.84	0	0	102.286	0	0	111.4913	0	0	113.7211	0
– 18–	100	0	0	92	0	0	93.84	0	0	102.286	0	0	111.4913	0	0	102.572	0
– 19–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	79.42618	0	0	86.57453	0
– 20–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	79.42618	0	0	81.0147	0
– 21–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	79.42618	0	0	73.07208	0
– 22–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	88.05946	0	0	95.98481	0
– 23–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	88.05946	0	0	89.82065	0
– 24–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	88.05946	0	0	81.0147	0
– 25–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	94.10275	0	0	102.572	0
– 26–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	94.10275	0	0	95.98481	0
– 27–	100	0	11.5923	92	0	11.9401	93.84	0	12.2983	86.3328	1	12.6672	94.10275	0	0	86.57453	0
– 28–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	79.42618	0	0	86.57453	0
– 29–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	79.42618	0	0	81.0147	0
– 30–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	79.42618	0	0	73.07208	0
– 31–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	88.05946	0	0	95.98481	0
– 32–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	88.05946	0	0	89.82065	0
– 33–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	88.05946	0	0	81.0147	0
– 34–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	94.10275	0	0	102.572	0
– 35–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	94.10275	0	0	95.98481	0
– 36–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	86.3328	0	0	94.10275	0	0	86.57453	0
– 37–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	92.2576	0	0	84.87699	0	0	92.51592	0
– 38–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	92.2576	0	0	84.87699	0	0	86.57453	0
– 39–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	92.2576	0	0	84.87699	0	0	78.08683	0
– 40–	100	0	13.5357	92	0	13.9417	84.64	1	14.36	92.2576	0	0	94.10275	0	0	102.572	0

WB! Status | WB!_Stochastic | Model | WB!_Hist

DEA: An SP Application with No Randomness



SP_DEA.xls [Compatibility Mode] - Microsoft Excel

DEA as a Scenario Planning Problem

Which schools efficiently convert their inputs into outputs?

1) Core Model for one school

Case=	6		(Selected output)		Selected input				Weights on Inputs and Outputs		
								Min wgt	0.0203444	0.0005	0.001682	0.004311
		Efficiency=	1		1	=	1	0.0005	>=	=>=	>=	>=
		(Maximize)							The inputs		and outputs	
			Weighted		Weighted					%not Low	Writing	Science
		School	output score		input score				Cost/pupil	income	score	score
	1	Bloom	1.0038069	<=	1.85073168				89.39	64.3	25.2	223
	2	Homewood	1.28477828	<=	1.804200413				86.25	99	28.2	287
	3	New_Trier	1.4161368	<=	2.249634848				108.13	99.6	29.4	317
	4	Oak_Park	1.29899696	<=	2.212232231				106.38	96	26.4	291
	5	York	1.3175876	=<=	1.317587603				62.4	96.2	27.2	295
	6	Elgin	1	=<=	1				47.19	79.9	25.5	222

2) Stage information

There are no stage 0 variables. We treat the "Case" as a random variable from 1 to 6. In stage 1, once we see the case we choose wgts to try to make the selected school look as efficient as possible, with the constraints that no school, using the same weights can have efficiency > 1. Efficien

Describe it as a Scenario Planning Model

WBSP_RAND Declare the 'Case' cell to be 'Random' in stage 1.

WBSP_VAR Declare the Weights to be Stage 1 recourse variables.

3) Distribution information. Random variable "Case" is distributed from 1 to 6.

WBSP_DIST_DISCRETE Declare that 'Case' has a discrete distribution over the 6 possible cases.

4) Sample size

WBSP_STSC We want to try all 6 scenarios.

1 6

5) Report on the cells of interest: Efficiency & weights

WBSP_REP

WB! Status | WB! Stochastic | Core_DEA | Customized

Report: DEA Efficiency

SP_DEA.xls [Compatibility Mode] - Microsoft Excel

C4 =‘WB! Stochastic’!C22

Efficiency Report

School	Efficiency	WgtCost	WgtIncome	WgtWriting	WgtScience
Elgin	1	0.020344	0.0005	0.001682	0.004311
Oak_Park	0.912128	0.001747	0.008481	0.0005	0.003089
New_Trier	0.96158	0.001691	0.008204	0.0005	0.002987
York	1	0.001884	0.009173	0.0005	0.003344
Homewood	0.909507	0.001763	0.008565	0.0005	0.00312
Bloom	1	0.0005	0.014857	0.0005	0.004428

WB! Status | WB! Stochastic | Core_DEA | Customized

Random Number Generation and Sampling

Ideas and Steps:

Uniform Random Number Generation

Arbitrary Distribution from Uniform

Variance Reduction, Quasi-random Numbers, Super Uniforms
Latin Hypercube Sampling, Antithetic Variates.

Correlated Random Numbers

Uniform Random Number Generators

LINDO API and What'sBest 10 provide:

1) Linear congruential, 31 bit,

2) Composite of linear congruentials with a long period,(default)

3) Mersenne Twister with long period.

Simple Linear Congruential, 31 bit Uniform Generator

```
IX = 742938285 * IX MOD 2147483647
 LSrand = IX/2147483647.0
```

The starting seed for the random number generator, regardless of which generator is used, can be selected by clicking on:

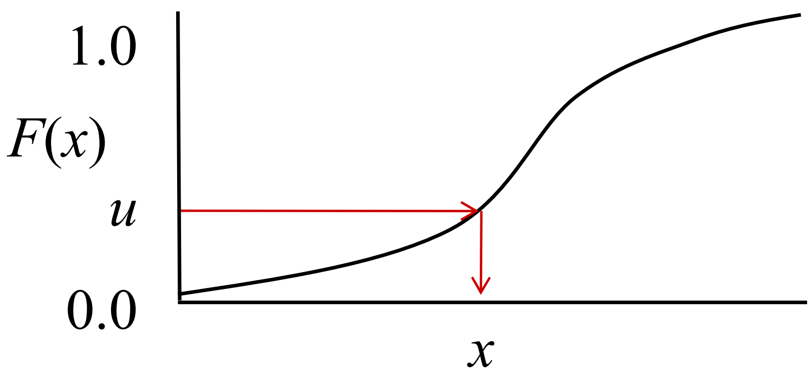
Add-Ins | WB! | Options | Stochastic Solver | Seed for Random Number Generator

Random Numbers from Arbitrary Distributions

Generating a random number from an arbitrary distribution, e.g., Normal, Poisson, Negative binomial…

1) Generate a uniform random number in (0, 1).

2) Convert the uniform to the desired distribution via the inverse transform of the cdf(cumulative distribution function.



Need to be able to invert

$u = F(x)$ *to*

$x = F^{-1}(u)$.

There are lots of methods for generating r.v.'s from a given distribution.

Why use the inverse transform method?

Additional Distributional Details

Distributions supported:

DISCRETE, DISCRETE_W (Emprical Multi-variate)

BETA	LOGARITHMIC
BINOMIAL	LOGISTIC
CAUCHY	LOGNORMAL
CHISQUARE	NEGATIVEBINOMIAL
EXPONENTIAL	NORMAL
F_DISTRIBUTION	PARETO
GAMMA	POISSON
GEOMETRIC	STUDENTS_T
GUMBEL	TRIANGULAR
HYPERGEOMETRIC	UNIFORM
LAPLACE	WEIBULL

Correlations supported:

Pearson, Spearman, Kendall

Sampling: Latin Hypercube

If we need more than one observation from a univariate distribution, use Latin Hypercube sampling.

Basic idea: If taking a sample of size N, choose one draw randomly from each Nth percentile.
This is easy to do if Inverse Transform Method is used.

Key feature: A given possible outcome has a probability of being chosen equal to its population probability.
So the sample is an unbiased sample.

Latin Hypercube Sampling

SPNewsBoyU.xls [Compatibility Mode] - Microsoft Excel

Stochastic/Scenario Optimization of Newsvendor Problem in What's*Best* *(Linear version)*

Given all costs and prices, in
Stage 0 we must decide how many newspapers to stock. In
Stage 1, in the beginning, unknown demand is revealed to us, and finally in
Stage 1, at the end, we compute our sales and the resulting profit.

1) Core model:

Row	A	B	C	I	J	K	L	M
7	CP = Purchase cost/unit=	30						
8	H=Holding cost/(unit leftover)=	10			***Add stochastic data here***			
9	P=Shortage cost/(unit unsatisfied demand)=	5		**2) Stage information**				
10	V=revenue per unit sold=	60			WBSP_VAR		(S is a stage 0 decision)	
11	S=Stock level(stage 1 decision)=	4.346	<<== Stage 0 decision.				**3) Distribution information**	
12	D=Demand(stage 2 random variable)=	8.9722	<<== Stage 1 random demand.		WBSP_RAND		WBSP_DIST_UNIFORM	
13	LS= Lost sales=	4.6262	<<== Stage 1 (recourse) decision.				0	Lower Limit
14	[LSGE] LS >= D - S (constraint)	=>=	<<== Stage 1 constraint.		WBSP_VAR		10	Upper Limit
15	[IDEF] I=Inventory=S-D+LS=	0	<<== Stage 1 decision and constraint.					
16	[IGE0] I >= 0 (constraint)	=>=	<<== Stage 1 non-negativity constraint.	**4) Sample size**	***Stage***	***Scenarios***		
17	[TCDEF] TC = Total cost of goods = CP * S =	130.38	<<== Stage 0 cost computation.		WBSP_STSC			
18	[THDEF] TH = Total Holding cost=H*I =	0	<<== Stage 1 holding cost computation.		1	10		
19	[TSDEF] TS = Total Shortage cost= P*LS=	23.131	<<== Stage 1 shortage cost computation.					
20	[VIDEF] VI = Revenue = V*(D-LS)=	260.76	<<== Stage 1 revenue computation.					
21	Profit, expected value, [To be maximized] =			**5) Reporting cells**				
22	[TPDEF] TP = VI - TC - TH - TS =	107.249	<<== Stage 1 expected value (maximize)		WBSP_REP			

Overview:

The user enters only a generic scenario 1.

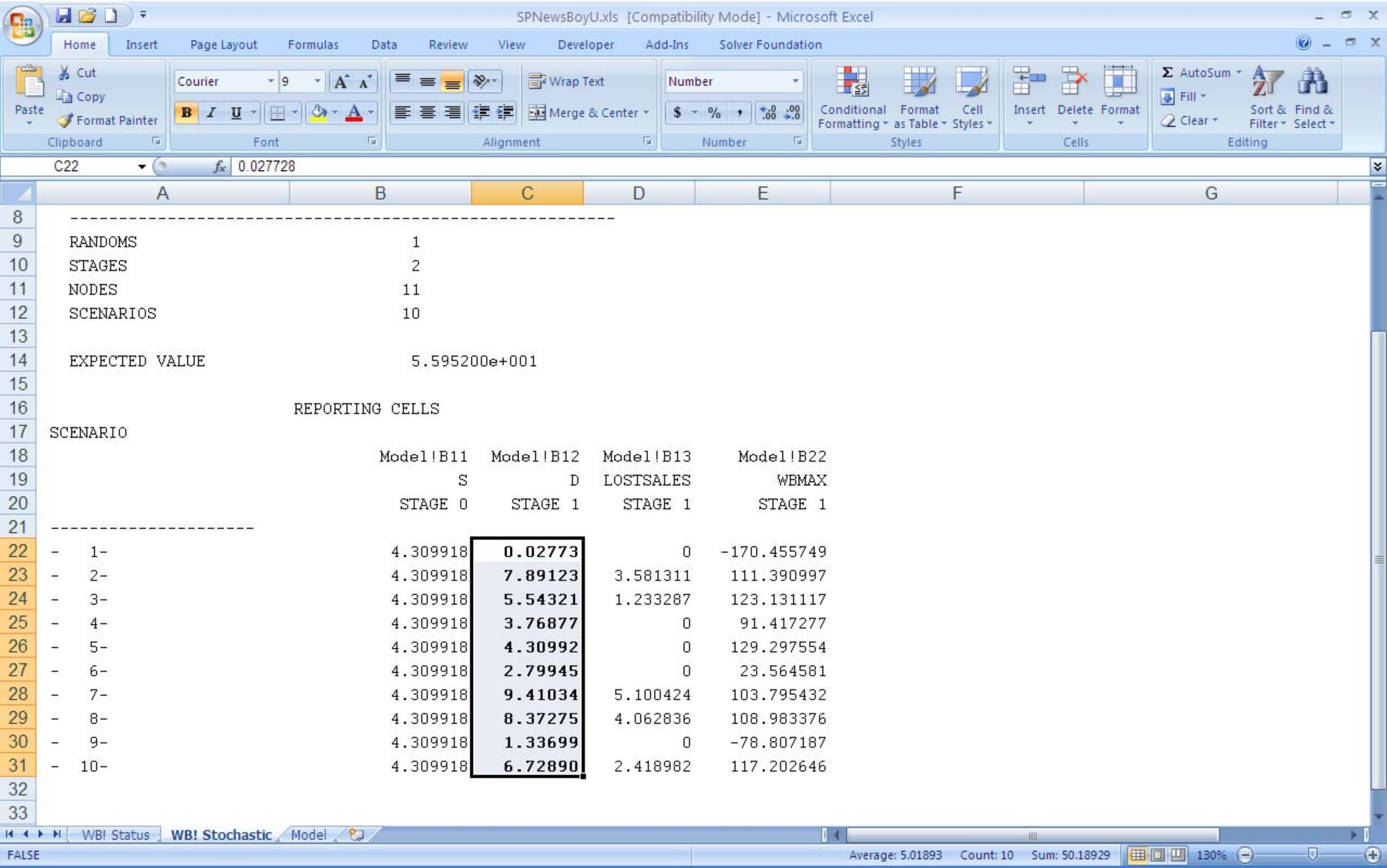
The other scenarios are generated "behind the scenes" during model generation, with the additional features that:

a) Cells designated as stage 0 decision variables are constrained to be equal in all scenarios,

b) Cells designated as stage 1 random variables(of some specified distribution) are replaced by a random variable in each scenario

c) the behind the scenes objective is to maximize net profit averaged over all scenarios.

LHS Illustrated, Notice "Super uniformity"



SPNewsBoyU.xls [Compatibility Mode] - Microsoft Excel

C22 0.027728

RANDOMS 1
STAGES 2
NODES 11
SCENARIOS 10

EXPECTED VALUE 5.595200e+001

REPORTING CELLS

SCENARIO

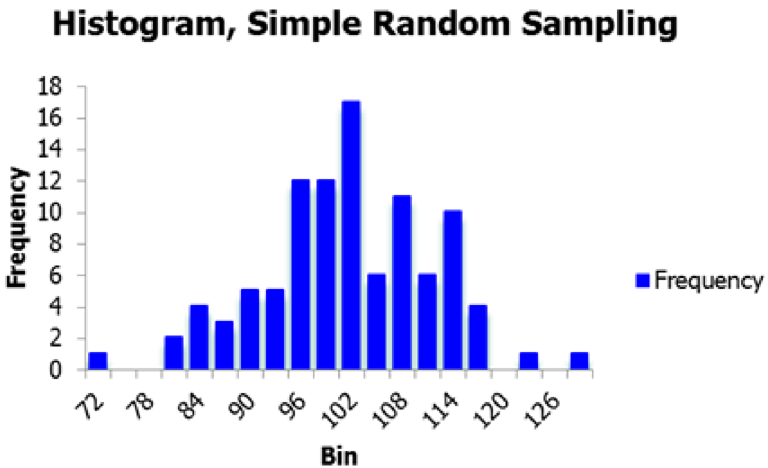
	Model!B11	Model!B12	Model!B13	Model!B22
	S	D	LOSTSALES	WBMAX
	STAGE 0	STAGE 1	STAGE 1	STAGE 1
- 1-	4.309918	0.02773	0	-170.455749
- 2-	4.309918	7.89123	3.581311	111.390997
- 3-	4.309918	5.54321	1.233287	123.131117
- 4-	4.309918	3.76877	0	91.417277
- 5-	4.309918	4.30992	0	129.297554
- 6-	4.309918	2.79945	0	23.564581
- 7-	4.309918	9.41034	5.100424	103.795432
- 8-	4.309918	8.37275	4.062836	108.983376
- 9-	4.309918	1.33699	0	-78.807187
- 10-	4.309918	6.72890	2.418982	117.202646

WB! Status | WB! Stochastic | Model

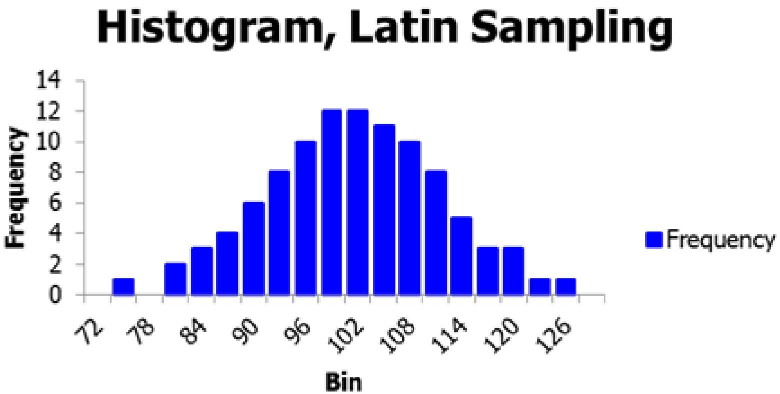
Average: 5.01893 Count: 10 Sum: 50.18929

Latin HyperCube vs. Simple Random Sampling

Generated a sample of 100 Normal demands with Mean = 100, SD = 10;



Mean = 100.31, SD = 10.14;



Mean = 99.98, SD = 9.98;

LHS Benefits, Optimistic Bias of Estimates from SP

If n = sample size, there is an optimistic "optimization" bias of the order of $(n-1)/n$ in the objective function value from simple SP.

Using LHS tends to reduce this bias, as well as the variance of the estimate. Some examples:

	Simple random sampling		LHS	
Problem	Mean	S. Error	Mean	S. Error
Newsvendor[1] **Min** cost, n =1000, r = 100;	5546.7	28.83	5547.2	9.86
Multi-product inv. [1] with random yield and partial substitution, **Max** profit, n =256, r = 100	189902	3162	189173	1275

[1]Yang, 2004.

Correlated Random Variables in SP

Three ways of measuring correlation:

Pearson

Define:

$$\bar{x} = \sum_{i=1}^{n} x_i / n; \qquad s_x = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 / (n-1)};$$

$$\rho_s = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) / (n s_x s_y);$$

Spearman Rank

Same as Pearson, except x_i and y_i replaced by ranks,
Minor adjustments when there are ties.

Kendall Tau Rank

$$\rho_\tau = \sum_{i=1}^{n} \sum_{k=i+1}^{n} 2 * sign[(x_i - x_k)(y_i - y_k)] / [n(n-1)]$$

LINDO SYSTEMS INC.

Advantages of Rank, and Copulas

If two random variables are Normal distributed, then it is relatively straightforward to generate them so they have a specified correlation (Pearson).

Challenge:
If two random variables have an arbitrary distribution, it is not so easy to give them a specified correlation.
Things are easy if we use rank correlation.
The rank correlation of two random variables is unchanged by a monotonic increasing transformation, e.g.,
Generating
Normal random variables from Uniform random variables
by the inverse cdf transformation method
does not change the rank correlation of the random variables.
The transformed Normals have the same rank correlation as the original uniforms.

Rank Correlation and Copulas

The Gaussian Copula is a way of generating set of d random variables, each with arbitrary marginal distribution, but having a specified d by d rank correlation matrix.

Procedure:

1) Generate a sample of size n of d Normal random variables having a specified rank correlation matrix. This is relatively easy.
2) Convert each of the d Normal random variables to uniforms with the transformation: $u_{ij} = F_{\text{normal}}(x_{ij})$.
3) Convert each uniform to the desired target marginal distribution with the inverse transform: (Steps 2 & 3 preserve rank correlation.)

$$y_{ij} = F_j^{-1}(u_{ij}).$$

The Gaussian Copula has been named as a culprit in the mortgage securities meltdown because of false confidence in a math model…..

LINDO SYSTEMS INC.

Kendall vs. Spearman Rank Correlation

+The Kendall correlation has a simple probabilistic interpretation. If (x_1, y_1) and (x_2, y_2) are two observations on two random variables that have a Kendall correlation of ρ_k, then the probability that the two random variables move in the same direction is $(1+ \rho_k)/2$. That is:

$$\text{Prob}\{(x_2 - x_1)*(y_2- y_1) > 0\} = (1+ \rho_k)/2.$$

For example, if the weekly change in the DJI and the SP500 have a Kendall correlation of 0.8, then the probability that these two indices will change in the same direction next week is (1+0.8)/2 = 0.9.

+The Spearman coefficient seems to be finer grained. E.g., the possible values for various sample sizes are:

Sample size	Kendall #Outcomes	Kendall Possible values	Spearman #Outcomes	Spearman Possible values
2	2	-1, +1	2	-1, +1
3	4	-1, -1/3, +1/3, +1	4	-1, -1/2, +1/2 1
4	7	-1, -2/3, …, +2/3, 1	11	-1, -4/5, …,+4/5, +1
5	11	-1, -4/5, …, +4/5, +1	21	-1, -9/10, …, +9/10, +1
6	16	-1, -91/105, -77/105,…,+1	36	-1, -99/105, -93/105,…,+1
. . .				

Also, Spearman matrix is always positive definite.

LINDO SYSTEMS INC.

Correlation Specification in What'sBest

SP_gas_linear_corr.xls [Compatibility Mode] - Microsoft Excel

M17 =WBSP_CORR_PEARSON(N19:O20,B13:B14)

Stochastic/Scenario Optimization of Gas Company Problem in What's*Best* *(Linear version)*

see http://stoprog.org/index.html?spintroduction.html

Given all costs and prices, in

Stage 0 we must decide how much gas to buy now to satisfy demand now, with excess to put into storage

Stage 1, in the beginning, unknown demand is revealed to us, and finally in

Stage 1, at the end, purchase additional gas if needed, probably at a high price.

1) The Core Model

P0 = Purchase cost/unit now=	5	
H=Holding cost/(unit leftover)=	1	
S0=Amount to purchase/stock this period=	195.81963	<<== Stage 0 decision.
D0=Demand this period=	100	
I0=Inventory left over= S0 - D0 =	95.8196302	<<== Stage 1 inventory.
P1 = Price next period=	7.74324096	<<== Stage 1 random price.
D1 = Demand next period=	196.633079	<<== Stage 1 random demand.
S1 = Amount to purchase next period=	100.813448	<<== Stage 1 (recourse) decision.
S1 >= D1 - I0 (constraint)	=>=	<<== Stage 1 constraint.
TC = Total cost of goods = P0*S0+P1*S1 =	1759.72097	<<== Stage 1 cost computation.
TH = Total Holding cost=H*I0 =	95.8196302	<<== Stage 1 holding cost computation.
Total cost = TP + TH =	1855.5406	<<== Stage 0 expected value to be minimized.

Add stochastic data here

2) Staging Information

WBSP_VAR				
	3) Distribution		Mean	StdDev
WBSP_RAND	WBSP_DIST_NORMAL		6.2	1.1
	WBSP_DIST_NORMAL		143	35
WBSP_VAR				
WBSP_STSC		WBSP_CORR_PEARSON		
			Correlation Matrix	
			1	0.9
			0.9	1

4) Sampling

Stage	Scenario
1	9

5) Reporting

Reporting cells (optional):

WBSP_REP

Overview:

The user enters only a generic scenario 1.

The other scenarios are generated "behind the scenes" during model generation.

with the additional features that:

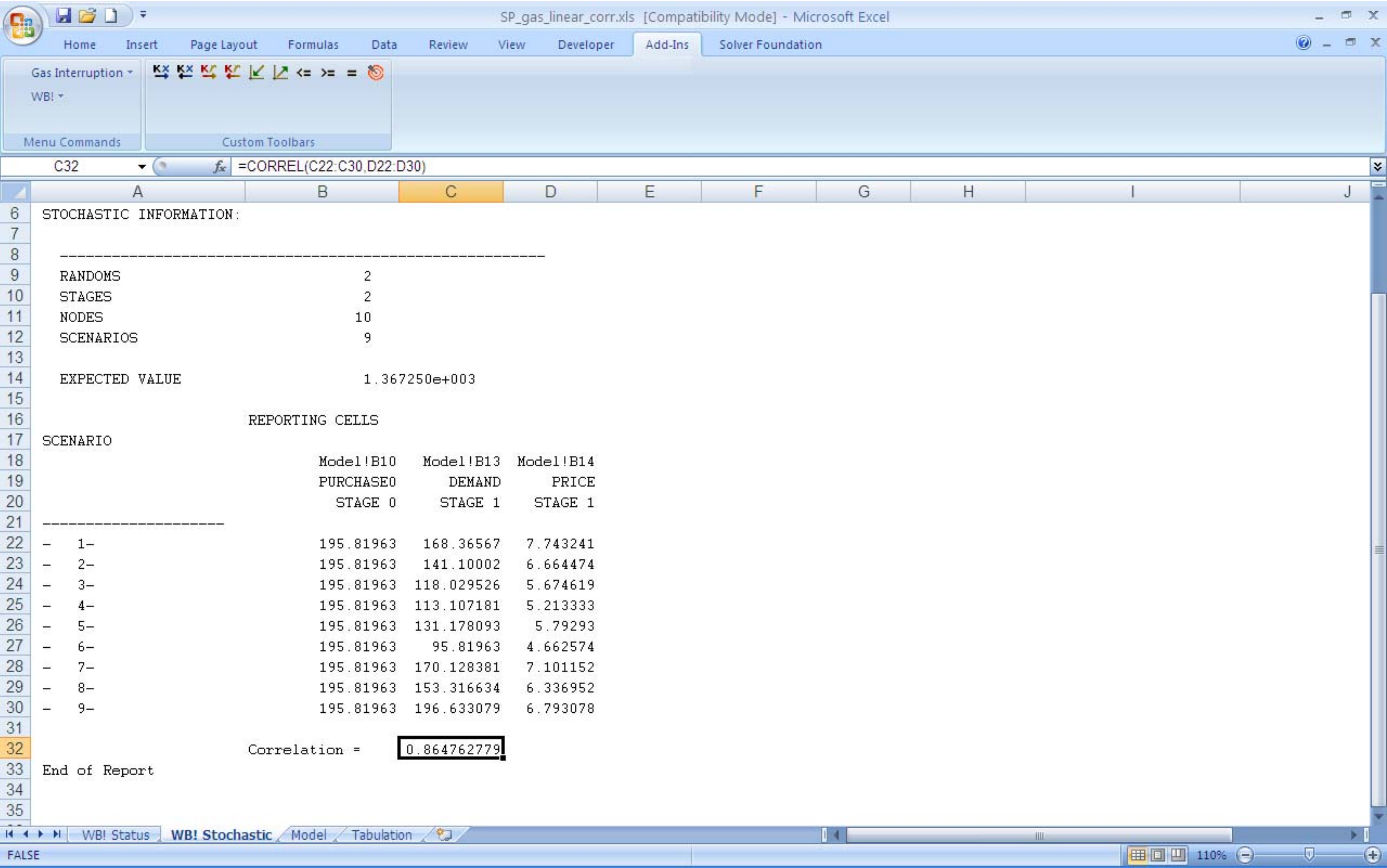
a) Cells designated as stage 0 decision variables are constrained to be equal in all scenarios

b) Cells designated as stage 1 random variables(of some specified distribution) are replaced by a random variable in each scenario.

c) the behind the scenes objective is to maximize net profit averaged over all scenarios.

WB! Status | WB! Stochastic | Model | Tabulation

Correlation Specification, cont.



STOCHASTIC INFORMATION:

RANDOMS	2
STAGES	2
NODES	10
SCENARIOS	9
EXPECTED VALUE	1.367250e+003

REPORTING CELLS

SCENARIO	Model!B10 PURCHASE0 STAGE 0	Model!B13 DEMAND STAGE 1	Model!B14 PRICE STAGE 1
– 1–	195.81963	168.36567	7.743241
– 2–	195.81963	141.10002	6.664474
– 3–	195.81963	118.029526	5.674619
– 4–	195.81963	113.107181	5.213333
– 5–	195.81963	131.178093	5.79293
– 6–	195.81963	95.81963	4.662574
– 7–	195.81963	170.128381	7.101152
– 8–	195.81963	153.316634	6.336952
– 9–	195.81963	196.633079	6.793078

Correlation = 0.864762779

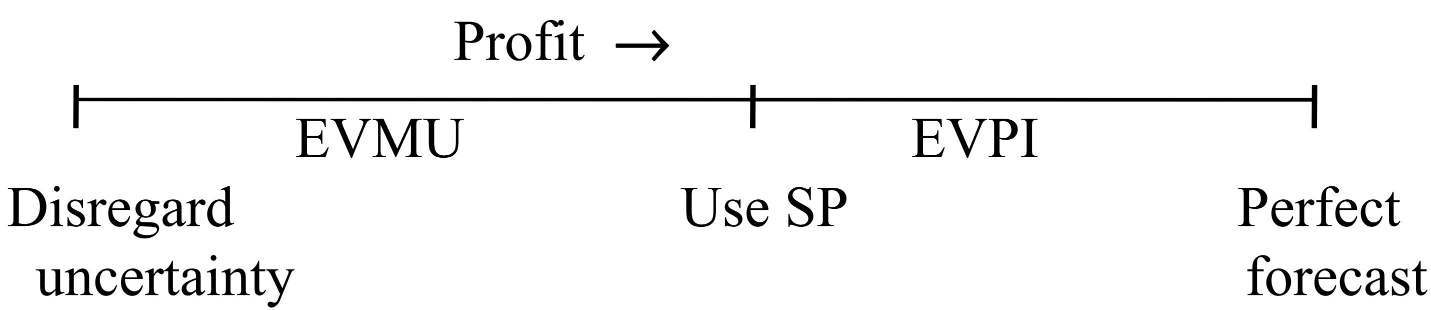
End of Report

How much is Uncertainty Costing us? EVPI and EVMU

EVPI (Expected Value of Perfect Information)
= Expected increase in profit if we know the future in advance.

EVMU (Expected Value of Modeling Uncertainty)
= Expected decrease in profit if we replaced each random variable by a single estimate and act as if this value is certain.

Typical single estimate is the estimated mean.
Why might you rather use the median?*

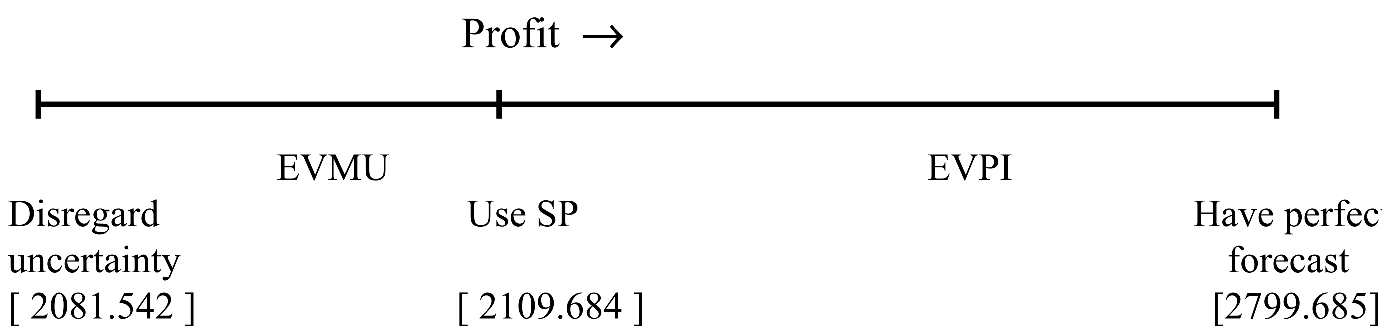


*We estimate that country *X* will have 1.823 aircraft carriers in 2012..

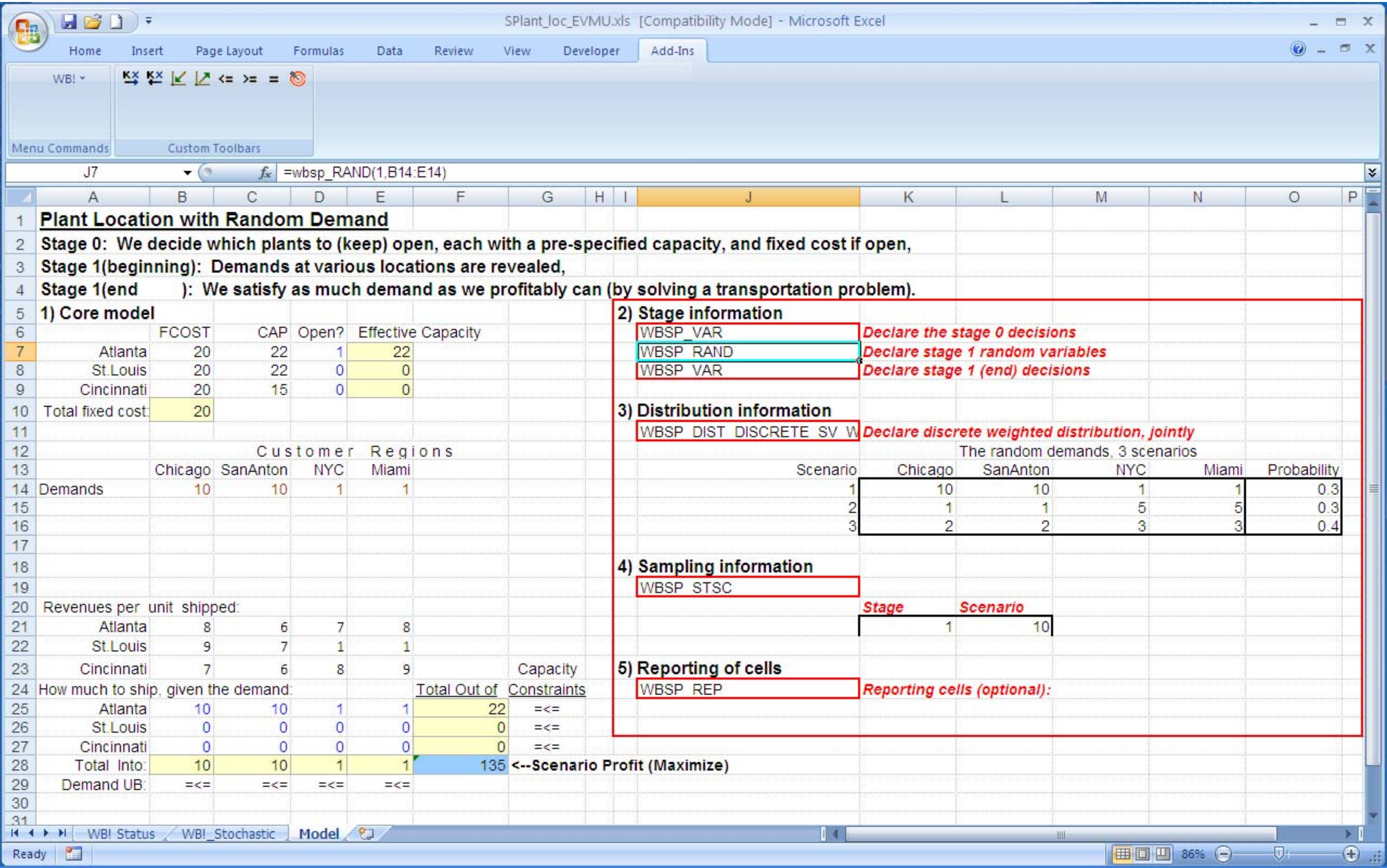
Expected Value of Better Modeling and/or Forecasting

EVMU and EVPI are provided in What'sBest! 10 for the Newsvendor model considered previously. The solution summary section is:

```
Objective (EV):                                        2109.684
Wait-and-see model's objective (WS):                   2799.685
Perfect information (EVPI = |EV - WS|):                 690.0007
Policy based on mean outcome (EM):                     2081.542
Modeling uncertainty (EVMU = |EM - EV|):                 28.14211
```



EVPI and EVMU: A Capacity Planning Example



EVPI and EVMU: Capacity Planning Example Output

SPlant_loc_EVMU.xls [Compatibility Mode] - Microsoft Excel

A31 — End of Report

```
Expected Value (EV)                                          82.40
Expected Value of Wait-and-See (EVWS)                        88.80
Expected Value using Expected Value Policy (EVEVP)           71.70
Expected Value of Perfect Information (=|EVWS-EV|)            6.40
Expected Value of Modeling Uncertainty (=|EV-EVEVP|)         10.70
```

REPORTING CELLS

		Model!B14	Model!C14	Model!D14	Model!E14	Model!B25	Model!C25	Model!D25	Model!E25	Model!B26	Model!C26	Model!D26	Model!E26	Model!]
		DEM_CHI	DEM_SAN	DEM_NYC	DEM_MIA	ATL_CHI	ATL_SAN	ATL_NYC	ATL_MIA	STL_CHI	STL_SAN	STL_NYC	STL_MIA	CIN_
SCENARIO	PROBABILITY	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAGE 1	STAG
- 1-	0.3	10	10	1	1	10	10	1	1	0	0	0	0	
- 2-	0.3	1	1	5	5	1	1	5	5	0	0	0	0	
- 3-	0.4	2	2	3	3	2	2	3	3	0	0	0	0	

End of Report

WB! Status | WB!_Stochastic | Model

EVPI Computations: Capacity Planning Example

```
If we know future only probabilistically..
                   Expected total profit =   82.40
 Plants to open:
  ATL

"Wait and See" Analysis, Perfect Information:
If we know scenario is 1,     then Profit=  142.00  (Probability=0.3)
 Plants to open:
       STL

If we know scenario is 2,     then Profit=   78.00  (Probability=0.3)
 Plants to open:
            CIN

If we know scenario is 3,     then Profit=   57.00  (Probability=0.4)
 Plants to open:
            CIN
                                            ------
 Expected Profit with Perfect Information    88.80  (=.3*142 + .3*78 + .4* 57)
 Simple Expected Profit                      82.40
 Expected Value of Perfect Information(EVPI)= 6.40
```

Notice Atlanta not optimal for any scenario!

LINDO SYSTEMS INC.

EVMU Computations: Capacity Planning Example

```
If we act as if mean demand is certain...
 The demand vector is:
    4.1    4.1     3     3
 Plants to open:
           CIN
Actual expected profit with this configuration= 71.7

    Expected Profit Modeling uncertainty=   82.40
    Expected Profit using expected values=  71.70
    Expected Value of Modeling Uncertainty= 10.70
```

EVPI Continued

If EVPI = 0 does this mean the value of doing SP = 0?

….we can buy this flexible facility for just a little more…

EVMU, When is it zero?

Can we predict when EVMU = 0?

E.g.,

Situation 1:

The price we get for our products are random variables.

Situation 2:

The demands for our products are random variables.

EVMU, Using Median vs. Mean

The default is to use the Mean.

+ Mean is intuitive for most people.

-Mean is undefined for some distributions, e.g., Cauchy.
Median is always defined for univariate distributions.

-Mean may not make sense for some situations,
e.g., discrete distribution. The average result of roll of a die is 3.5.
A fractional mean may not make sense. Median can always be chosen to be an actual possible outcome.

EVMU and EVPI, True vs. Estimated

A fine point: If the true number of scenarios is large, or infinite, and we use sampling, then the values for EVPI and EVMU reported are estimates rather than true values.

Computing Approximate Confidence Intervals

How confident should we be statistically,
of the results of an SP optimization?

Issue 1) There is an optimistic bias of the order of $(n\text{-}1)/n$ in the objective function value from an SP optimization. The optimization chooses the policy best for the sample observed.

Issue 2) If we use Latin Hypercube sampling, then the samples are correlated*, so an estimate of standard deviation among the samples based on the assumption of independence is wrong.

For modest size sample sizes, these two effects can be notable.
See the next slide for example.

*Generally negatively correlated. An observation or result far below the median will be compensated by an observation far above.

Approximate Confidence Intervals, an Example

```
! Newsvendor model;
 MU = 1000; ! Mean demand for the one period;
 SD =  300; ! Standard deviation in demand;
  V =  140; ! Revenue/unit sold;
  C =   60; ! Cost/unit purchased;
  P =    0; ! Penalty/unit unsatisfied demand;
  H =  -40; ! Holding cost/unit left in inventory;
  N = 15;   ! Number of scenarios sampled in the SP optimization.
              The 15 is chosen for illustrative purposes only, not necessarily
              a recommended sample size;
```

We repeated or replicated the above 15-sample SP 1000 times. For each replication we computed

a) the observed average profit, *xbar*;

b) the traditional "unbiased" estimate of the population standard deviation by

$[\Sigma_i (x_i - xbar)^2/(n-1)]^{0.5}$, and,

c) a 90% coverage interval for *xbar*, estimating the standard deviation of *xbar* by

$s = [\Sigma_i (x_i - xbar)^2/(n(n-1))]^{0.5}$.

For each replication we recorded whether the computed confidence interval in fact covered the true expected profit of $71,601. Results for the 1000 replications are shown below.

Sampling method	Mean profit	Mean sample standard deviation	Actual 90% confidence interval coverage
Random	$72,127	$25,945	.898
LHS	$71,595	$26,761	1.000
True/Analytical	$71,601		

LINDO SYSTEMS INC.

Approximate Confidence Intervals, Comments

Some things to note:

1) Because of the modest* number of scenarios, $n = 15$, SP with simple random sampling seriously overestimates the expected profit by $526. SP with LHS actually, by chance, slightly underestimates, by $6, the true expected profit.

2) The sample standard deviation under LHS is substantially less of an underestimate of the (unknown) population standard deviation in profit than is that under simple random sampling.

3) The confidence intervals computed under simple random sampling do not quite achieve the desired 90% coverage, perhaps because the intervals are not correctly centered because of the optimistic bias in *xbar*.

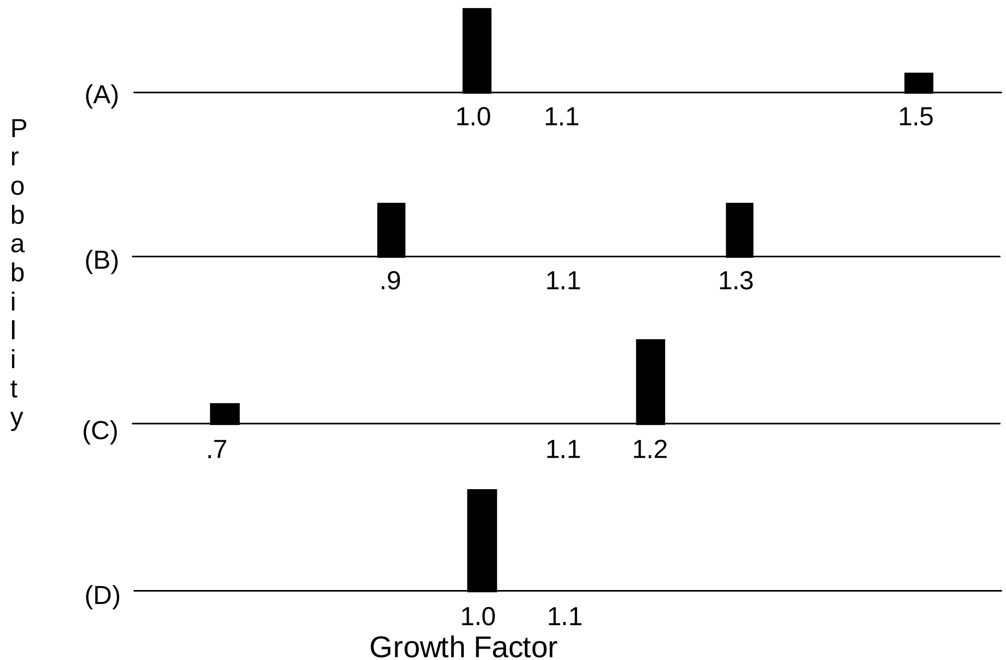
4) The confidence intervals from SP with LHS are extremely conservative, and in fact achieve 100% coverage,

* Roughly, a bias of $n/(n-1)$.

LINDO SYSTEMS INC.

Measures of Uncertainty: Variance, Risk, Utility,…

Which alternative investment : A, B, C, or D do you prefer?



Probabilities: A) .8, .2; B) .5, .5; C) .2, .8; D) 1.0. What are mean and s.d.?

Utility Function Approach to Measuring Risk

$U(w)$ = utility or value of having wealth w,

When w is a random variable, we want to
maximize $E[U(w)]$.
Qualitatively, if
$E[w_1] = E[w_2]$
but w_1 is "riskier" than w_2, what would we expect about
$E[U(w_1)]$ vs. $E[U(w_2)]$?

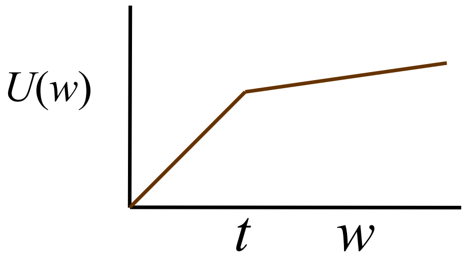
Reasonable features of $U(\ )$:
F1) Monotonic (strictly?) increasing.
"More is better",
Implies: a dominated random variable cannot be preferred.

F2) Concave(strictly?)
"Next *\$* not as useful as the previous *\$*"

Utility Functions, Popular Examples

May also specify a threshold *t,* and parameter *b.*

1) Downside: $U(w) = w - b*\max(0, t-w)$; $0 \le b \le 1$;



2) Quadratic: $U(w) = w - b*(t-w)^2$; $0 \le b$;

3) Power: $U(w) = (w^b - 1)/b$; $b \le 1$;

4) Log: $U(w) = \log(w)$, (Limit of Power utility as $b \rightarrow 0$); so-called "Kelly criterion".

LINDO SYSTEMS INC.

GM Model: Capacity Planning Under Uncertainty

Plant configuration decisions,

GM had too much capacity.

Needed to close or refocus an unknown number of plants.

Essential Structure:

Maximize expected profit contribution – cost of reconfiguration;

Cannot produce more in a plant than installed capacity;

Cannot sell more of a product than is demanded in a scenario.

GM SP Model, Special Features & Computations

+ Unsatisfied demand for a product transfers to other products according to a substitution matrix. One dozen products.

Key parameters:

c_{pv} = cost per unit to produce vehicle *v* in plant *p* (only possible if plant is open),

τ_{vw} = fraction of unsatisfied demand for vehicle v that transfers to vehicle *w*, (data from surveys),

$CAP_{p\sigma}$ = capacity of plant *p* in configuration *σ*,

Key variables:

x_{spv} = number of units of vehicle *v* produced in plant *p* in scenario *s*.

Other features:

+ Infinite final period.
+ Downside risk

LINDO SYSTEMS INC.

GM Model: Inventory Balance Constraint

The key constraints in words are:

For each scenario *s*

For each product (or vehicle) *v:*

$$Production_{vs} + Unsat_{sv} = Demand_{sv} + Transfer\_in_{sv};$$

For each vehicle *v* and *w* in scenario *s:*

$$Transfer\_from\_to_{svw} \le \tau_{vw} * Unsat_{sv};$$

For each plant *p* and configuration *σ:*

$$Total\_production_{sp} \le CAP_{p\sigma} * y_{p\sigma}$$

Downside Risk

$$penalty_s \geq threshold - profit_s ;$$

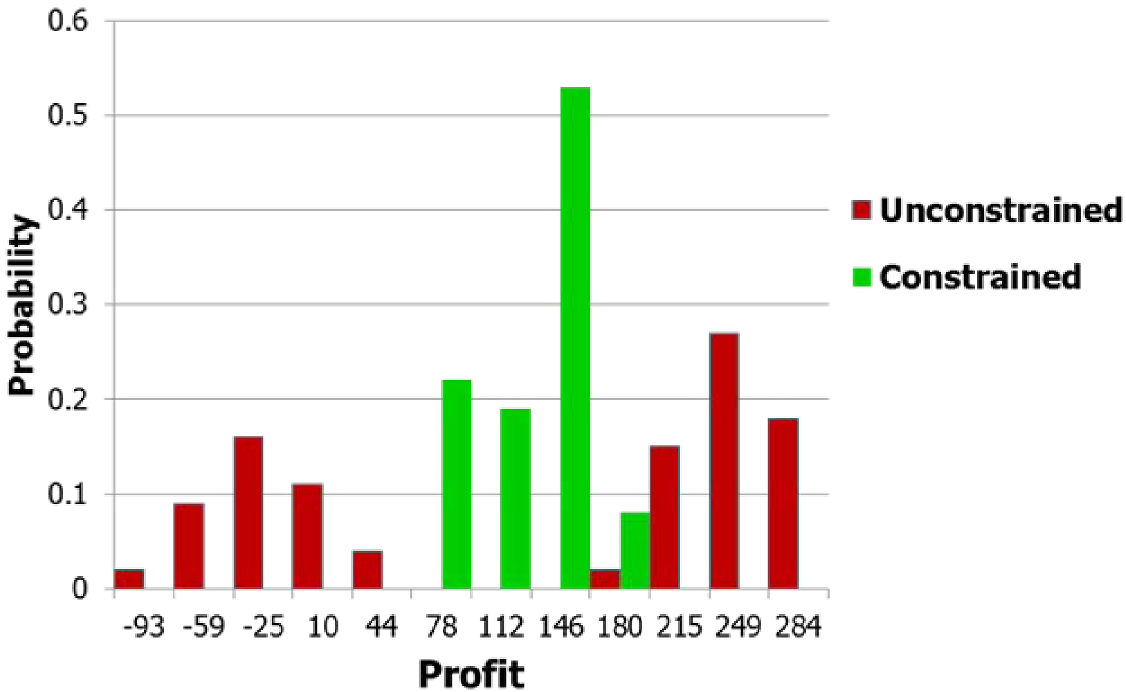
Expected downside risk constraint:

$$\sum_s Prob_s penalty_s \leq tolerance;$$

Both *threshold* and *tolerance* are parameters.

Effect of putting a constraint on Downside Risk

$$\sum_s \; Prob_s \, penalty_s \leq \; tolerance;$$



Airline Crew Scheduling, Deterministic Case

Approach used by many(most?) major airlines: Enumerate all interesting work patterns for a crew for a work period, e.g., day, week.

Variables:

y_p = 1 if crew work pattern p is used.

A work pattern is a sequence of flight legs.

Parameters:

a_{ip} = 1 if work pattern p includes flight leg i,

The deterministic, core model:

$$\text{Min } \Sigma_p \, c_p \, y_p \, ;$$

For each flight segment i, it must be covered by some pattern p:

$\Sigma_p \, a_{ip} \, y_p = 1$; Stage 1b constraints, for each scenario s:

Airline Crew Scheduling Under Uncertainty

A triggering delay may occur on a flight leg because of bad weather, equipment failure, etc.
A cascade delay can occur on a flight leg because of an earlier delay of one of the three entities* needed to execute a flight leg.

The SP approach (Air New Zealand, Yen & Birge)

Stage 0: Select a set of work patterns to use, the y_p.

Stage 1a: Random triggering delays occur.

Stage 1b: Compute the implied cascade delays and their costs.

*Plane, crew, passengers

Airline Crew Scheduling Under Uncertainty

How can the crew schedule chosen affect (cascade) delays?

If a flight leg is delayed(triggering or cascade), it could directly delay up to three immediately following flight legs:
1) A flight leg that needs the same plane,
2) A flight leg that needs the same crew,
3) A flight leg that needs a significant number of the same passengers.

If a work pattern keeps the crew on the same plane between two successive flight legs, then type 2 delay does not cause additional delay. So good work patterns from an uncertainty point of view keep the crew on the same plane.

Airline Crew Scheduling Under Uncertainty, details

Parameters:

R = set of leg pairs (i,j) for which i must arrive before j departs, because of plane or passengers,

w_{ijp} = 1 if leg i provides the crew for leg j under pattern p,

Stage 1a random parameters:

t_{is} = total flight time of leg i under scenario s,

Stage 1b decision variables:

d_{is} = departure time of leg i under scenario s,

r_{is} = arrival time or "ready for next leg" time of leg i, scenario s,

Stage 1b constraints, for each scenario s:

$r_{is} \geq d_{is} + t_{is}$, ! Flight time;

$d_{js} \geq r_{is}$ for i,j in R; ! Plane connection;

$d_{js} \geq \Sigma_i \Sigma_p w_{ijp}\, y_p\, r_{is}$; ! Crew connection (can be linearized);

LINDO SYSTEMS INC.

Airline Crew Scheduling Under Uncertainty, Full Formulation

! Minimize weighted combination of explicit cost + delay, where θ specifies the tradeoff between explicit costs and delays;

$$\text{Min } \Sigma_p \, c_p \, y_p + \theta \, \Sigma_i \Sigma_s \, d_{is};$$

! Stage 0 decisions and constraints,
For each flight segment i, it must be covered by some pattern p:

$$\Sigma_p \, a_{ip} \, y_p = 1;$$

$$y_p = 0 \text{ or } 1;$$

! Stage 1b constraints, to compute departure times, d_{is} , as a result of random leg times, t_{is} , for each scenario s…;

$r_{is} \geq d_{is} + t_{is}$, ! Ready time = departure + flight time;

$d_{js} \geq r_{is}$ for i,j in R; ! Plane connection;

$d_{js} \geq \Sigma_i \Sigma_p w_{ijp} \, y_p \, r_{is}$; ! Crew connection (can be linearized);

Metal Blending, The Problem

Stochastic Complication:

The composition (% C, %Si, %Cr, % Mn, etc.) of input materials, typically scrap, is a random parameter, i.e., known only approximately.

Stage 0:

Choose amounts x_j of various input materials, each containing a random fraction a_{ij} of target component i so as to approximately get mixture into target interval for component i.

Stage 1, beginning:

Melt mixture and observe actual composition for each i;

Stage 1, end:

Add additional, more pure and more expensive materials to move any wayward quality measures to within tolerance.

Recourse decision must be quick, < 1 min.

LINDO SYSTEMS INC.

Acknowledgments

This presentation benefited from the comments of *Sue Lisowski*.

References

Atlihan, M., K. Cunningham. G. Laude, and L. Schrage(2010), "Challenges in Adding a Stochastic Programming/Scenario Planning Capability to a General Purpose Optimization Modeling System", in *A Long View of Research and Practice in Operations Research and Management Science*, Springer, vol. 148, editors Sodhi, M. and C. Tang, pp.117-134.