ORIE 4820: Spreadsheet-Based Modeling and Data Analysis
Health Insurance Comparison Lab Exercise
Spring 2013

During this exercise, we will build a simplified decision tool that can be used to help individuals decide which Health Insurance Plan to select among the available options provided. The primary purpose is to give you practice in ***modeling decision problems with a fixed set of alternatives*** as well as ***incorporating uncertainty into evaluation models***. Although we will make extensive use of some of the Excel functions and tools we have already seen, we will also cover a number of new functions and tools that will enable us to ***facilitate sensitivity analysis***, ***visualize the impact*** of decisions and outcomes in a clear and compact way, and ***identify the "best" decision alternative*** according to user-specified parameters.

The template file for this exercise is ***Health-Insurance-Comparison.xlsm***. Download a copy of the file from the course Blackboard site and *save it on your computer*. There is one worksheet in the workbook and 6 Sections to this exercise in total:

- **Sections 1-3** focus on the *Health Insurance Plan Comparison* area
- **Sections 4-6** focus on the *Risk Analysis* area

<u>Topics and tools we will cover:</u>

- Developing a ***model to evaluate the consequences*** (i.e., total cost to individual) associated with a particular outcome (i.e., annual expense level) for all possible decision alternatives (i.e., plan choices)
- Using a ***one-way data table*** to enumerate the consequences for *every* possible outcome and decision alternative
- Creating a ***chart with data series of different types***
- Using ***data validation*** to restrict cell values
- Using ***conditional formatting*** to color code table entries
- Computing ***expected values*** for each decision alternative using a table of consequences and their corresponding probabilities: **sumproduct**
- Retrieving data using ***lookup and reference*** functions: **index, match**
- Using ***conditional logic*** functions as ***array functions*** to identify the "best" decision alternative meeting certain criteria
- Using a ***two-way data table*** to determine how varying the values of two different parameters affects the "best" decision alternative

If you have problems or questions at any point during the exercise, please raise your hand.

Background:

Many large companies are able to offer their employees a number of different insurance options to cover health care costs. Typically, a company will contract with one or more insurance providers that offer ***plans with different features*** (e.g., premiums, deductibles, co-payments, co-insurance, cost limits) to accommodate individuals with different risk tolerances and different types of life situations, including age, marital and dependent status, known health concerns, etc. Although in reality plan features can be quite complex (including differing coverage levels for specific types of services), for purposes of this exercise we will focus on *four parameters* for each health insurance plan alternative:

- Premium: The fixed cost that an individual pays each year for the privilege of having insurance coverage. This amount must be paid ***regardless of whether any health care expenses are incurred*** during the year.
- Annual Deductible: The amount of health care expenses that an individual is ***wholly responsible for*** in a given year. This can be thought of as the expense level at which the insurance plan "kicks in" each year.
- Co-insurance Percentage: The percentage of health care expenses ***beyond the deductible level*** that the individual is responsible for.
- Out-of-Pocket Limit: This feature limits the ***out-of-pocket*** amount that an individual is required to pay for health care expenses in a given year (i.e., the amount paid towards the deductible and co-insurance, but ***not*** the premium). Once this limit is reached, the insurance plan ***fully covers*** all additional health care expenses incurred during the year.

For example, suppose that Bob is covered by an insurance plan with the following features:

- Premium = $600
- Annual Deductible = $200
- Co-insurance Percentage = 20%
- Out-of-Pocket Limit = $2,500

If Bob incurs *no* health care expenses during the year, then the *total cost to him* would be the premium of $600.

If Bob incurs $1,200 of health care expenses during the year, then the *total cost to him* would be:

$600	+	$200	+	[20% * ($1,200-$200)]	=	**$1,000**
(Premium)		(Deductible)		(Co-insurance)		(Total Cost)

In this case, Bob's out-of-pocket expenses are $400, well below the limit of $2,500.

If Bob incurs $12,000 of health care expenses during the year, then because his out-of-pocket expenses ($200 + [20% * ($12,000-$200)] = $2,596) would *exceed* the Out-of-Pocket Limit, the *total cost to him* would be:

$600	+	$2,500	=	**$3,100**
(Premium)		(OOP Limit)		(Total Cost)

Before we begin, there are several things to note about the *Health Insurance Comparison* worksheet:

- This worksheet contains four ***named cell ranges*** (three single cell parameters, and one column of input parameters). To view all of the defined names in the workbook, along with their associated range addresses and properties, from the ribbon, select Formulas->Defined Names->Name Manager.
- The named cells Anticipated_Annual_Expense (F11), Cost_Threshold (R24), and Threshold_Tolerance (R26) have forms attached to them to facilitate data entry by the user (one has a ***scroll bar***, and two have ***spinners***). Recall how this is done: from the ribbon, select Developer->Insert, and under Form Controls, select the appropriate icon. Then "draw" the form anywhere on the worksheet. To attach it to a particular cell, right-click on the form, select Format Control, and set the parameters accordingly on the Control tab.

  Note: The spinner associated with the Threshold_Tolerance percentage is actually linked to the cell *beneath* Threshold_Tolerance and not the cell itself. This has been done as a work-around to account for the fact that *spinners and scroll bar forms can only change their linked cell values in integer increments*, not fractional or percentage values such as those needed for Threshold_Tolerance.
- ***Text concatenation*** has been used in several places to create dynamic labels or titles. To see this, press Ctrl-` to toggle between regular and formula views of the worksheet.
- A short VBA macro attached to the ActiveX control SpinButton_Tolerance has been written to enable decimal level incrementing of the Threshold_Tolerance cell. (Form controls are restricted to integer-valued increments.) To see the code, press Alt-F11 to open the Visual Basic Editor (VBE). In the left-hand side project window, double-click on ***Sheet 2*** within the object hierarchy. We will work more with VBA code in the weeks to come.

Section 1: Constructing the Total Cost Calculator

In the *Health Insurance Plan Comparison* area, we will develop a model to compare four insurance plan choices side-by-side for an entire spectrum possible annual health care expense levels. Given the plan parameters listed in the *Health Insurance Plans* table and the user-entered Anticipated_Annual_Expense, the table beneath will show the *Total Annual Cost to Individual Under Each Plan*. As we will see shortly, this table will be used as a *calculator* to facilitate other model computations. To set up the calculator:

(1) Enter the appropriate formula in cell C15 to ***compute the total cost*** for Plan1 assuming that Anticipated_Annual_Expense has been incurred:
"=C$5+MIN(C$8,IF(Anticipated_Annual_Expense<=C$6,Anticipated_Annual_Expense,C$6+C$7*(Anticipated_Annual_Expense-C$6)))"

(2) Copy this across to cells D15:F15.

(3) ***Populate cell G15*** with the *lowest total cost* among the values in C15:F15.

(4) ***Populate cell H15*** with the *corresponding plan name* using the **index** and **match** functions:
"=INDEX(C14:F14,MATCH(G15,C15:F15,0))"

Section 2: Computing Lowest Costs Using a One-Way Data Table

The calculator you just developed computes the total costs for all plans and identifies the lowest cost plan for a ***given*** Anticipated_Annual_Expense. In order to make an informed comparison of the plans, however, we need to understand how the lowest cost plan varies across an *entire spectrum* of possible Anticipated_Annual_Expense levels. We can accomplish this easily using a ***one-way data table*** in the area beneath cell C29.

Recall that a ***one-way data table*** allows the user to specify an entire set of different "what-if" values for a *single input parameter* (in our case, Anticipated_Annual_Expense levels). For each input value, the table *returns one or more computed quantities* that we specify with *formulas in the top row of the table* (row 31). To populate this table with the *total cost under each plan*, along with the *lowest cost plan* for the range of 151 different Anticipated_Annual_Expense levels specified in B32:B182:

(1) ***Populate cell C31*** with "=C15". Note that this formula simply refers to the "calculator table" cell with the computed total cost for Plan1. This is the output value that will populate this column of the data table *for each designated input value in column B.*

(2) Copy this across to cells D31:H31, so that all of the "calculator table" cells are captured.

(3) We are now ready to ***populate the data table***:

   a. Highlight the table range B31:H182 (the rectangle containing the column of input values and the top row cells you just populated, but *not* the text headers). From the ribbon, select Data->Data Tools->What-If Analysis->Data Table…

   b. Select F11 as the Column Input cell (i.e., each value in the input *column* will effectively be "substituted" into Anticipated_Annual_Expense), leave the Row Input cell blank, and press OK. The table should now be populated with the resulting values.

   Note: Once populated, you cannot alter the output range content of a data table (e.g., by inserting or deleting rows). However, you *can* format the cells, and you *can* change the top row formulas or the column input parameters. If the workbook has the Calculation option set to "Automatic", then the data table will automatically update anytime an input parameter or defining formula is changed.

Section 3: Constructing a Chart with Different Series Chart Types

(1) Using the data in the table we just populated, we will now ***construct a chart*** to visually compare the plans for different annual expense levels: Using the Ctrl-key, select the ranges C30:G30 and C32:G182, and insert a line chart (first option under Insert->Chart->Line).

(2) The result will not be exactly what we want, so we have to ***do some clean up***:

   a. Right click on the chart, and Select Data…

   b. Change the Horizontal Axis Labels to be the expense levels in cells B32:C182.

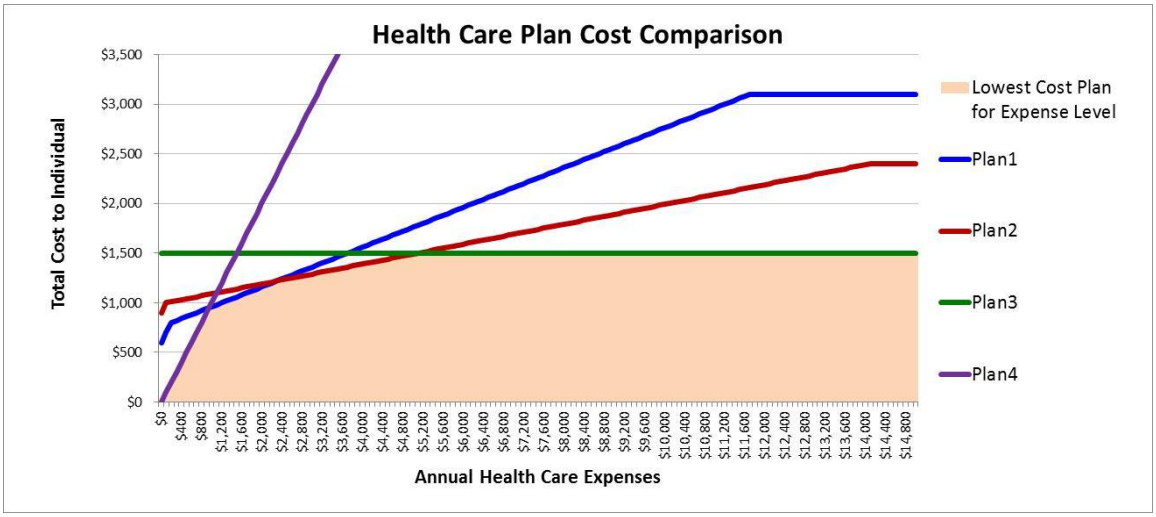
   c. Next, right click on the "Lowest Cost Plan for Expense Level" data series in the chart and select Change Series Chart Type… Select the first Column chart and click "OK".

d. Again, right click on the "Lowest Cost Plan for Expense Level" data series in the chart and select Format Data Series… Under Series Options, set the Gap Width to "No Gap", and set the Fill to be a color that contrasts with the other series lines.

e. To adjust the scale of the vertical axis, right click on the axis, and select "Format Axis"... Adjust the Maximum setting as appropriate.

f. To adjust the line widths of the other data series, right click on each in turn, select "Format Data Series", and adjust the Line Style.

(3) Click on the chart, then using Chart Tools-> Layout, use the appropriate options to ***update the chart title and axes*** accordingly:



We will now move on the *Risk Analysis* Area.

Section 4: Evaluating a Risk Profile Using a One-Way Data Table

In order for an individual to make an informed decision about what plan to choose, it is necessary to have an understanding of that individual's *risk profile*: i.e., how *likely* it is that the individual will incur different levels of health care expenses, as well as how tolerant that individual is of the possibility of paying high costs in any given year.

The first two columns of the top table in this area are input parameters intended to capture the first part of the risk profile. Although we will not attempt to anticipate every possible error that a user might make in entering this information, we will guard against a few common mistakes:

(1) Use ***data validation*** to insure that the probabilities entered in the first column are all between 0% and 100%.

(2) Use ***conditional formatting*** to change the color of the total cell to indicate when 100% has been reached (exactly).

(3) ***Populate cell R5*** with "=C15", and copy this across to cells S5:W5. As before, these formulas simply capture the values of the "calculator table" cells.

(4) We are now ready to ***populate the data table***:

a. Highlight the table range Q5:W15 (the rectangle containing the column of input values and the top row cells you just populated, but *not* the text headers). From the ribbon, select Data->Data Tools->What-If Analysis->Data Table…

b. Select F11 as the Column Input cell (i.e., each value in the individual's expense column will effectively be "substituted" into Anticipated_Annual_Expense), leave the Row Input cell blank, and press OK. The table should now be populated with the resulting values.

(5) Using the data in the populated table, ***create a line chart*** with 4 data series showing the total costs for each plan by annual expense level, then ***add a column data series on a secondary axis*** to show the corresponding probabilities associated with each annual expense level:

a. Right click on the chart and Select Data… and under the Legend Entries, click Add.

b. Under Series Name, type "Likelihood", and under Series Values, select the range P6:P15, then click "OK".

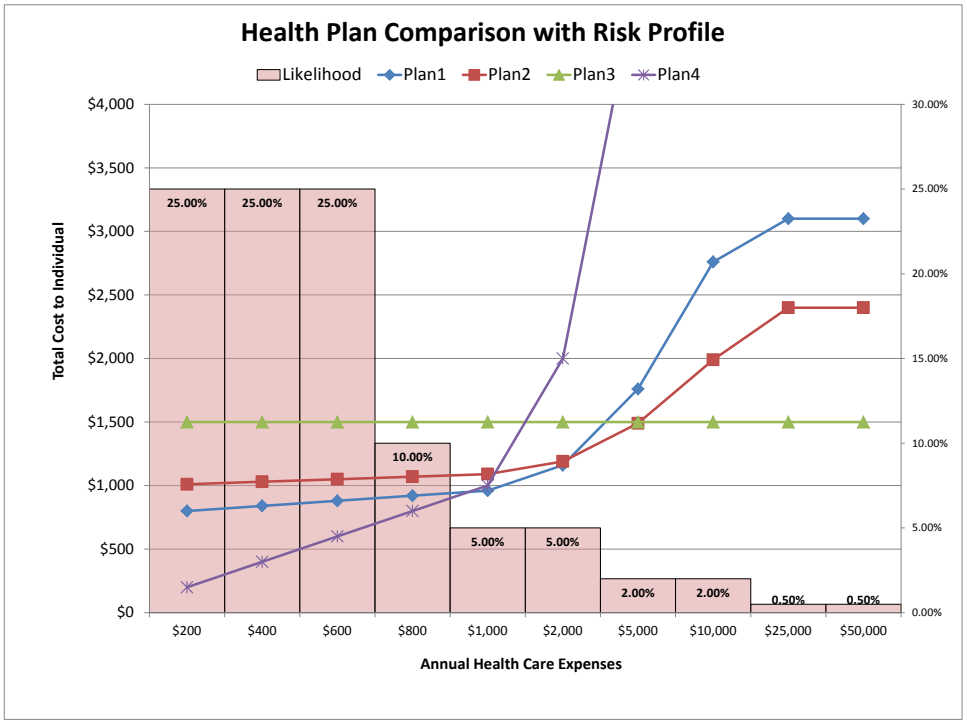
c. Set the Horizontal Axis Labels to be the range Q6:Q15, and click "OK".

d. Right click on the "Likelihood" data series within the chart and select Format Data Series… Under Series Options, select Secondary Axis.

e. Again, right click on the "Likelihood" data series and select Change Series Chart Type… Select the first Column chart and click "OK".

f. Again, right click on the "Likelihood" data series and select Format Data Series… Under Series Options, set the Gap Width to "No Gap", and set the Fill to be a color that contrasts with the other series lines.

g. Click on the chart, then using Chart Tools-> Layout, use the appropriate options to ***update the chart title and axes*** accordingly:



Section 5: Computing Metrics

The data table you just populated lists all of the possible *costs* that an individual might have to pay under each plan, as well as the *likelihoods* that each of these costs will be realized in any given year. These elements, taken together, allow us to compute many different *metrics* that can be the basis for an individual's decision.

(1) ***Average*** or ***expected costs*** are one possible metric that an individual might use to determine which insurance plan to choose. The ***expected annual cost level*** associated with a particular plan is defined as:

$$E[\text{Cost for Plan}] = \sum_{\substack{\text{All possible cost values k} \\ \text{under the Plan}}} k \cdot P(\text{Cost} = k)$$

a. Provided that the sum of the entered likelihoods (cell P16) is 100%, let's compute the expected costs for each plan in cells R19:U19 using the **sumproduct** function:

sumproduct(*list_of_values, list_of_probabilities*)

b. Note that ***conditional formatting*** has already been set on cells R19:U19 to highlight the cell having the *lowest* of the four values (i.e., a "Bottom 1" rule). To see this, highlight the cells, select Home->Styles->Conditional Formatting, then Manage Rules.

(2) A risk-averse individual might want to *minimize the maximum cost* that he might have to pay in any one year, so using the ***minimax*** metric to select a plan is another possibility.

a. In cell R20, compute the maximum cost that might have to be paid under Plan1, making sure to check that cell *P16 is 100%* and *only considering those rows that have a nonzero probability*. You can do this by entering the following formula as an ***array formula*** (**Ctrl-Shift-Enter**):

"=IF($P$16<>100%,"",MAX(IF(Probabilities>0,R$6:R$15,"")))"

b. Copy this formula to cells S20:U20. Again, a Bottom 1 ***conditional formatting*** rule has been entered for cells R20:U20.

(3) A more optimistic individual may select the plan that has the lowest cost *the majority of the time*. That is, the plan with the ***highest probability of having the lowest cost*** in a given year.

a. In cell R21, compute the *probability that Plan1 will have the lowest cost in any given year*: "=IF($P$16<>100%,"",SUMIF($W$6:$W$15,R$18,Probabilities))"

b. Copy this formula to cells S21:U21. Note that a Top 1 ***conditional formatting*** rule has been entered for cells R21:U21 to highlight the cell having the *highest* of the four values.

(4) Finally, it is reasonable to expect that an individual may want to know the ***probability that his cost in any given year will exceed a particular dollar threshold***.

a. In cell R22, compute the *probability that the cost in any given year will exceed the Cost_Threshold* specified in cell R24 (under Plan 1):

"=IF($P$16<>100%,"",SUMIF(R$6:R$15,">"&Cost_Threshold,Probabilities))"

b. Copy this formula to cells S22:U22. A Bottom 1 ***conditional formatting*** rule has been entered for cells R22:U22.

(5) Here is an example of a decision rule that takes multiple factors into account simultaneously:

Choose that plan that has the lowest average cost, <u>subject to the restriction</u> that the probability that that plan exceeds the Cost_Threshold is "reasonably low".

What does "reasonably low" mean? In cell R26 (Threshold_Tolerance), we will let the user specify his or her own personal risk tolerance for exceeding the Cost_Threshold. In cells V24 and V25, we want to ***indicate to the user which plan (if any) is "best"*** according to the above rule:

a. It is easiest to start with the cost cell here. In cell V25, enter the following as an ***array formula***: "=MIN(IF($R$22:$U$22>Threshold_Tolerance,"",$R$19:$U$19))"

b. In cell V24, enter the following as an ***array formula***:

"=IFERROR(INDEX($R$18:$U$18,MATCH(V25,IF($R$22:$U$22>Threshold_Tolerance, "",$R$19:$U$19),0)),"None")"

Note that the ***iferror*** function is being used to handle the case where no plan satisfies the Threshold_Tolerance restriction.

<u>Section 6: Measuring Sensitivity of the Best Plan to Threshold and Tolerance</u>

As a final piece to this exercise, we want to give the user an idea of ***how sensitive his or her "best" plan is to the Cost_Threshold and Threshold_Tolerance*** that are specified. We will do this using a ***two-way data table*** to indicate the "best" plan for an *entire matrix of combinations of Cost_Threshold and Threshold_Tolerance*.

A ***two-way data table accepts a row of candidate values for one input parameter and a column of candidate values for a second input parameter*** (i.e., Threshold_Tolerance and Cost_Threshold). For each combination of parameter values, the data table ***determines one output value*** (i.e., the "best" plan) and populates the associated matrix entry with this value. Directly beneath the Metric Table, a matrix has been created specifically for this purpose. To populate the matrix:

(1) Enter "=V24" in cell T31 to ***specify the resulting "best" plan***.

(2) Highlight the ***entire matrix*** starting from cell T31 (T31:AM82) and from the ribbon select Data->Data Tools->What-If Analysis->Data Table…

(3) Select R26 (Threshold_Tolerance) as the Row Input cell and R24 (Cost_Threshold) as the Column Input cell -- these specify the cells into which the pairs of values effectively will be "substituted" -- and press OK. The matrix will automatically populate with "best" plan values. ***The data table values will update automatically if any parameters are changed, or if any of the row or column input parameter values are changed.***

(4) Finally, note that ***conditional formatting*** has been set on the table cells to color code the various plan names that populate it. The following process was used:

a. Highlight the cells in the table (not including the row and column headers).

b. From the ribbon, select Home->Styles->Conditional Formatting.

c. Select New Rule, and in the upper Rule Type box, highlight "Format only cells that contain".
d. In the area "Format only cells with", select "Specific Text", "containing", and specify "Plan1".
e. Click Format, and choose your favorite cell fill color.
f. Click OK, then Apply.
g. Repeat the above steps for Plans 2, 3, and 4, specifying a different color each time.

IMPORTANT: ***DATA TABLES CANNOT BE NESTED***. That is, you cannot populate a data table with entries in which the calculation of the table entries depends upon the changing values in *another* data table. (Well, you can try, but Excel will not return the correct values.)