**Instructor Guide**

**Practical Spreadsheet Risk Modeling for Management**

**Lehman, Groenendaal, and Nolder**

It is our experience that instructor guides are rarely read, except for the solutions to text problems. So, we will keep this guide short and to the point. The exercise solutions are available in separate files, so this document contains only some advice, notes, and additional comments on particular exercises. **Please do not distribute or post solutions to the exercises, as these have a way of finding their way into the public domain.** The solution files each have a tab showing the results (based on 10,000 simulations with a manual random seed set = 0).

***General***

We have not provided PowerPoint slides to accompany the book, as it is often not good pedagogy – especially in the case of this book’s material. Spreadsheet modeling lends itself to a hands-on laboratory type approach. The exercises are challenging and educational and a good portion of course time should be spent working with students on their solutions to problems. This is at odds with a PowerPoint, lecture style approach to the subject. Instructors that want to prepare presentation materials can use the book’s figures as they wish – all of the figures are provided on this website.

The exercises are intentionally challenging (some problems will challenge even experienced risk analysts) and are meant to require the student to think through the entire problem in order to formulate a solution strategy. One approach we have found useful is to permit students to hand in late or redone homework after it is covered in class. If the solutions are not posted, it is valuable for students to build a model on their own, even if they have seen it solved by the instructor. Partial credit can be given for late work as an incentive for students to obtain a sense of mastery over the subject matter, even if they were not successful when first attempting the exercises on their own.

If a course has sufficient time, a course project (either individual or in small groups) can be a very good learning experience for both students and instructors. Students frequently choose projects that are too difficult to do within the time available in a course, so it is worth making sure that their topics are sufficiently narrow and well-defined to be feasible to handle within the course.

***Chapter 1 Material***

Depending on the audience, this material can be assigned for self-study, covered quickly, or could require a week of class time. It all depends on the students’ level of experience with building spreadsheet models. There are ample examples of pitfalls for novice modelers, as well as abundant evidence of the frequency of spreadsheet errors. Good visualization is essential to building useful models. For inexperienced modelers, a careful visualization can make the difference between building a functional model and not being able to begin a model at all.

Identifying variables concretely, as measurable quantities, is also important. For example, when modeling alternative supply chains, an objective of the cost per unit delivered (or the net present value of the cost per unit delivered) is better than a more diffuse objective such as “optimize the supply chain.” It is worth covering the textbook examples (cellular pricing, retirement) and having students work on some or all of the chapter exercises.

***Chapter 1 Exercise Notes***

1.1 It was intended that the office visit copays apply independently of whether or not the deductible has been met, but not independent of the out-of-pocket maximum. The interplay of these insurance elements is often complex and does not always work the same way. Of course, the assumption that is made will affect the numbers – the model in the solution makes this simplifying assumption.

1.2 The NPV of the savings is calculated assuming the entire cash flow of annual saving occurs in one “payment” halfway through the year.

1.3 The question asks for the additional NPV of obtaining a further degree, conditional upon already having a college degree. The solution assumes that wages for ages 20 through 24 would also be growing (at the same rate as the post-age 25 wages).

1.5 This problem may not be clear if you are not familiar with network effects. The underlying idea is that a network’s value increases with the number of people on the network. Intrinsic values refer to the value each individual has, conditional on everybody being part of the network: these intrinsic values will differ because people place different importance on the service. The realized value will depend on how many adopters there are – the simplest formulation (used in this exercise) is that the realized value is the product of the intrinsic value and the fraction of the population that adopts – this produces a quadratic equation for the realized value. The key feature of the solution is the critical mass: when the number of subscribers (users) is less than this level, the marginal user’s value is less than its cost and subscribership will be reduced (ultimately to zero). When it is greater than critical mass, it will continue to grow – to the stable mass market equilibrium level.

1.8 It is worth comparing the models using annual versus monthly calculations as they differ considerably. If you look ahead to chapter 2, you will see that the use of internal rates of return causes a number of problems since Excel produces errors when the IRR is sufficiently negative because the cumulative cash flows never become positive (which happens often in Exercise 2.8). The XIRR function is much more flexible than the IRR function, but has the same problem. Instructors may find it useful to digress on the more general properties of rates of return and their relationship with net present value. NPV is preferred to IRR since the latter has theoretical issues (non-uniqueness) as well as computational issues. However, the interpretation of IRR (when it can be relied on) is neat and does get around the thorny issue of how to specify the discount rate for computing NPV.

***Chapter 2 Material***

If Monte Carlo simulation is new to students, then this chapter should be covered carefully and slowly. If students are more familiar with simulation, this chapter could also be self-studied or covered quickly. The text examples and chapter exercises continue those from chapter 1 – it is valuable to see how much simulation adds to the richness of a model. The retirement example is a particularly good example, since the model in chapter 1 appears to provide a comfortable income level for retirement. When the investment return is uncertain, this comfort rapidly disappears.

When students are introduced to simulation software (whether it is ModelRisk or one of the other packages available), it is advisable to walk them through the options available in the simulation outputs. Frequency histograms are the most common view, but it is also useful to look at the cumulative distribution. Boxplots can be valuable, particularly when comparing (overlaying) a number of uncertain quantities. The statistics provide numerical results, but aside from the precision, rarely convey as much as the visual displays.

An additional note on the retirement planning example is warranted. The model in the book takes the retirement age as given. It is worth discussing whether retirement age is an objective, a decision, a parameter, or an uncertain parameter. Indeed, any of these is a plausible way to model this variable. Many people view the retirement age as their choice – others believe that the retirement age is their goal (to make it as early as possible). For others, health and labor market conditions may force retirement upon them. There is no single correct way to view this variable – ultimately, it depends on the decision context and what question we’re trying to answer with the model. Another guide is that the role of retirement age should not be buried within the analysis. That is, a good model should elucidate the role that retirement age plays.

***Chapter 2 Exercise Notes***

2.1 The way that tax treatment of health insurance affects the choice of plan is quite important. Tax deductibility of plan premiums lead to enhanced demand for “gold-plated” insurance. This, in turn, contributes to the cost of medical care (in the US).

2.2 Our solution assumes that the annual cash flow is received in one “payment” halfway through each year.

2.3 The question asks for the model to be run for two different interest rates – one indicative of federal loans and the other typical for private loans. The solution uses these two discount rates, and the NPV of these educational investment decisions is lower when the discount rate is higher. A more subtle (and more correct) analysis would only use the loan discount rate while the loan is being paid off – after that, the discount rate should reflect the appropriate opportunity cost for this individual. The exercise does not provide this information, so the solution simply uses two different discount rates to estimate the NPV.

Note that the solution spreadsheet embodies the possibility of unemployment within the formulas for the incremental cash flows. While it gives correct answers it may not be the best way to build the spreadsheet. A clearer method would be to create two columns – one for whether or not a high school graduate is unemployed, and the other for the possible unemployment of the college graduate. These would be simple VoseBernoulli distributions. Then, the earnings and foregone earnings can just be multiplied by (1-unemployment) so that when unemployment = 0, foregone earnings = 0. It is worth pointing out how this is a better modeling approach in terms of logical flow and clarity of troubleshooting the model.

2.8 One important thing to note is the dramatic difference between results when modeled as annual versus monthly cash flows. The latter is preferred. The XIRR function is not as well known as the IRR function, but is more useful in this case. XIRR can handle irregular cash flows, such as the monthly cash flow version of this problem. The IRR function runs into problems when encountering significant negative values (specifically, when cumulative cash flows never become positive) – Excel often reports errors as the result. This is particularly important with Monte Carlo simulation, since iterations may produce relatively extreme numerical results. In other words, the base case in the spreadsheet may look ok, but when running a large number of simulations, some iterations will produce very different numerical results. NPV is much preferred to internal rates of return and this is one reason. When using internal rates of return, it is important to keep track of the cases where errors result (rates of return are negative and large in absolute value). As a result, it is not meaningful to report the mean internal rate of return.

The XIRR function uses a default initial guess of 0.1. This can cause problems: the number of internal rates of return will equal the number of times the signs of the cash flows change. Then, the initial guess can impact which internal rate of return is calculated. In the present exercise, the cash flows always change sign once (from negative to positive), so this problem will not emerge. However, XIRR will not perform well (it produces values of 0.0000%) if the default initial guess is used but a large negative XIRR is the correct calculation. For this reason, it is best to use an initial guess = .1\*sign(sum of cash flows). Thus, when cash flows are negative (in summation), an initial guess of -0.1 is used. XIRR appears to produce no errors as a result. **Note that this correction appears in the Textbook Errata document**.

***Chapter 3 Material***

The use of Objects is unique to ModelRisk – however, much of the material in this chapter is relevant to any Monte Carlo simulation software. The need to model frequency/severity problems correctly may require more work for some software programs than for others, but it is important to understand that the simple approach of multiplying a frequency distribution (such as the Binomial or Poisson) by a severity distribution is generally incorrect. Simple multiplication is incorrect because the random value drawn from the severity distribution during each simulation iteration would implicitly assume that all of the events during that iteration have the same severity level when, in fact, a realistic scenario would be that each event has its own random severity. Correct modeling requires that the severity vary across events within a single iteration. If a frequency/severity model is built incorrectly, it could dramatically overestimate the risk (or variability) of the results. The Aggregate feature in ModelRisk accomplishes this for the user – in other programs, this may require additional modeling.

For simple problems, Objects may seem like an additional, and unnecessary, step. They are. However, there is still value in using Objects (and then the VoseSimulate function to simulate these Objects) for clarity and troubleshooting. The Objects can be kept in a separate section of the spreadsheet so that all of the uncertain variables appear in a concise area where they can be easily reviewed and/or modified. The calculation section of the spreadsheet then need not be touched in order to change the nature of the uncertainty of a particular parameter.

Objects are required for many of ModelRisk’s features, including the Aggregate functionality.

This chapter focuses on an insurance problem, although the Exercises show the wide applicability of frequency/severity modeling. The outcome variable in the insurance models is the average policy premium – this is done because insurance premiums are familiar to virtually all students. It is worth noting that insurance companies do not focus on this – their objective is the total loss function.

When using some of the advanced ModelRisk features (such as multidimensional aggregate modeling with deductibles), Excel may freeze after the model is run. A *hard reset* often overcomes this problem. This little known Excel feature requires simultaneously pressing Shift+Cntl+Alt+F9.

***Chapter 3 Exercise Notes***

3.1 A more realistic version of this problem would involve modeling each day of the year, including arrival times, queuing, overflow handling, etc. While this could be done in a spreadsheet, such discrete event simulation is best handled by specialized software designed for this task. For this more simple version of the problem, assume that supply and demand are even/flat over the course of a year and that all rooms are equally capable of handling patients. With the given input values, part (c) is not very interesting since occupancy is always above 90%. It might be better to ask students to change this part of the problem to estimate the probability that occupancy will exceed 90% and 95%.

3.3 The rate parameter (lambda) for the Poisson frequency distribution should be the mean number of occurrences = (claim frequency percentage)\*(% choosing each policy)\*(number of drivers).

3.4 This exercise can be solved with or without simulation. If the AggregateFFT method is used, then no simulation is required and this is a great demonstration of the power of Objects. Of course, this exercise can also be solved using the Aggregate Monte Carlo method. Note that a complete solution to this exercise is quite difficult and is best achieved using a Markov Chain analysis (covered in Chapter 6, section 6.8. Such an analysis could be used to examine how an initial stock of customers (for example, start with 1000 normal customers) would evolve over time, given transition probabilities derived from the given information (this is shown in the solution as an alternative solution). If you don’t want to cover Markov Chains, it is advisable to tell students to analyze the loyalty programs for two types of customers: normal and gold, but not to worry about a complete analysis (which would require information about the relative numbers of the two types of customers).

3.7 This is a model that may produce errors when you try to run it. If so, using the “hard reset” feature of Excel (Shift + Ctrl + Alt + F9) should allow it to work. This is a good example to point out this not well known Excel feature.

There are a number of interesting features of this exercise. Part a can be solved by either using the Aggregate Monte Carlo or the Aggregate FFT (this is a good example to show how to verify that the latter is a good approximation to the former – by looking at the FFT window’s statistics).

Part c also has two alternative methods. Without using objects, we can either model each of the customers separately, or utilize the Central Limit Theorem. In the present example, all four methods (two version of part a and two versions of part c) yield very similar results.

Students have found this to be an engaging exercise as most have encountered similar problems as both service providers and as customers. The exercise can easily be expanded to consider the relative costs of having an over-capacity or under-capacity of cashiers. It is also possible to build at extended example using different times of day with different customer arrival patterns.

***Chapter 4 Material***

There is a lot of material in this chapter. There are many distributions, both continuous and discrete, in Monte Carlo simulation software and these are related to each other in many subtle ways. Instructors may wish to explain the theory behind these distributions further – the chapter merely introduces the most commonly used distributions.

The first case in the chapter involves expert opinion for which ModelRisk has an extensive set of modeling tools. When building risk models, data should be used wherever possible. Most Monte Carlo software permits extensive fitting of distributions to data, but most simulation models will require the use of some expert opinion. The valuation example shows how a number of expert opinions can be combined (and weighted). The PERT distribution is the most common choice for expert opinion – it should be compared with the Triangular distribution. The latter is easy to understand and can be derived geometrically. However, the mean is quite sensitive to the minimum and maximum values – and the latter values are often difficult for experts to confidently estimate. The mean of the PERT distribution, on the other hand, is less severely affected by changes in the minimum and maximum values.

One note about the valuation example that bears discussion in class is the use of a terminal value. This is common practice in finance, but must be used carefully. Assuming terminal growth rates that exceed the expected growth rate of the economy is questionable. It is difficult to build a model that supports the current valuation of Apple without such an assumption, however.

The VaR case serves several purposes. It shows how to fit a distribution to data. It also shows how parameter uncertainty may be included in this fit. It is worth developing an intuitive understanding of how parameter uncertainty is modeled. We call this second order uncertainty. A bootstrap method is used, where the empirical data is assumed to be one random sample from the underlying distribution. Different random samples are then constructed and the best fitting parameters of that distribution are estimated. When there is a lot of data, inclusion of parameter uncertainty does not have a dramatic impact. However, when the data is sparse, the effect is more pronounced.

We use the continuous log-return rather than the discrete percentage change to measure daily price changes. It is worth demonstrating how close these two measures are (when the percentage change is small). ModelRisk uses Information Criteria to rank distribution fits rather than simpler criteria, such as the Kolmogrov Smirnov or Anderson Darling statistics. The latter are easier to calculate and more frequently encountered, but not as powerful. Instructors may wish to expand on this discussion.

We then show how uncertainty of the distribution fit may be incorporated. That is, a number of distributions may fit the data with varying degrees of precision, but we really don’t know which distribution is the correct one. Bayesian Model Averaging permits a number of distributions to be used – weighted according to how well each fits the given data. Model uncertainty need not increase the uncertainty in the outcomes – it may actually reduce it. This depends on which one of several distributions fits the data “best” and whether that distribution has a lot of variability relative to the other potential distribution fits. In any case, it should be stressed that visual examination of the distributions is required – it is not advisable to merely use the numerical information criteria alone to decide what distribution to use.

Instructors can easily spend additional time covering the major distributions and relationships among these. One type of distribution bears special consideration – these are the nonparametric distributions – CumulA, CumulD, Ogive, and Relative. These provide empirical fits to input data without assuming any particular functional form.

Section 4.8 on Bayesian Estimation is marked as optional but is worthwhile if you have time for it. The example just scratches the surface of Bayesian inference, but given the importance of this field, the time would be well spent. The example shows how the ability to learn after collecting data can affect the estimation of uncertainty. The result is that the forecast is more uncertain, not less. That is, the initial data is treated as a starting point, with the expectation that this will be updated with more information.

***Chapter 4 Exercise Notes***

4.1 A very interesting result obtains if the best fitting distributions for smokers and non-smokers are used (Logistic for non-smokers and Normal for smokers). If the Logistic is used for both, then the odds ratio of having either low birthweight babies or very low birthweight babies is greater for smokers than for non-smokers. When the differing distributions are used, the odds ratio of having very low birthweight babies is actually lower for smokers than non-smokers. However, this is not the result of a difference between the groups, but a difference between the Logistic and Normal distributions. The Normal distribution makes extreme events (such as very low birthweight babies) quite unlikely compared with the Logistic distribution. In this problem, using the Logistic distribution for both smokers and non-smokers probably gives the more accurate picture.

The solution file contains a number of extensions that are worth exploring. Deriving confidence intervals for the odds ratio requires building the model somewhat differently (using the VoseProb function so that an odds ratio can be calculated for each simulation iteration). An alternative approach is a Bayesian approach using the Beta distribution. The parameters for the Beta distribution are the number of events + 1 and the number of observations – the number of events +1.

Finally, the results can be compared with the results of a logistic regression – those show a mean odds ratio of 2.816 and a 95% confidence interval of (1.678, 4.835) for low birthweight babies. A logistic regression cannot be performed for the very low birthweight babies because there are no observed occurrences in the data set. It is worth emphasizing that point, as the simulation clearly permits exploration of events that do not happen in the available data.

4.2 Note that there are two data files. One has the complete monthly flight data for the US (a big file) and the other shows the extracted data for DEN – ORD flights. If instructors with to cover Excel’s data management tools (filtering, pivot tables, etc.), then they might want to use the bigger file and show students how to extract the needed data.

4.3 Building this spreadsheet is greatly simplified if Objects are used.

4.4 Note that the fitted distributions should be truncated to prevent negative numbers during simulation. This is a subtle, but common, occurrence (and error) in simulation modeling. During Monte Carlo simulations, extreme sets of variable values may result in unrealistic calculations – such as a negative fraction of houses with price cuts. An easy way to avoid this is to truncate the distribution so that negative values are not permitted. The probability of obtaining negative values is then reallocated to the rest of the probability distribution to preserve the relative probabilities of the allowed values. It is worth pointing out that the need to truncate distributions may be a sign that the wrong distribution is being used. In this exercise, the Beta distribution could be selected (it is near the top of the list of best fitting distributions to this data), which only permits fractional values. The solution provided does not use the Beta distribution, but it might be worth comparing the results.

4.5 This exercise is quite complex, but the required steps are shown in the text. Behind the detail, the exercise essentially compares two methods of using the given data – a fitted Lognormal distribution, and a constructed distribution to match the data given in the table. The latter constructs the distribution using a Step Uniform distribution along with a VLOOKUP function.

4.6 This exercise can be solved using simulation or by using Objects without simulation.

4.8 This exercise is fun and will certainly drive some students crazy. You can find websites where students can play this game which might provide a nice complement to the simulated solution (one example can be found at http://www.stat.sc.edu/~west/javahtml/LetsMakeaDeal.html). Three solutions are presented in the spreadsheet. The simpler (but not simplest) solution assumes that the initial door choice is always door #1 (this does not alter the results of model). The more complex solution (requiring more complicated logical statements) makes the initial door choice random. The much easier solution forgoes modeling what happens to each door and just models which door is chosen initially, which door has the price, which door Monte reveals, and which door the new choice is.

***Chapter 5 Material***

This chapter covers a lot of ground, particularly if instructors wish to cover regression analysis in detail (and if students do not have a good statistics background). It begins with the simplest type of relationship to model – logical relationships (IF, AND, OR, VLOOKUP, etc.). These can be quite powerful and can be used to model many types of relationships. Simpler models are always preferable, but when the relationships become dependent on continuous values, then more complex tools will be required. This chapter proceeds from simple to increasing complexity.

ModelRisk uses copulas to model correlations. ModelRisk implements the correlations generated by copulas through the U parameter in almost all available distributions. U parameters represent percentiles of a distribution. Thus, any type of correlation can be modeled by the relationship between these percentiles in two (or more) distributions. If two distributions are perfectly correlated, then if the 86th percentile of one distribution is simulated, the 86th percentile of the other will also be simulated. If there is no correlation between the distributions, then the percentile in the first will bear no relation to the percentile of the second that is simulated. Any intermediate relationship (including more complex correlation patterns) can be represented through the joint simulation(via copulas) of these U factors. Including the U factors in the distributions models the desired relationship between them.

Linear correlation is one of the simplest correlation patterns, but is a bit harder to incorporate in ModelRisk than in other software since it requires that it be implemented via a copula. The Normal copula is used for linear correlation in ModelRisk.

With multiple variables, the VoseValidCorrmat function is used to ensure that a correlation matrix is consistent. This is necessary since assumed pairwise correlation patterns will generally cause implied correlations between other variables that may not match the correlation entered by the modeler (for example, if variable A and variable B are correlated and B and C are correlated, then A and C will also be correlated). This is particularly true as the number of variables increases.

A number of more complex copulas are provided in ModelRisk and the examples in this chapter illustrate their importance. Home price changes have proven to be correlated, but the relationship is stronger when prices are decreasing significantly than at other times. Copulas can incorporate such patterns. Indeed, the Empirical copula can recreate almost any correlation pattern.

It should be noted that only the Normal copula and T copula permit the strength of correlations to vary across different pairs of variables. The other copulas can capture more general correlation patterns, but at the cost of losing the ability to vary the correlation strengths.

The next topic is regression modeling. Instructors may wish to supplement the material in the chapter since the use of regression models is usually an entire course. One particular area to expand upon is the use of linear regression to capture nonlinear relationships. The chapter shows how to do regression modeling in Excel and points out the main statistical assumptions and features in a regression model. The focus, however, is on simulation of regression relationships. This is important and natural, since the sample data describes a less-than-perfect relationship, and the regression output provides information about how (im)precise that relationship is. A bootstrap method permits the estimated regression line to be simulated. While the Linear Regression functionality of Excel provides nicely formatted output, the LINEST function is more useful for simulation purposes, since it automatically adjusts as the input data changes (which it will as a simulation is run). The LINEST function is not formatted nicely and not particularly easy to enter properly. Note that you may have to show students how to activate the Analysis Toolpak within Excel since it is not automatically activated upon installation.

The final (optional) section of the chapter provides a particular regression method for correlating variables when modeling expert opinion. The Envelope Method essentially uses a (regression) line to fit data solicited from an expert about how their choice of one variable may be impacted by their choice of another variable.

***Chapter 5 Exercise Notes***

5.1 Note that the probability of profit assuming success at phase 1 and phase 2 is the same because there is only a profit if the product passes all stages.

5.2 It is probably worth pointing out that this data is not from a random sample. It is the population of air routes for one three-month period and the analysis treats this data as if it is a random sample of data for other time periods. Though this is not really correct, it is common practice in the social sciences and business where controlled random experiments are not as easy to conduct as in the sciences. Part (b) is probably best done using the Regression Analysis tool in Excel, but part (c) will require use of the LINEST function. You should overlay the output cells to compare the results with and without parameter uncertainty (though they are quite close in this case).

5.5 Overlaying the box plots for the forecasts with and without parameter uncertainty is a good example of how modeling parameter uncertainty increases the variability of the forecast.

5.6 This example can be used for a more extended discussion of population biology models. The figure in the spreadsheet is a good example of quadratic regression and the bootstrap method for including uncertainty about the regression model.

5.8 This example can lead to an extended discussion of power laws and the long tail phenomenon. Some students will find the use of logarithms difficult, but this is a good example to demonstrate their usefulness.

***Chapter 6 Material***

Another lengthy chapter. ModelRisk contains many time series analysis tools, but it is important to realize that these do not automatically capture trend or seasonality. While the time series models can be fit to the raw data, the data should be cleaned of trend and seasonality before attempting to fit these models. These models attempt to describe the random variation in the data series, not the systematic variation – and trend and seasonality are systematic. ModelRisk does not automatically remove trend or seasonality. Indeed, while some software packages claim to do this, they do not correctly model the autocorrelation over time. Some software packages do not include time series modeling at all.

So, it is best to de-trend and de-seasonalize the data before attempting to fit time series models to it. We accomplish this through the use of Pivot Tables and we use the average seasonal factors and average trends. More sophisticated approaches are possible, particularly if the data are extensive. For the cases in this chapter, there is little basis for complex modeling of how seasonality or trend is changing over time, so the use of simple averages is justified.

When fitting the models to the data, it is best to show more than one forecast line (5 works well) in the time series fit window, and to use the Generate button a few times to see how the forecasts may vary. Just as with distribution fitting, it is critical to visually check the fit and not just rely on the Information Criteria to gauge the fit.

To paste the forecasts into the spreadsheet, the output range can be specified in the dialog box and the pasted cells will be an array formula. This means that you cannot specify these as outputs in ModelRisk (since Excel does not permit modification of array formulas). So, to display the forecast values, the desired cells should be replicated by creating duplicate cells – which can then be set as outputs (or as an output range to display a trend chart).

Alternatively, the time series Object can be placed in the spreadsheet. Then, the VoseTimeSimulate function should be used to simulate a forecast. This differs from the VoseSimulate function in that the autocorrelation is included when simulating.

Two other features of the Time Series dialog window should be noted. First, when modeling a time series of period to period changes (rather than the raw data), the LogReturns box should be checked (ModelRisk will normally automatically detect this). Second, when forecasting a series that starts at a particular point in time, the LastValue box should be checked and a value (or cell address) included to show where the forecast values start.

Once the forecast model is fit, the trend and seasonality need to be added back in.

The stock price case provides a good example where there is no trend or seasonality, but a systematic pattern still exists. Volatility trends are common in stock data and ModelRisk contains a number of time series models that can capture this.

The oil price example illustrates a limitation of time series models. The best fitting models do not really perform any better than naïve forecasts due to the volatile past history of these prices. Still, the example is useful for showing how time series methods can be combined with other modeling (expert opinion or relationship modeling) to produce hybrid forecast models.

The final (optional) section on Markov Chains is actually quite easy. All that is required is a transition probability matrix – and the assumption that the historic transitions do not impact the transition probabilities.

***Chapter 6 Exercise Notes***

6.3 This exercise provides an opportunity for a more extended discussion and analysis of different types of financial options. Once the forecasting model is built, it is easy to examine a variety of different options and simulate their values. The strangle option is but one example.

6.5 Note that the fractions that are forecast cannot be less than zero or greater than one.

6.6 It is important to stress that time series methods are not really designed to model 800,000 years of history. Still the data is interesting and important and the mechanical steps of constructing a forecast model are feasible.

6.8 The forecasts in this question perform much better if trends are removed before fitting a time series model to the data. Since the time scales for the two events are so different, for viewing a graph, it is best to normalize the data first (we indexed them so that the 1896 data equals 100 for both series).

***Chapter 7 Material***

This chapter can easily be expanded if instructors wish to cover optimization in more detail (as many courses do). We have chosen not to incorporate different types of optimization (linear, integer, etc.) since the optimization within ModelRisk can handle any sort of optimization problem (although it will not generally be the most efficient solution method).

This chapter explores optimization and simulation as complementary tools of analysis. Most decision problems are aimed at optimizing something in an uncertain environment. Use of simulation and optimization together permits a rich analysis of the potential tradeoffs between risks and rewards.

A single case is explored in this chapter – if instructors wish to expand the analysis, we would suggest including additional cases. The airline pricing problem is a prototypical example which has been extensively analyzed. Indeed, it is the seminal example for revenue management pricing.

Simulation tables can be useful for exploring risk/return tradeoffs and finding an appropriate solution without invoking an optimization algorithm. It is often useful to do this before running an actual optimization. It is important (at least when using OptQuest) to use an iterative procedure. When possible, run a discrete optimization before running a continuous one. And, run a series of discrete simulations with increasingly narrow ranges and smaller steps to converge on the solution. OptQuest takes time to run and it is much more efficient to proceed in a series of such steps rather than trying to obtain an exact solution in one step.

It is important to conceptually understand the difference between decisions, objectives, constraints, and requirements. The last two, in particular, are easily confused. But constraints act on decision variables and requirements act on objectives. Constraints can sometimes be modeled directly in a spreadsheet model rather than being specified as optimization constraints. Requirements illustrate the power of doing optimization within simulation since the optimal solution can be derived subject to a designated tolerable level of uncertainty.

Note that OptQuest can be used in the absence of uncertainty by selecting “without simulation” in the dialog. It has the advantage of being applicable to any sort of optimization problem, while Excel’s Solver is limited in the types of problems it can handle.

The final (optional) section on Stochastic Dominance is rather specialized though not difficult. It can be used to compare some probability distributions, though it often produces inconclusive results.

***Chapter 7 Exercise Notes: note that the solution spreadsheets are set up to handle the differing conditions in each part of the problem by enabling the appropriate target cells, decision variables, constraints, and requirements while disabling the others.***

7.1 The total budget should be entered as a requirement, not a constraint because it is a calculated cell rather than a direct constraint on the decision variables. This highlights the difference between the common language usage of the word “constraint” and the OptQuest meaning of the term.

7.2 Part (a) can be solved using Solver (answer: 62.33442 miles per hour). For Part (b), the optimum speed is 65. There is no feasible solution to Part (c).

7.3 This problem can provide the basis for an extended discussion of financial planning. Instructors can vary the time horizons and risk tolerances to explore how these affect optimum portfolios.

7.4 This is a discrete event model so it is rather complicated to build in a spreadsheet. It is feasible, however, and instructive about both the capabilities and limitations of spreadsheets.

7.5 As currently structured, the solution for all 3 parts of the problem is to do projects A,B,D, and F. To make the problem more interesting, some parameters will need to be changed. One suggestion is to change the probabilities to (0.6, 0.65, 0.8, 0.5, 0.75, 0.6, 0.85). If you do this, you should insert these values into cells E3:E9 of sheet 7.5c. The solution will now be to fund projects A,B,E,F. The moral of the story is that uncertainty sometimes changes the optimal solution and sometimes does not. In part (b), replacing certain values with uncertain and symmetric distributions does not alter the solution that maximizes the expected value. In part (c), changing the probabilities of success is a form of asymmetric uncertainty and does result in a change in the optimal solution.

7.6 This exercise can be used for an extended discussion of the potential conflicts between profitability and biological sustainability. The shorter the time horizon and/or the higher the discount rate, the greater is the probability of extinction when optimizing the NPV of the fish catch.

7.8 As stated in the book part d is superfluous since there is no probability that total revenue will be less than $50,000 in part c. The errata sheet says to change part c to $55,000. In that case, the solution in part d (shown in the answer spreadsheet) is to no put any seats in tier A. Intuitively, to reduce the risk this much requires eliminating the volatile tier A demand from the calculation. The mean total revenues drop from $64,826 to $62,964, as a result.

*If you discover errors in the exercises, solutions, or text, or have suggestions to pass along regarding the evolution and use of this book, please send them to Dale Lehman at* [*dlehman@alaskapacific.edu*](mailto:dlehman@alaskapacific.edu)*. Thanks for using our book and we hope you enjoy it.*