Chapter 7
Modeling and Simulation: Part 1

Contents

7.1 Introduction

The previous six chapters have provided us with a general idea of what a model is and how it can be used. Yet, if we are to develop good model building and analysis skills, we still need a more formal description, one that permits us a precise method for discussing models. Particularly, we need to understand a model's structure, its capabilities, and its underlying assumptions. So, let us carefully reconsider the question—what is a model? This might appear to be a simple question, but as is often the case, simple questions can often lead to complex answers. Additionally, we need to walk a fine line between an answer that is simple, and one that does not

H. Guerrero, *Excel Data Analysis*, https://doi.org/10.1007/978-3-030-01279-3_7

oversimplify our understanding. Albert Einstein was known to say—"Things should be made as simple as possible, but not any simpler." We will heed his advice.

Throughout the initial chapters, we have discussed models in various forms. Early on, we broadly viewed models as an attempt to capture the behavior of a system. The presentation of quantitative and qualitative data in Chaps. 2 and 4 provided visual models of the behavior of a system for several examples: sales data of products over time, payment data in various categories, and auto sales for sales associates. Each graph, data sort, or filter modeled the outcome of a focused question. For example, we determined which sales associates sold automobiles in a specified time period, and we determined the types of expenditures a college student made on particular days of the week. In Chaps. 3 and 5, we performed data analysis on both quantitative and qualitative data leading to models of general and specific behavior, like summary statistics and *PivotTables*. Each of these analyses relied on the creation of a model to determine behavior. For example, our paired t-Test for determining the changes in average page views of teens modeled the number of views before and after website changes. In all these cases, the model was the way we *arranged*, *viewed*, and *examined* data.

Before we proceed with a formal answer to our question, let's see where this chapter will lead. The world of modeling can be described and categorized in many ways. One important way to categorize models is related to the circumstances of their *data availability*. Some modeling situations are **data rich**; that is, data for modeling purposes exists and is readily available for model development. The data on teens viewing a website was such a situation, and in general, the models we examined in Chaps. 2, 3, 4, 5, and 6 were all data rich. But what if there is little data available for a particular question or problem—a **data poor** circumstance? For example, what if we are introducing a new product that has no reasonable equivalent in a sales market? How can we model the potential success of the product if the product has no sales history and no related product exists that is similar in potential sales? In these situations, modelers rely on models that *generate* data based on a set of underlying assumptions. Chaps. 7 and 8 will focus on these models that can be analyzed by the techniques we have discussed in our early chapters.

Since the academic area of Modeling and Simulation is very broad, it will be necessary to divide the topics into two chapters. Chapter 7 will concentrate on the basics of modeling. We will learn how models can be used and how to construct them. Also, since this is our first formal view of models, we will concentrate on models that are less complex in their content and structure. Although uncertainty will be modeled in both Chaps. 7 and 8, we will deal explicitly with uncertainty in Chap. 8. Yet, for both chapters, considering the uncertainty associated with a process will help us analyze the risk associated with overall model results.

Chapter 8 will also introduce methods for constructing *Monte Carlo* simulations, a powerful method for modeling uncertainty. Monte Carlo simulation uses random numbers to model the probability distributions of outcomes for uncertain variables in our problems. This may sound complicated, and it can be, but we will take great care in understanding the fundamentals—simple, but not too simple.

7.1.1 What Is a Model?

Now, let us go back to our original question—what is a model? To answer this question, let us begin by identifying a broad variety of model types.

1. **Physical model**: a physical replica that can be operated, tested, and assessed—e.g. a model of an aircraft that is placed in a wind-tunnel to test its aerodynamic characteristics and behavior.
2. **Analog model**: a model that is analogous (shares similarities)—e.g. a map is analogous to the actual terrestrial location it models.
3. **Symbolic model**: a model that is more abstract than the two discussed above and that is characterized by a symbolic representation—e.g. a financial model of the US economy used to predict economic activity in a unique economic sector.

Our focus will be on symbolic models: models constructed of mathematical relationships that attempt to mimic and describe a process or phenomenon. Of course, this should be of no surprise since this is exactly what Excel does, besides all its clerical uses like storing, sorting, manipulating, and querying data. Excel, with its vast array of internal functions, is used to represent phenomenon that can be translated into mathematical and logical relationships.

Symbolic models also permit us to observe how our decisions will perform under a set of model conditions. We can build models where the conditions within which the model operates are assumed to be known with certainty. Then the specific assumptions we have made can be changed, and the changed conditions applied to the model. Becoming acquainted with these models is our goal.

We can also build models where the behavior of model elements is uncertain, and the range of uncertainty is built directly into the model. This is the goal of Chap. 8. The difference between the two approaches is subtle, but under the first approach, the question that is addressed is—if we impose these specific conditions, what is the resulting behavior of our model? It is a very focused approach. In the latter approach we incorporate a broader array of possible conditions into the model and ask—if we assume these possible conditions, what is the full array of outcomes for the model? Of course, this latter approach is much broader in its scope.

The models we will build in this chapter will permit us to examine complex decisions. Imagine you are considering a serious financial decision. Your constantly scheming neighbor has a business idea, which for the first time you can recall, appears to have some merit. But the possible outcomes of the idea can result either in a huge financial success or a colossal financial loss, and thus the venture is very risky. You have a conservatively invested retirement portfolio that you are considering liquidating and reinvesting in the neighbor's idea, but you are cautious, and you wonder how to analyze your decision options carefully before committing your hard-earned money. In the past you have used intuition to make choices, but now the stakes are extremely high because your neighbor is asking you to invest the *entire* value of your retirement portfolio. The idea could make you a multi-millionaire or a penniless pauper at retirement.

Certainly, in this situation it is wise to rely on *more* that intuition. Chapters 7 and 8 will describe procedures and tools to analyze the risk in decision outcomes, both good and bad. As we have stated, this chapter deals with risk by answering questions related to *what* outcome occurs *if* certain conditions are imposed. In the next chapter we will discuss a related, but more powerful, method for analyzing risk—**risk profiles**. Risk profiles are graphical representations of the risk associated with decision strategies or choices. They make explicit the many possible outcomes of a complex decision problem, along with their estimated probability of occurrence. For example, consider the risk associated with the purchase of a one-dollar lottery ticket. There is a very high probability, 99%, that you will *lose* the dollar invested; there is also a very small probability, 1%, that you will *win* one million dollars. This risk profile is shown in Fig. 7.1. Note that the *win* outcome, $999,999, is the $1 million net of your $1 investment for the lottery ticket. Now let's turn our attention to classifying models.

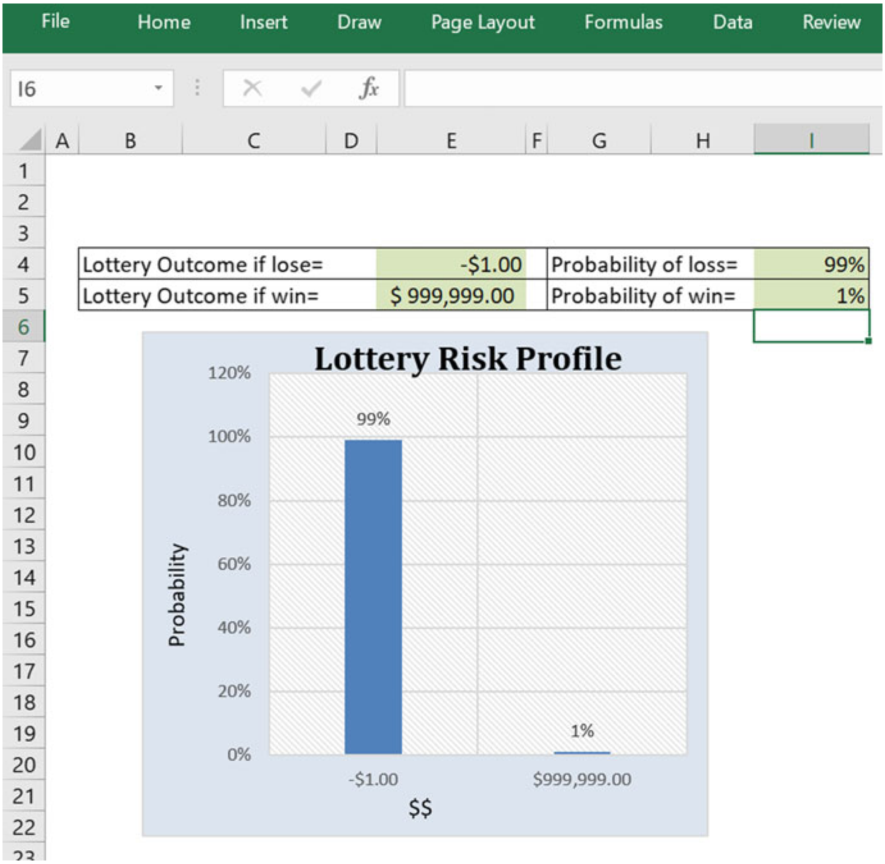


Fig. 7.1 Lottery risk profile

7.2 How Do We Classify Models?

There are ways to classify models other than by the circumstances within which they exist. For example, earlier we discussed the circumstances of data rich and data poor models. Another fundamental classification for models is as either deterministic or probabilistic. A deterministic model will generally ignore, or assume away, any uncertainty in its relationships and variables. Even in problems where uncertainty exists, if we reduce uncertain events to some determined value, for example an average of various outcomes, then we refer to these models as deterministic. Suppose you are concerned with a task in a project that you believe to have a 20% probability of requiring 2 days, a 60% probability of 4 days, and a 20% probability of 6 days. If we reduce the uncertainty of the task to a single value of 4 days, the average and the most likely outcome, then we have converted an uncertain outcome into a deterministic outcome. Thus, in deterministic models, all variables are assumed to have a specific value, which for the purpose of analysis remains constant. Even in deterministic models, if conditions change we can adjust the current values of the model and assume that a new value is known with certainty, at least for analysis. For example, suppose that you are trying to calculate equal monthly payments due on a mortgage with a term (30 years or 360 monthly payments), an annual interest rate (6.5%), a loan amount ($200 K), and a down-payment ($50 K). The model used to calculate a constant payment over the life of the mortgage is the PMT() financial function in Excel. The model returns a precise value that corresponds to the deterministic conditions assumed by the modeler. In the case of the data provided above, the resulting payment is $948.10, calculated by the function PMT(0.065/12,360,150,000). See Fig. 7.2 for this calculation.

Now, what if we would like to impose a new set of conditions, where all PMT() values remain the same, except that the annual interest rate is now 7%, rather than

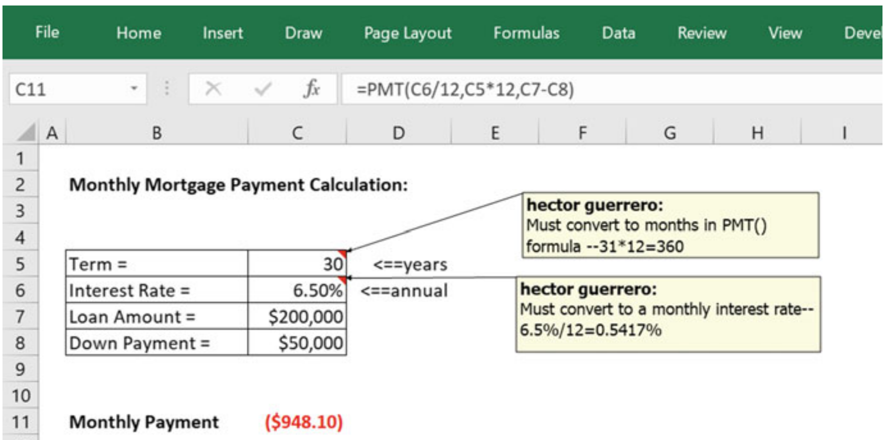


Fig. 7.2 Model of mortgage payments with rate 6.5%

6.5%. This type of *what-if* analysis of deterministic models helps us understand the potential variation in a deterministic model, variation that we have assumed away. The value of the function with a new interest rate of 7% is $997.95 and is shown in Fig. 7.3. Thus, deterministic models can be used to study uncertainty, but only through the manual change of values.

Unlike deterministic models, probabilistic models explicitly consider uncertainty; they incorporate a technical description of how variables can change, and the uncertainty is embedded in the model structure. It is generally the case that probabilistic models are more complex and difficult to construct because the explicit consideration of the uncertainty must be accommodated. But despite the complexity, these models provide great value to the modeler; after all, almost all important problems contain some elements of uncertainty.

Uncertainty is an ever-present condition of life and it forces the decision maker to face several realities:

1. First and foremost, we usually make decisions based on what we currently know, or think we know. We also base decisions and actions on the outcomes we expect to occur. Introducing uncertainty for both existing conditions *and* the outcomes resulting from actions can severely complicate decision making. Yet, it is usually the case that uncertainty applies equally to perceived, present conditions and anticipated, future outcomes.
2. It is not unusual for decision makers to delay or abandon decision making because they feel they are unable to deal with uncertainty. Decision makers often believe that taking *no* action is a superior alternative to making decisions with highly uncertain problem elements. Of course, there is no guarantee of this. Not acting can be just as damaging as acting under difficult to model uncertain circumstances.

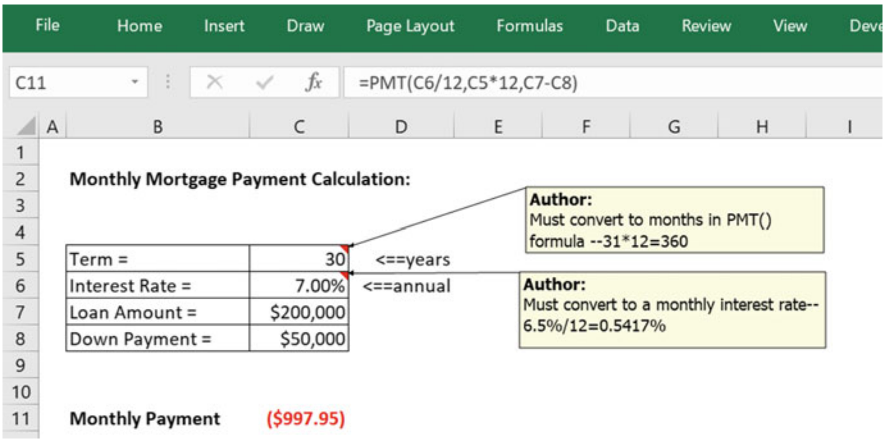


Fig. 7.3 Model of mortgage payments with rate 7.0%

3. Decision makers who incorporate a better understanding of uncertainty in their modeling and how uncertainty is related to the elements of their decision problems are far more likely to achieve better results than those who do not.

So, how do we deal with these issues? We do so by *systematically* dealing with uncertainty. This suggests that we need to understand a number of important characteristics about the uncertainty that surrounds our problem. In Chap. 8 we will see precisely how to deal with these problems.

7.3 An Example of Deterministic Modeling

Now, let us consider a relatively simple problem from which we will create a deterministic model. We will do so by taking the uncertain problem elements and converting them into deterministic elements. Thus, despite the uncertainty the problem contains, our approach will be to develop a deterministic model.

A devoted parish priest, Fr. Moses Efia, has an inner-city parish in a large American city, Baltimore, MD. Fr. Efia is faced with a difficult situation. His poor parish church, Our Lady of Perpetual Succor (OLPS), is scheduled for removal from the official role of Catholic parishes in the Baltimore dioceses. This means that the church and school that served so many immigrant populations of Baltimore for decades will no longer exist. A high-rise condominium will soon replace the crumbling old structure. Fr. Efia is from Accra, Ghana and he understands the importance of a community that tends to the needs of immigrants. Over the decades, the church has ministered to German, Irish, Italian, Filipino, Vietnamese, Cambodian, and most recently Central American immigrants. These immigrants have arrived in waves, each establishing their neighborhoods and communities near the church, and then moving to other parts of the city as economic prosperity has taken hold.

Fr. Efia knows that these *alumni* of OLPS have a great fondness and sense of commitment to the church. He has decided to save the church by holding a fundraising event that he calls Vegas Night at OLPS. His boss, the Archbishop has strictly forbidden Fr. Efia to solicit money directly from past and present parishioners. Thus, the event, appropriately named to evoke Las Vegas style gambling, is the only way to raise funds without a direct request of the parish alumni. The event will occur on a Saturday afternoon after evening mass, and it will feature several games of fortune. The Archbishop, a practical and empathetic man, has allowed Fr. Efia to invite alumni, but he has warned that if he should notice anything that suggests a direct request for money, he will cancel the event. Many of the alumni are now very prosperous and Fr. Efia hopes that they will attend and open their pockets to the event's games of chance.

7.3.1 A Preliminary Analysis of the Event

Among one of his strongest supporters in this effort is a former parishioner who has achieved considerable renown as a risk and data analyst, Voitech Schwartzman. Voitech has volunteered to provide Fr. Efia with advice regarding the design of the event. This is essential since an event based on games of chance offers no absolute guarantee that OLPS will make money; if things go badly and luck frowns on OLPS, the losses could be disastrous. Voitech and Fr. Efia decide that the goal of their *design* and *modeling* effort should be to construct a tool that will provide a forecast of the revenues associated with the event. In doing so, the tool should answer several important questions. Can *Vegas Night at OLPS* make money? Can it make too little revenue to cover costs and cause the parish a serious financial problem? Can it make too much revenue and anger the Archbishop?

Voitech performs a simple, preliminary analysis to help Fr. Efia determine the design issues associated with *Vegas Night at OLPS* in Table 7.1. It is a set of questions that he addresses to Fr. Efia regarding the event and the resolution of the issues raised by the questions. You can see from the nature of these questions that Voitech is attempting to lead Fr. Efia to think carefully about how he will design the event. The questions deal specifically with the types of games, the sources of event revenues, and the turn-out of alumni he might expect.

This type of interview process is typical of what a consultant might undertake to develop a model of the events. It is a preliminary effort to *define* the problem that a client wants to solve. Fr. Efia will find the process useful for understanding the choices he must make to satisfy the Archbishop's concerns—the games to be played, the method for generating revenues, the attendees that will participate, etc. In response to Fr. Efia's answers, Voitech notes the resolution of the question or the steps needed to achieve resolution. These appear in the third column of Table 7.1. For example, in question 4 of Table 7.1, it is resolved to consider a number of possible attendance fees and their contribution to overall revenues. This initial step is critical to the design and modeling process, and is often referred to as the **model or problem definition phase**.

In the second step of the model definition phase, a *flow diagram* for planning the OLPS event process is generated. This diagram, which provides a view of the related steps of the process, is shown in Fig. 7.4. Question 1 was resolved by creating this preliminary diagram of the process, including all its options. Since the event in our example is yet to be fully designed, the diagram must include the options that Fr. Efia believes are available. This may not always be the case. It is possible that in some situations you will be provided a pre-determined process that is to be modeled, and as such, this step will not include *possible* design options. The answers to Voitech's questions and the discussion about the unsettled elements of the game permit Voitech to construct a relatively detailed process flow map of the event and its options.

The process flow model, at this point, does not presume to have all questions related to the design of the event answered, but by creating this diagram Fr. Efia can

Table 7.1 Simple analysis of Fr. Efia's risk related to Vegas night at OLPS

Voitech's question to Fr. Efia	Fr. Efia's answer	Resolution (if any)
1. How do you envision the event?	I'm not sure. What do you think? There will be food and gambling. The archbishop is not happy with the gambling, but he is willing to go along for the sake of the event	Let's create a diagram of the potential process—See Fig. 7.4
2. What games will be played?	• The bowl of treachery • Omnipotent two-sided die • Wheel of outrageous Destiny	We will have to think about the odds of winning and losing in the games. We can control the odds
3. Will all attendees play all games?	I don't know. What do you think? But I do know that I want to make things simple. I will only allow attendees to play a game once. I don't really approve of gambling myself, but under these special circumstances a little won't hurt	Let's consider a number of possibilities–attendees playing all games only once at one end of the spectrum, and at the other end, attendees having a choice of the games they play and how often they play. I am just not sure about the effect on revenue here
4. Will the games be the only source of income?	No. I am also going to charge a fee for attending. It will be a cover charge of sorts. But, I don't know what to charge. Any ideas?	Let's consider a number of choices and see how it affects overall revenues
5. How many alumni of OLPS will attend?	It depends. Usually the weather has a big effect on attendance	Let's think carefully about how the weather will affect attendance
6. Will there be any other attraction to induce the OLPS alumni to attend?	I think that we will have many wonderful ethnic foods that will be prepared by current and past parishioners. This will not cost OLPS anything. It will be a contribution by those who want to help. We will make the food concession *all-you-can-eat*. In the past this has been a very powerful inducement to attend our events. The local newspaper has even called it the *Best Ethnic Food Festival* in the city!	I urge you to do so. This will make an entry fee a very justifiable expense for attendees. They will certainly see great value in the excellent food and the *all-you-can-eat* format. Additionally, this will allow us to explore the possibility of making the entry fee a larger part of the overall revenue. The archbishop should not react negatively to this

begin to comprehend the decisions he must make to execute *Vegas Night at OLPS*. This type of diagram is usually referred to as a **process flow map** because of the directed flow (or steps) indicated by the arrows. The rectangles represent steps in the process. For example, the *Revenue Collected* process indicates the option to collect an attendance or entry fee to supplement overall revenues. The diamonds in the diagram represent decision points for Fr. Efia's design. For example, the *Charge an Entry Fee?* diamond suggests that to finalize the event, Fr. Efia must either decide whether he will collect an entry fee or allow free admission.

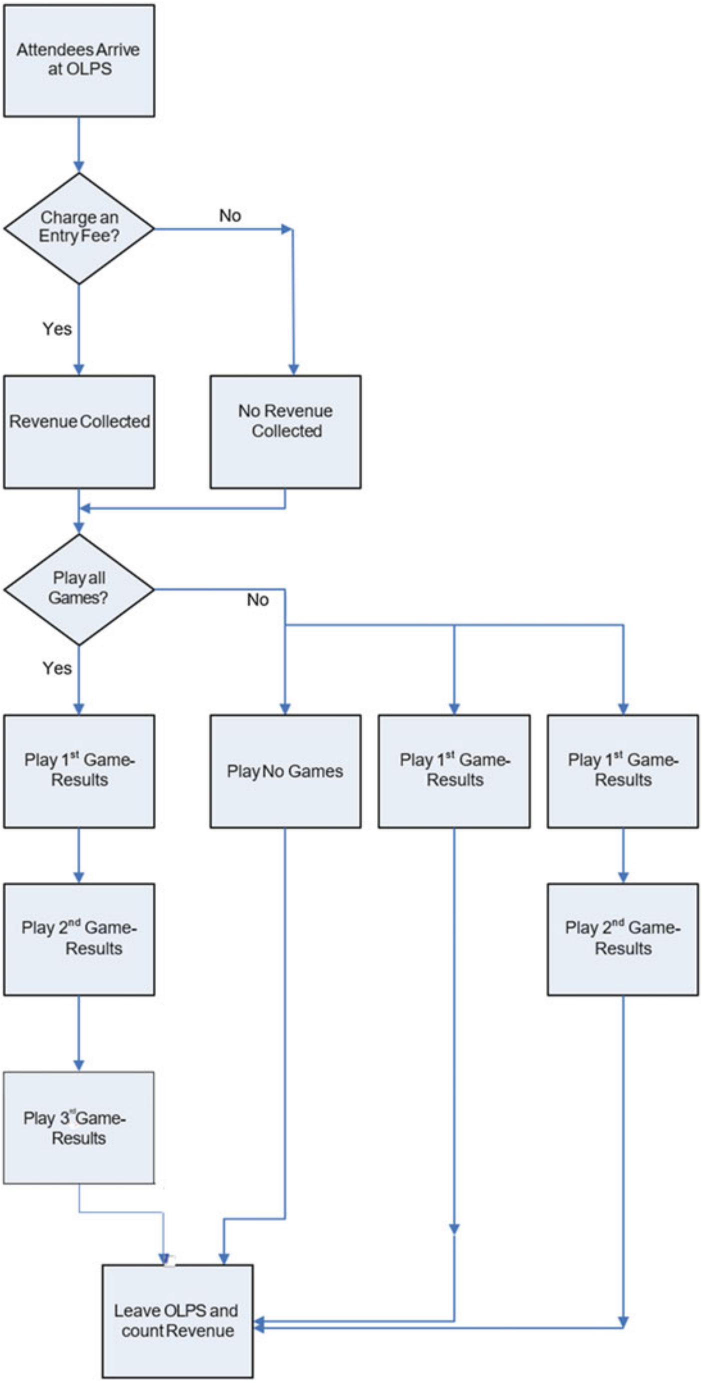


Fig. 7.4 Simple process flow planning and design model of OPLS

From this preliminary analysis, we can also learn where the risk related to uncertainty occurs. Fr. Efia can see that uncertainty is associated with a number of event processes: (1) the number of parishioners attending *Vegas Night at OLPS*, which is likely to be associated with weather conditions and the entry fee charged, and (2) the outcomes of the games (players winning or losing) which are associated with the odds that Fr. Efia and Voitech will set for the games. The question of setting the odds of the games is not included at this point but could be a part of the diagram. In this example, it is assumed that after these preliminary design issues are resolved we can return to the question of the game odds. The design process is usually iterative due to the complexity of the design task, so you may return to several of the resolved issues to investigate possible changes. Changes in one design issue can, and will, affect the design of other event elements. We will return to this problem later and see how we can incorporate uncertainty deterministically in an Excel based decision model.

7.4 Understanding the Important Elements of a Model

As we can see from the brief discussion of the OLPS event, understanding the processes and design of a model is not an easy task. In this section we will create a framework for building complex models. Let us begin by considering *why* we need models. First, we use models to help us analyze problems and eventually make decisions. If our modeling is accurate and thorough, we can greatly improve the quality of our decision making. As we determined earlier in our investment example, intuition is certainly a valuable personal trait, but one that may not be sufficient in *complex* and *risky* decision situations. So, what makes a problem complex? **Complexity** comes from:

1. the need to consider the interaction of many factors
2. the difficulty in understanding the nature and structure of the interactions
3. the uncertainty associated with problem variables and structure
4. the potentially evolving and changing nature of a problem.

To deal with complexity, we need to develop a formal approach to the modeling process; that is, how we will organize our efforts for the most effective and efficient modeling. This does not guarantee success in understanding complex models, but it contributes mightily to the possibility of a *better* understanding. It is also important to realize that the modeling process occurs in stages, and that one iteration through the modeling process may not be sufficient for completely specifying a complex problem. It may take several iterations with progressively more complex modeling approaches to finally arrive at an understanding of our problem. This will become evident as we proceed through the OLPS example.

So, let us take what we have learned thus far and organize the steps that we need to follow to perform effective and efficient modeling:

1. A **pre-modeling or design phase** that contributes to our preliminary understanding of the problem. This can also be called the *problem definition* phase. This step can take a considerable proportion of the entire modeling effort. After all, if you define the problem poorly, no amount of clever analysis will be helpful. At this stage the goal of the modeling effort should be made clear. What are we expecting from the model? What questions will it help answer? How will it be used and by whom?
2. A **modeling phase** where we build and implement a model that emerges from the pre-modeling phase. Here we refine our specification of the problem sufficiently to explore the model's behavior. At this point the model will have to be populated with very specific detail.
3. An **analysis phase** where we test the behavior of the model developed in steps (1) and (2), and we analyze the results. We collect data in this phase that the model produces under controlled experimental conditions, and then we analyze the results.
4. A **final acceptance phase** where we reconsider the model specification if the result of the analysis phase suggests the need to do so. At this point, we can return to the earlier phases until the decision maker achieves desired results. It is, of course, also possible to conclude that the desired results are *not* achievable.

7.4.1 Pre-modeling or Design Phase

In the pre-modeling or design phase, it is likely that we have not settled on a precise definition of our problem, just as Fr. Efia has not decided on the detailed design of his event. I refer to this step as the pre-modeling phase, since the modeling is generally done on paper and does not involve the use of a computer-based model. Fr. Efia will use this phase to make decisions about the activities that he will incorporate into *Vegas Night at OLPS*; thus, as we stated earlier, he is still defining the event's design. Voitech used the interview exercise in Table 7.1 to begin this phase. The resulting actions of Table 7.1 then led to the preliminary process flow design in Fig. 7.4. If the problem is already well defined, this phase may not be necessary. But, often the problem definition is not easy to determine without considerable effort. It is not unusual for this step to be the longest phase of the modeling process. And why not, there is nothing worse than realizing that you have developed an elegant model that solves the *wrong* problem.

7.4.2 Modeling Phase

Now, it is time to begin the second phase—*modeling*. At this point, Fr. Efia has decided on the basic structure of the events—the games to be played and their odds, the restrictions, if any, on the number of games played by attendees, whether an entry fee will be required, etc. A good place to begin the modeling phase is to create an **Influence Diagram** (IFD). IFDs are diagrams that are connected by directed arrows, much like those in the preliminary process flow planning diagram of Fig. 7.4. An IFD is a powerful tool that is used by decision analysts to specify *influences* in decision models. Though the concept of an IFD is relatively simple, the theoretical underpinnings can be complicated. For our example, we will develop two types of IFDs: one very simple and one considerably more complex.

We begin by identifying the factors—processes, decisions, outcomes of decisions, etc.—that constitute the problem. In our first IFD, we will consider the links between these factors and determine the type of influence between them, either positive (+) or negative (−). A **positive influence** (+) suggests that if there is an increase in a factor, the factor that it influences also has an increase; it is also true that as a factor decreases so does the factor it influences. Thus, they move in the same direction. For example, if I increase marketing efforts for a product, we can expect that sales will also increase. This suggests that we have a positive influence between marketing efforts and sales. The opposite is true for a **negative influence** (−): factors move in opposite directions. A negative influence can easily exist between the quality of employee training and employee errors—the higher the quality of training for employees the lower the number of errors committed by employees. Note that the IFD does not suggest the intensity of the influence, only the direction.

Not all models lend themselves to this simple form of IFD, but there will be many cases where this approach is quite useful. Now, let's apply the IFD to Fr. Efia's problem. Voitech has helped Fr. Efia to create a simple IFD of revenue generation for *Vegas Night at OLPS*. It is shown in Fig. 7.5. Voitech does so by conducting another interview and having Fr. Efia consider more carefully the structure of the event, and how elements of the event are related. To understand the movement of one factor due to another, we first must establish a scale for each factor, from negative to positive. The negative to positive scale used for the levels of *weather quality* and *attendee good luck* is *bad to good*. For *attendance* and *revenue*, the scale is quite direct: higher levels of attendance or revenue are positive and lower levels are negative. The IFD in Fig. 7.5 provides an important view of how revenues are generated, which of course is the goal of the event. Fr. Efia has specified six important factors: Weather Quality, Attendance, Attendee Luck or Fortune in Gambling, Entry Admission Revenue, Gambling Proceeds Revenue, and Total Revenue.

Some factors are uncertain, and others are not. For example, weather and attendee fortune are uncertain, and obviously he hopes that the weather quality will be good (+) and that attendee good fortune will be bad (−). The effect of these two conditions will eventually lead to greater revenues (+). Entry Admission Revenues are known with certainty once we know the attendance, as is the Total Revenue once we determine Entry Admission Revenue and Gambling Proceeds Revenue.

Note that the model is still quite general, but it does provide a clear understanding of the factors that will lead to either success or failure for the OLPS event. There is no final commitment, yet, to several the important questions in Table 7.1; for example, questions 2—*odds of the games*, and question 3—*will all attendees play all games*. But, it has been decided that the games mentioned in question 2 will all be

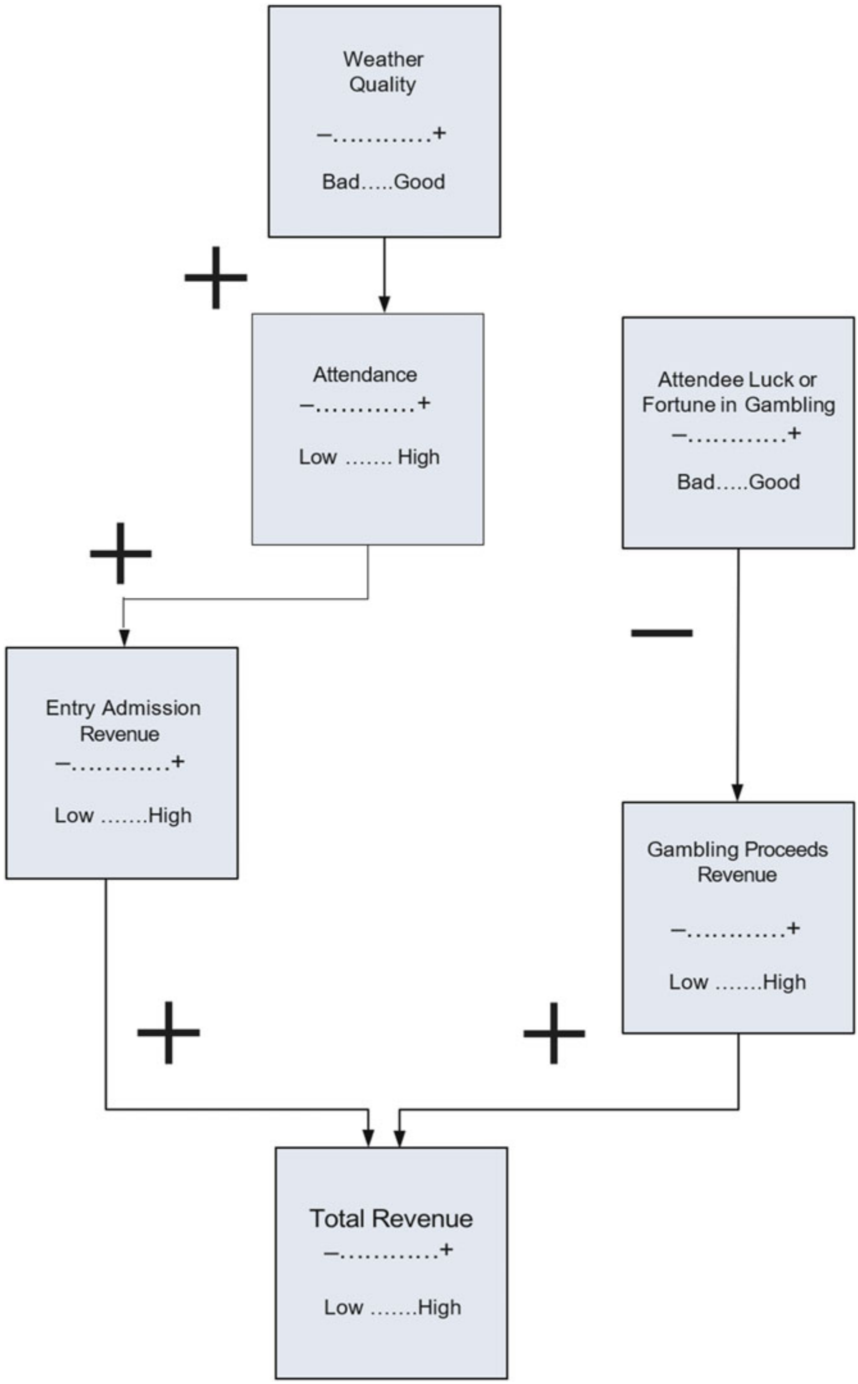


Fig. 7.5 Simple revenue generation influence diagram of OLPS

a part of the event, and that an entry admission fee *will* be charged. The admission fee will supplement the revenues generated by the games. This could be important given that if the games generate a loss for OLPS, then entry admission revenues could offset them. Since Fr. Efia can also control the odds of these games, he eventually needs to consider how the odds will be set.

So, in summary, what does the IFD tell us about our problem? If Fr. Efia wants the event to result in larger revenues, he now knows that he will want the following conditions:

1. Good weather to attract a higher attendance
   (a) we have little control of the weather
   (b) we can schedule events in time periods where the likelihood of good weather is more likely
   (c) the exact effect of weather on attendance is uncertain
2. Poor attendee luck leading to high gambling revenues
   (a) we do have control of an attendee's *luck* by setting the odds of games
   (b) a fair game has 50–50 odds, and an example of a game favoring OLPS is 60–40 odds—60% of the time OLPS wins and 40% the attendee wins
   (c) the odds in favor of OLPS can be set be higher, depending on what attendees will tolerate
3. Charge an entry admission to supplement gambling revenue
   (a) entry fee is a guaranteed form of revenue based on attendance, unlike the gambling revenue which is uncertain
   (b) entry fee is also a delicate matter: charging too much might diminish attendance, and it may also appear to be a direct request for funds, that the Archbishop is firmly against
   (c) the entry fee can be justified as a fee for food and refreshments

As you can see, this analysis is leading to a formal design of the event (a formal problem definition). Just what will the event look like? At this point in the design, Voitech has skillfully directed Fr. Efia to consider all the important issues related to revenue generation. Fr. Efia must make some difficult choices at this point if he is going to eventually create a model of the event. Note that he will still be able to change his choices during testing the model, but he does need to settle on a particular event configuration to proceed with the modeling effort. Thus, we are moving toward the final stages of phase 2, the *modeling phase*.

Voitech is concerned that he must get Fr. Efia to make some definitive choices for specifying the model. Both men meet at a local coffee shop, and after considerable discussion, Voitech determines the following final details for *Vegas Night at OLPS*:

1. There will be an entry fee of $10 for all those attending. He feels this is a reasonable charge that will not deter attendance. Given the array of wonderful ethnic foods that will be provided by parishioners, this is quite a modest charge

for entry. Additionally, he feels that weather conditions are the most important determinant for attendance.

2. The date for the event is scheduled for October 6th. He has weather information forecasting the weather conditions for that October date: 15% chance of rain, 40% chance of clouds, and a 45% chance of sunshine. Note that these weather outcomes are **mutually exclusive** and **collectively exhaustive**. They are mutually exclusive in that there is *no overlap* in events; that is, it is either rainy or cloudy or sunny. They are collectively exhaustive in that the sum of their probabilities of occurrence is equal to 1; that is, these are *all* the outcomes that can occur.
3. Since weather determines attendance, Voitech interviews Fr. Efia with the intent to determine his estimates for attendance given the various weather conditions. Based on his previous experience with parish events, Fr. Efia believes that if weather is *rainy*, attendance will be 1500 people; if it is *cloudy*, attendance is 2500; if the weather is *sunshine*, attendance is 4000. Of course, these are subjective estimates, but he feels confident that they closely represent likely attendance.
4. The selection of the games remains the same—Omnipotent Two-Sided Die (O2SD), Wheel of Outrageous Destiny (WOD), and the Bowl of Treachery (BT). To simplify the process, and to comply with Fr. Efia's wishes to limit gambling (recall he does not approve of gambling), he will insist that every attendee must play all three games and play them *only* once. Later he may consider relaxing this condition to permit individuals to do as they please—play all, some, none of the games, and to possibly repeat games. This relaxation of play will cause Voitech to model a much more complex event by adding another factor of uncertainty: the unknown number and type of games each attendee will play.
5. He also has set the odds of attendees winning at the games as follows: probabilities of winning in O2SD, WOD, and BT, are 20, 35, and 55%, respectively. The structure of the games is quite simple. If an attendee wins, Fr. Efia gives the attendee the value of the bet (more on this in 6.); if the attendee loses, then the attendee gives Fr. Efia the value of the bet. The logic behind having a single game (BT with 55%) that favors the attendees is to avoid having attendees feel as if they are being exploited. He may want to later adjust these odds to determine the sensitivity of gambling revenues to the changes.
6. All bets at all games are $50 bets, but he would also like to consider the possible outcomes of other quantities, for example $100 bets. This may sound like a substantial amount of money, but he believes that the very affluent attendees will not be sensitive to these levels of bets.

7.4.3 Resolution of Weather and Related Attendance

Now that the *Vegas Night at OLPS* is precisely specified, we can begin to model the behavior of the event. To do so, let us first use another form of influence diagram, one that considers the *uncertain events* associated with a process. This diagramming approach is unlike our initial efforts in Fig. 7.5, and it is quite useful for identifying the complexities of uncertain outcomes. One of the advantages of this approach is its simplicity. Only two symbols are necessary to diagram a process: a rectangle and a circle. The rectangle represents a step or decision in the process, e.g. the arrival of attendees or the accumulation of revenue. The circle represents an *uncertain* event and the outputs of the circle are the anticipated results of the event. These are the symbols that are also used in **decision trees**, but our use of the symbols is slightly different from those of decision trees. Rectangles in decision trees represent decisions, actions, or strategies. In our use of these symbols, we will allow rectangles to also represent some state or condition; for example, the collection of entry fee revenue, or the occurrence of some weather condition, like rain. Figure 7.6 shows the model for this new IFD modeling approach.

In Fig. 7.6, the flow of the IFD proceeds from top to bottom. The first event that occurs in our problem is the resolution of the uncertainty related to weather. How does this happen? Imagine that Fr. Efia awakens early on October 6th and looks out his kitchen window. He notices the weather for the day. Then he assumes that the weather he has observed will persist for the entire day. All of this is embodied in the circle marked *Weather Condition* and the resulting arrows. The three arrows represent the possible resolution of *weather condition* uncertainty, each of which leads to an assumed, deterministic number of participants. In turn, this leads to a corresponding entry fee revenue varying from a low of $15,000 to a high of $40,000. For example, suppose Fr. Efia observes *sunshine* out of his kitchen window. Thus, *weather condition* uncertainty is resolved, and 4000 attendees are expected to attend *Vegas Night at OLPS*, resulting in $40,000 in entry fees.

7.4.4 Attendees Play Games of Chance

Next, the number of attendees determined earlier will participate in each of the three games. The attendees either win or lose in each game; an attendee win is bad news for OLPS, and a loss is good news for OLPS. Rather than concerning ourselves with the outcome of each individual attendee's gaming results, an *expected* outcome of revenues can be determined for each game and for each weather/attendee situation. An **expected value** in decision analysis has a special meaning. Consider an individual playing the WOD. On each play the player has a 35% chance of winning. Thus, the average or expected winnings on any single play are $17.50 ($50 ∗ 0.35) and the losses are $32.50 ($50 ∗ [1–0.35]). Of course, we know that an attendee either wins or loses and that the outcomes are either $50 or $0. The expected values represent a weighted average: outcomes weighted by the probability of winning or losing. Thus, if a player plays WOD 100 times, the player can *expect* to win $1750 (100 ∗ $17.50) and Fr. Efia can *expect* to collect $3250 (100 ∗ 32.50). The expected values should be relatively accurate measures of *long* term results, especially given the large

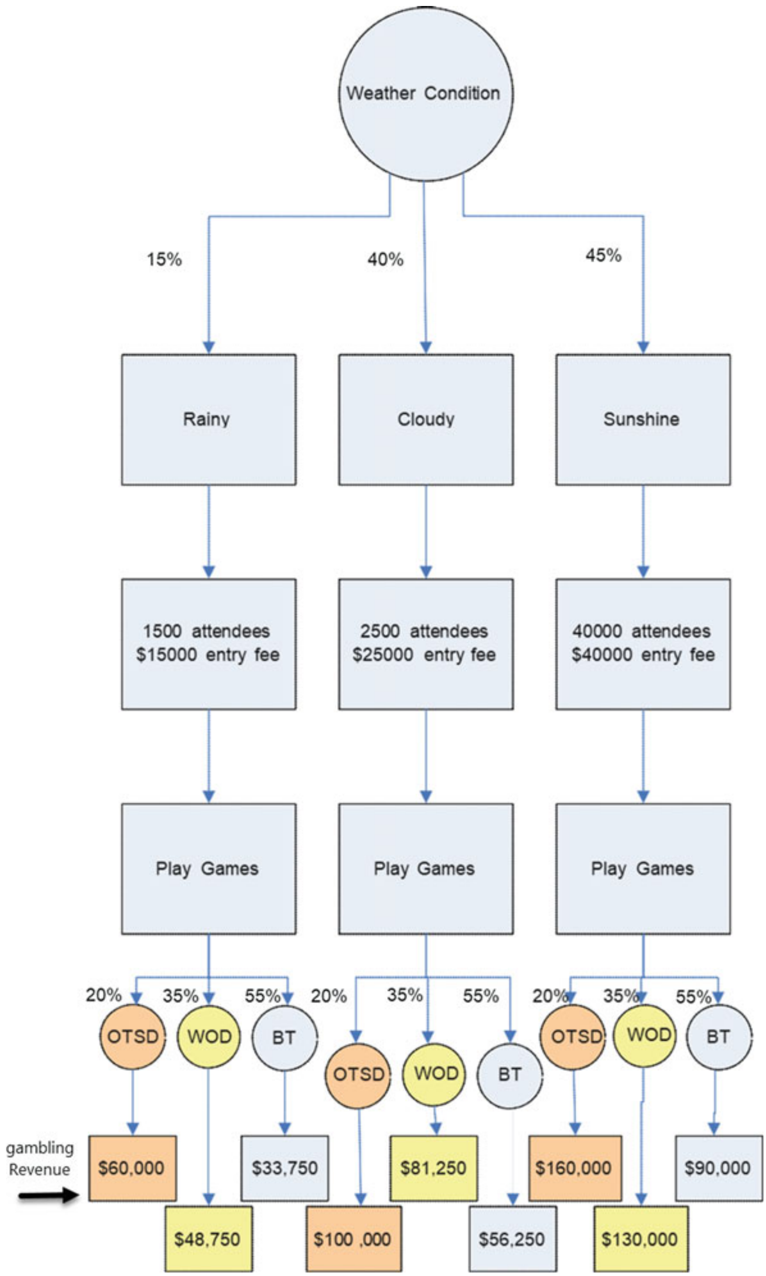


Fig. 7.6 IFD for Fr. Efia's final event configuration

quantity of attendees, and this should permit the averages for winning (or losing) to be relatively close to odds set by Fr. Efia.

At this point we have converted some portions of our probabilistic model into a deterministic model; the probabilistic nature of the problem has not been abandoned, but it has been modified to permit the use of a deterministic model. The weather remains probabilistic because we have a distribution of probabilities and outcomes that specify weather behavior. The outcomes of attendee gambling also have become deterministic. To have a truly probabilistic model we would simulate the outcome of every play for every player. We have chosen not to simulate each uncertain event, but rather, to rely on what we *expect* to happen as determined by a weighted average of outcomes. Imagine the difficulty of simulating the specific fortune, or misfortune, of each game for each of the thousands of attendees.

These assumptions simplify our problem greatly. We can see in Fig. 7.6 that the gambling revenue results vary from a low of $33,750[1] for the BT in rainy weather to a high of $160,000 for OTSD in sunshine. The range of total revenue (entry fee and gambling revenue for a given weather condition) varies from a low of $157,500[2] for rainy weather and a high of $420,000[3] for sunshine.

7.4.5 Fr. Efia's What-if Questions

Despite having specified the model quite clearly to Voitech, Fr. Efia is still interested in asking numerous what-if questions. He feels secure in the basic structure of the games, but there are some questions that remain, and they may lead to adjustments that enhance the event's revenue generation. For example, what if the entry fee is raised to $15, $20, or even $50? What if the value of each bet is changed from $50 to $100? What if the odds of the games are changed to be slightly different from the current values? These are all important questions because if the event generates too little revenue it may cause serious problems with the Archbishop. On the other hand, the Archbishop has also made it clear that the event should not take advantage of the parishioners. Thus, Fr. Efia is walking a fine line between too little revenue and too much revenue. Fr. Efia's what-if questions should provide insight on *how* fine that revenue line might be.

Finally, Voitech and Fr. Efia return to the goals they originally set for the model. The model should help Fr. Efia determine the revenues he can expect. Given the results of the model analysis, it will be up to him to determine if revenues are too low to halt the event or too high and attract the anger of the Archbishop. The model also should allow him to experiment with different revenue generating conditions. This

[1] 1500 ∗ $50 ∗ (1–0.55) = $33,750……and……4000 ∗ $50 ∗ (1–0.20) = $160,000.

[2] $60,000 + $48,750 + $33,750 + $15,000 = $157,500 (game revenue plus attendance fee).

[3] $160,000 + $130,000 + $90,000 + $40,000 = $420,000 (game revenue plus attendance fee).

important use of the model must be considered as we proceed to model building. Fortunately, there is a technique that allows us to examine the questions Fr. Efia faces. The technique is known as **sensitivity analysis**. The name might conjure an image of a psychological analysis that measures an individual's emotional response to some stimuli. This image is in fact quite similar to what we would like to accomplish with our model. Sensitivity analysis examines how sensitive the model *output* (revenues) is to changes in the model *inputs* (odds, bets, attendees, etc.). For example, if I change the entry fee, how will the revenue generated by the model change; how will gambling revenues change if the attendee winning odds of the WOD are changed to 30% from the original 35%? One of these changes could contribute to revenue to a greater degree than the other—hence the term *sensitivity* analysis. Through sensitivity analysis, Fr. Efia can direct his efforts toward those changes that make the greatest difference in revenue.

7.4.6 Summary of OLPS Modeling Effort

Before we proceed, let us step back for a moment and consider what we have done thus far and what is yet to be done in our study of modeling:

- *Model categorization*—we began by defining and characterizing models as deterministic or probabilistic. By understanding the type of model circumstances we are facing, we can determine the best approach for modeling and analysis.
- *Problem/Model definition*—we introduced several *paper modeling* techniques that allow us to refine our understanding of the problem or problem design. Among these were process flow diagrams that describe the logical steps that are contained in a process, Influence Diagrams (IFD) that depict the influence of and linkage between model elements, and even simple interview methods to probe the understanding of issues and problems related to problem definition and design.
- *Model building*—the model building phase has not been described yet, but it includes the activities that transform the paper models, diagrams, and results of the interview process into Excel based functions and relationships.
- *Sensitivity analysis*—this provides the opportunity to ask what-if questions of the model. These questions translate into input parameter changes in the model and the resulting changes in outputs. They also allow us to focus on parameter changes that have a significant effect on model output.
- *Implementation of model*—once we have studied the model carefully we can make decisions related to execution and implementation. We may decide to make changes to the problem or the model that fit with our changing expectations and goals. As the modeling process advances, we may gain insights into new questions and concerns, heretofore not considered.

7.5 Model Building with Excel

In this section, we will finally convert the efforts of Voitech and Fr. Efia into an Excel based model. Excel will serve as the programming platform for model implementation. All their work, thus far, has been aimed at conceptualizing the problem design and understanding the relationships between problem elements. Now, it is time to begin translating the model into an Excel workbook. Figure 7.6 is the model we will use to guide our modeling efforts. The structure of the IFD in Fig. 7.6 lends itself quite nicely to an Excel based model. We will build the model with several requirements in mind. Clearly, it should permit Fr. Efia flexibility in revenue analysis; to be more precise, one that permits sensitivity analysis. Additionally, we want to use what we have learned in earlier chapters to help Fr. Efia understand the congruency of his decisions and the goals he has for *Vegas Night at OLPS*. In other words, the model should be user friendly and useful for those decisions relating to his eventual implementation of *Vegas Night at OLPS*.

Let's examine Fig. 7.6 and determine what functions will be used in the model. Aside from the standard algebraic mathematical functions, there appears to be little need for highly complex functions. But, there are numerous opportunities in the analysis to use functions that we have not used or discussed before, for example control buttons that can be added to the **Quick Access Toolbar** menu via the Excel Options Customize tool menu—**Scroll Bars**, **Spinners**, **Combo Boxes**, **Option Buttons**, etc. Alternatively, we can find this new functionality in the **Developer Menu**[4] of the ribbon. We will see later that these buttons are a very convenient way to provide users with control of spreadsheet parameter values, such as attendee entry fee and the value of a bet. Thus, they will be useful in sensitivity analysis.

So how will we begin to construct our workbook? The process steps shown in Fig. 7.6 represent a convenient layout for our spreadsheet model. It also makes good sense that spreadsheets should flow either left-to-right or top-to-bottom in a manner consistent with process steps. I propose that left-to-right is a useful orientation, and that we should follow all our Feng Shui inspired best practices for workbook construction. The uncertain weather conditions will be dealt with deterministically, so the model will provide Fr. Efia outcomes for the event *given* one of the three weather conditions: rainy, cloudy, or sunshine. In other words, the model will not generate a weather event: a weather event will be *assumed*, and the results of that event can then be analyzed. The uncertainty associated with the games also will be handled deterministically through the use of *expected values*. We will assume that precisely 20% of the attendees playing OTSD will win, 35% of the attendees playing WOD will win, and 55% of those playing BT will win. These probabilistic winning percentages will rarely be exactly, 20, 35, and 55% in reality, but if there are many attendees, the percentages should be close to these values.

Figure 7.7 shows the layout of the model. For the sake of simplicity, I have placed all analytical elements—brain, calculations, and sensitivity analysis on a single

[4]Go to... File/Options/Customize Ribbon/Developer.

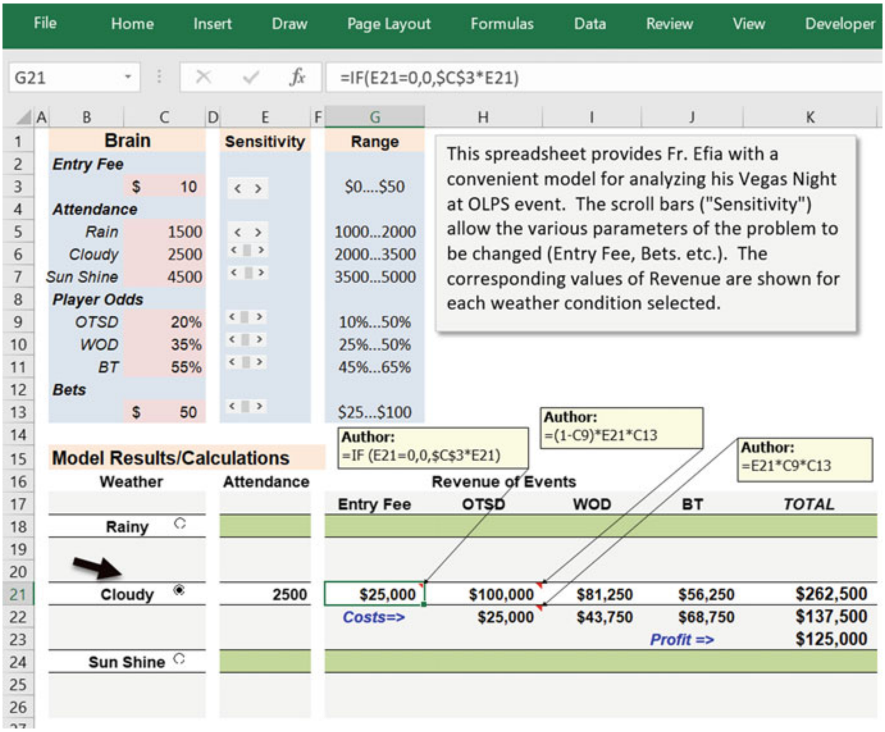


Fig. 7.7 Spreadsheet model for Vegas night at OPLS

worksheet. If the problem were larger and more complex, it probably would be necessary to place each major part of the model on a separate worksheet. We will discuss aspects of the spreadsheet model in the following order: (1) the basic model and its calculations, (2) the sensitivity analysis that can be performed on the model, and (3) the controls that have been used in the spreadsheet model (scroll bars and options buttons) for user ease of control.

7.5.1 Basic Model

Let us begin by examining the general layout of Fig. 7.7. The *Brain* is contained in the range B1 to C13. The *Brain* for our spreadsheet model contains the values that will be used in the analysis: Entry Fee, Attendance, Player (odds), and Bets. Note that besides the nominal values that Fr. Efia has agreed upon, on the right there is a *Range* of values that provide an opportunity to examine how the model revenue varies with changes in nominal values. These ranges come from Voitech's discussion with Fr. Efia regarding his interest in the model's sensitivity to change. The values currently available for calculations are in column C and they are referenced in the Model Results/Calculations section. The values in cells marked *Range* are text entries that are not meant as direct input data. They appear in G1:G13. As changes are made to the nominal values, they will appear in column C. Later I will describe how the scroll bars, in column E, are used to control the level of the parameter input without having to key-in new values, much like you would use the volume control scroll bars to set the volume of your radio or stereo.

The bottom section of the spreadsheet is used for calculations. The Model Results/Calculations area is straight forward. A weather condition (one of three) is selected by depressing the corresponding Option Button—the circular button next to the weather condition, which contains a black dot when activated. Later, we will discuss the operation of option buttons, but for now, it is sufficient to say that these buttons, when grouped together, result in a unique number to be placed in a cell. If there are 3 buttons grouped, the numbers will range from 1 to 3, each number representing a button. This provides a method for a specific condition to be used in calculation: Rainy = 1, Cloudy = 2, and Sunshine = 3. Only one button can be depressed at a time, and the Cloudy condition, row 21, is the current condition selected. All this occurs in the area entitled Weather.

Once the Weather is selected, the Attendance is known given the relationship Fr. Efia has assumed for Weather and Attendance. Note that the number in the Attendance cell, E21 of Fig. 7.7, is 2500. This is the number of attendees for a Cloudy day. As you might expect, this is accomplished with a logical *IF* function, and is generally determined by the following logical *IF* conditions: *IF* value of button = 1 then 1500, else *IF* value of button = 2 then 2500, else 4500. Had we selected the Rainy option button then the value for attendees, cell E18, would be 1500. As stated earlier, we will see later how the buttons are created and controlled.

Next, the number of attendees is translated into an Entry Fee revenue (C3∗E21 = 10∗$2500 = $25,000) in cell G21. The various game revenues also are determined from the number of attendees. For example, OTSD revenue is the product of the number of attendees in cell E21 (2500), the value of each bet in cell C13 ($50), and the probability of an OLPS win (1–C9 = 0.80), which results in $100,000 (2500 ∗ $50 ∗ 0.80) in cell H21. The calculations for WOB and BT are $81,250[5] and 56,250.[6]

Of course, there are also payouts to the players that win, and these are shown as *Costs* on the line below revenues. Each game will have payouts, either to OLPS or the players, which when summed equal the total amount of money that is bet. In the case of Cloudy, each game has total bets of $125,000 ($50 ∗ 2500). You can see that if you combine the revenue and cost for each game, the sum is indeed $125,000, the total amount bet for each game. As you would expect, the only game where the costs (attendee’s winnings) are greater than the revenues (OLPS’s winnings) is the BT game. This game has odds that favor the attendee. The cumulative profit for the event

[5] 2500 ∗ $50 ∗ (1–0.35) = $81,250.

[6] 2500 ∗ $50 ∗ (1–0.55) = $56,250.

is the difference between the revenue earned by OLPS in cell K21 ($262,500) and the costs incurred in cell K22 ($137,500). In this case, the event yields a profit of $125,000 in cell K23. This represents the combination of *Entry Fee*, $25,000, and *Profit* from the games, $100,000.[7]

The model in Fig. 7.7 represents the basic layout for problem analysis. It utilizes the values for entry fees, attendance, player odds, and bets that were agreed to by Voitech and Fr. Efia. In the next section we address the issues of sensitivity analysis that Fr. Efia has raised.

7.5.2 *Sensitivity Analysis*

Once the basic layout of the model is complete, we can begin to explore some of the what-if questions that were asked by Fr. Efia. For example, what change in revenue occurs if we increase the value of a bet to $100? Obviously, if all other factors remain the same, revenue will increase. But will all factors remain the same? It is conceivable that a bet of $100 will dissuade some of the attendees from attending the event; after all, this doubles an attendee's total exposure to losses from $150 (three games at $50 per game) to $300 (three games at $100 per game). What percentage of attendees might not attend? Are there some attendees that would be more likely to attend if the bet increases? Will the Archbishop be angry when he finds out that the value of a bet is so high? The answers to these questions are difficult to know. The current model will not provide information on how attendees will respond since there is no economic model included to gauge the attendee's sensitivity to the value of the bet, but Fr. Efia can posit a guess as to what will happen with attendance and personally gauge the Archbishop's response. Regardless, with this model Fr. Efia can begin to explore the effects of these changes.

We begin sensitivity analysis by considering the question that Fr. Efia has been struggling with—how to balance the event revenues to avoid the attention of the Archbishop. If he places the odds of the games greatly in favor of OLPS, the Archbishop may not sanction the event. As an alternative strategy to setting *poor* player odds, he is considering increasing the entry fee to offset losses from the current player odds. He believes he can defend this approach to the Archbishop, especially considering the *all-you-can-eat* format for ethnic foods that will be available to attendees. But of course, there are limits to the entry fee that OLPS alumni will accept as reasonable. Certainly, a fee of $10 can be considered very modest for the opportunity to feast on at least 15 varieties of ethnic food.

So what questions might Fr. Efia pose? One obvious question is—How will an increase in the entry fee offset an improvement in player odds? Can an entry fee increase make up for lower game revenues? Finally, what Entry Fee increase will offset a change to fair odds: 50–50 for bettors and OLPS? Let us consider the Cloudy

[7]($100,000–$25,000) + ($81,250–$43,750) + ($56,250–$68,750) = $100,000.

scenario in Fig. 7.7 for our analysis. In this scenario total game revenue is $262,500 and cost is $137,500, resulting in overall profit of $125,000. Clearly, the entry fee will have to be raised to offset the lost gaming revenue if we improve the attendee's winning odds.

If we set the gaming odds to fair odds (50–50), we expect that the distribution of game funds to OLPS and attendees will be exactly the same, since the odds are now fair. Note that the odds have been set to 50% in cells C9, C10, and C11 in Fig. 7.8. Thus, the *only* benefit to Fr. Efia is Entry Fee, which is $25,000 as shown in cell K23. The fair odds scenario has resulted in a $100,000 profit reduction. Now, let us increase the Entry Fee to raise the level of profit to the desired $125,000. To achieve such a profit, we will have to set our Entry Fee to a significantly higher value. Figure 7.9 shows this result in cell K23. An increase to $50 per person in cell C3 eventually achieves the desired result. Although this analysis may seem trivial since the fair odds simply mean that profit will only be achieved through Entry Fee, in more complex problems the results of the analysis need not be as simple.

What if $50 is just too large a fee for Fr. Efia or the Archbishop? Is there some practical combination of a larger (greater than $10), but reasonable, Entry Fee, and some *nearly* fair odds that will result in $125,000? Figure 7.10 shows that if we set odds cell C9 to 40%, C10 to 40%, and C11 to 50%, for OTSD, WOD, and BT,

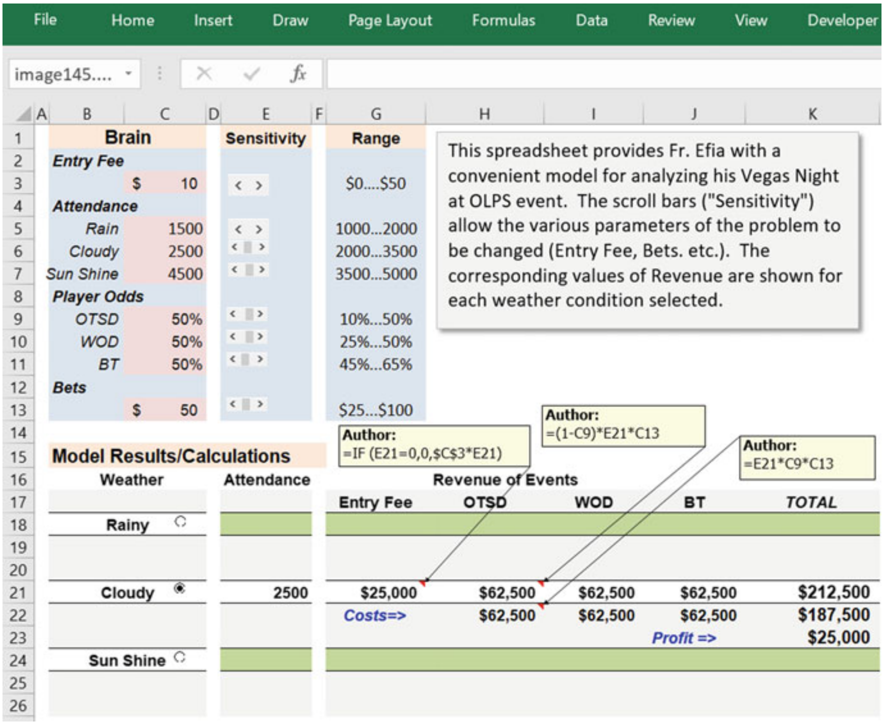


Fig. 7.8 Fair (50–50) odds for OPLS and attendees

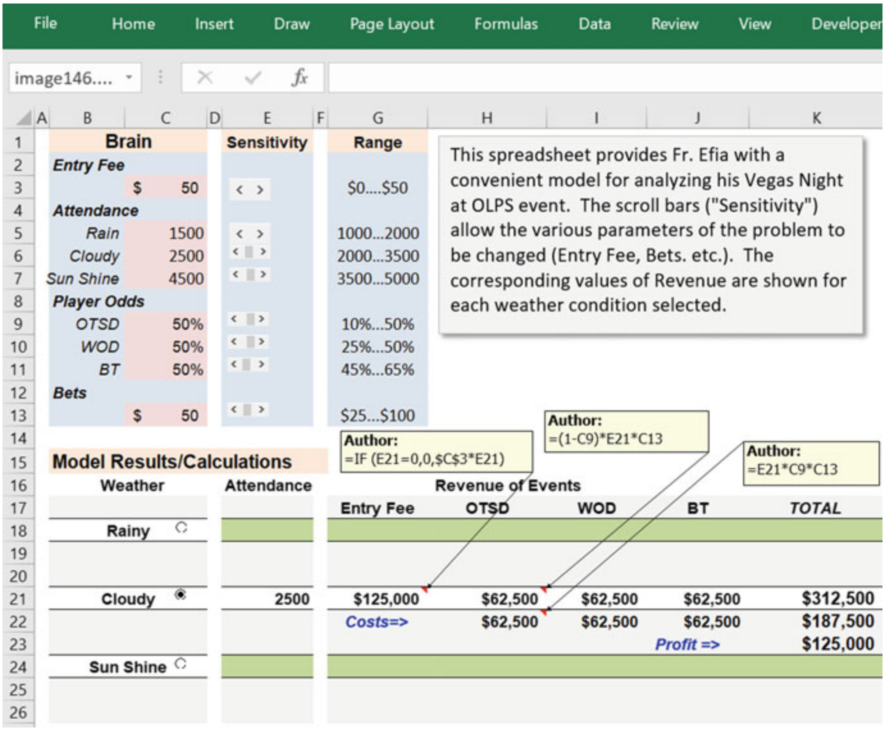


Fig. 7.9 Entry fee needed to achieve $125,000 with fair odds

respectively, and we also set the Entry Fee cell B3 to $30, we can achieve the desired results of $125,000 in cell K23. You can see that in this case the analysis is not as simple as before. There are many, many combinations of odds for three games that will result in the same profit. This set of conditions may be quite reasonable for all parties, but if they are not, then we can return to the spreadsheet model to explore other combinations.

As you can see, the spreadsheet model provides a very convenient system for performing sensitivity analysis. There are many other questions that Fr. Efia could pose and study carefully by using the capabilities of his spreadsheet model. For example, he has not explored the possibility of also changing the cost of a bet from the current $50 to some higher value. In complex problems, there will be many changes in variables that will be possible. Sensitivity analysis provides a starting point for dealing with these difficult *what-if* questions. Once we have determined areas of interest, we can take a more focused look at specific changes we would like to consider. To study these specific changes, we can use the **DataTable** feature in the Data ribbon. It is located in the What-If Analysis sub-group in the Data Tools group. The Data Table function allows us to select a variable (or two) and find the corresponding change in formula results for a given set of input values of the variable(s). For example, suppose we would like to observe the changes in the formula

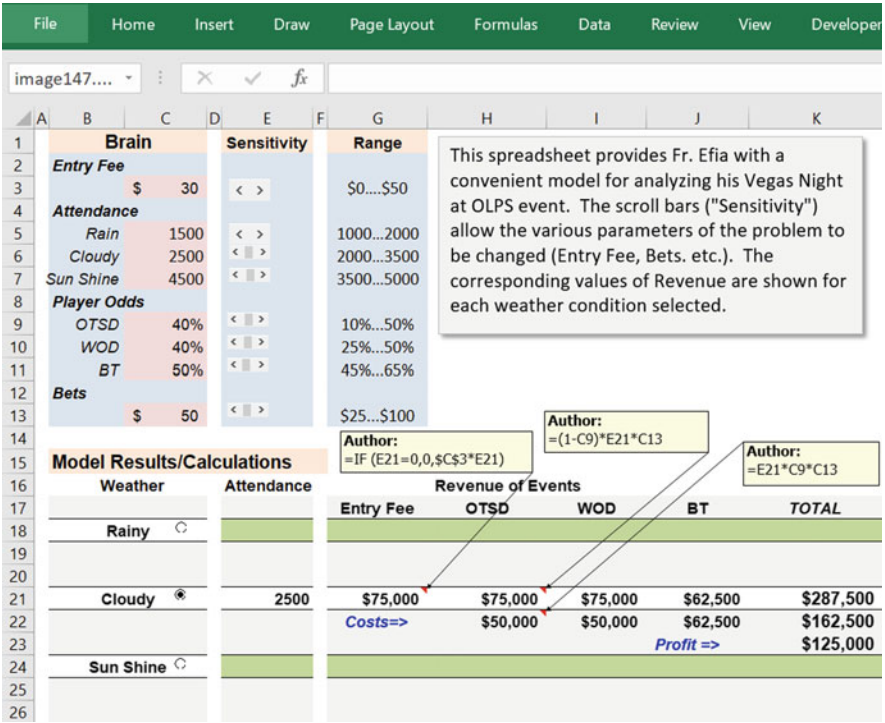
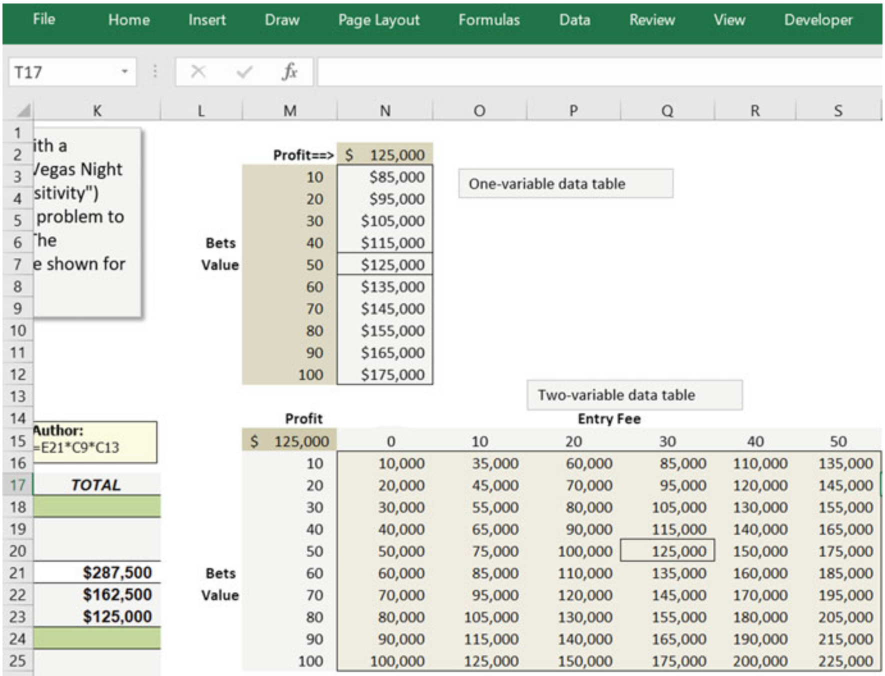


Fig. 7.10 Entry fee and less than fair odds to achieve $125,000

for *Profit* associated with our *Cloudy* scenario in Fig. 7.10, by changing the value of *Bets* in a range from $10 to $100. Figure 7.11 shows a single-variable Data Table based on $10 increment changes in *Bets*. Note that for $10 increments in Bets the corresponding change in *Profit* is $10,000. For example, the difference in the *Profit* when *Bet Value* is changed from $30 in cell N5 to $40 in cell N6 is $10,000 ($115,000–$105,000).

What about simultaneous changes in *Bets* and *Entry Fee*? Which of the changes will lead to greater increases in *Profit*? A two-variable data Table is also shown in Fig. 7.11, where values of *Bets* and *Entry Fee* are changed simultaneously. The cell with the rectangular border in the center of the table reflects the profit generated by the nominal values of *Bets* and *Entry Fee*, $50 and $30, respectively. From this analysis, it is clear that an increase in Bets, regardless of Entry Fee, will result in an increase of $10,000 in *Profits* for each $10 increment. We can see this by subtracting any two adjacent values in the column. For example, the difference between 125,000 in cell Q20 and 135,000 in cell Q21, for an *Entry Fee* of $30 and *Bets* of $50 and $60, is $10,000. Similarly, a $10 increment for *Entry Fee* results in a $25,000 increase in *Profits*, regardless of the level of the *Bet*. For example, 125,000 in cell Q20 and 150,000 in cell R20, for a *Bet* of $50 and *Entry Fee* of $30 and $40, is $25,000. This simple sensitivity analysis suggests that increasing *Entry Fee* is a more effective source of *Profit* than increasing *Bet*, for similar increments of change ($10).



Bets Value	Profit==> $ 125,000
10	$85,000
20	$95,000
30	$105,000
40	$115,000
50	$125,000
60	$135,000
70	$145,000
80	$155,000
90	$165,000
100	$175,000

Profit	Entry Fee					
Bets Value / $ 125,000	0	10	20	30	40	50
10	10,000	35,000	60,000	85,000	110,000	135,000
20	20,000	45,000	70,000	95,000	120,000	145,000
30	30,000	55,000	80,000	105,000	130,000	155,000
40	40,000	65,000	90,000	115,000	140,000	165,000
50	50,000	75,000	100,000	125,000	150,000	175,000
60	60,000	85,000	110,000	135,000	160,000	185,000
70	70,000	95,000	120,000	145,000	170,000	195,000
80	80,000	105,000	130,000	155,000	180,000	205,000
90	90,000	115,000	140,000	165,000	190,000	215,000
100	100,000	125,000	150,000	175,000	200,000	225,000

Fig. 7.11 One-variable and two-variable data table

Now, let us see how we create a single-variable and two-variable Data Table. The Data Table tool is a convenient method to construct a table for a particular cell formula calculation, and to record the formula's variation as one, or two, variables in the formula change. The process of creating a table begins by first selecting the formula that is to be used as table values. In our case, the formula is *Profit* calculation. In Fig. 7.12, we see the *Profit* formula reference, cell K23, placed in the cell N2. The *Bet Values* that will be used in the formula are entered in cells M3 through M12. This is a one variable table that will record the changes in *Profit* as *Bet Value* is varied from 10 to 100 in increments of 10. A one variable *Table* can take either a vertical or horizontal orientation. I have selected a vertical orientation that requires the *Bet Value* be placed in the column (M) immediately to the left of the formula column (N); for a horizontal orientation, the *Bet Value* would be in the row above the formula values. These are conventions that must be followed. The empty cells, N2:N12, will contain repetitive calculations of *Profit*. If a two-variable table is required, the variables are placed in the column to the left and the row above the calculated values. Also, the variables used in the formula must have a cell location in the worksheet that the formula references. In other words, the variable cannot be entered directly into the formula as a number, but must reference a cell location; for example, the cell references that are in the Brain worksheet.

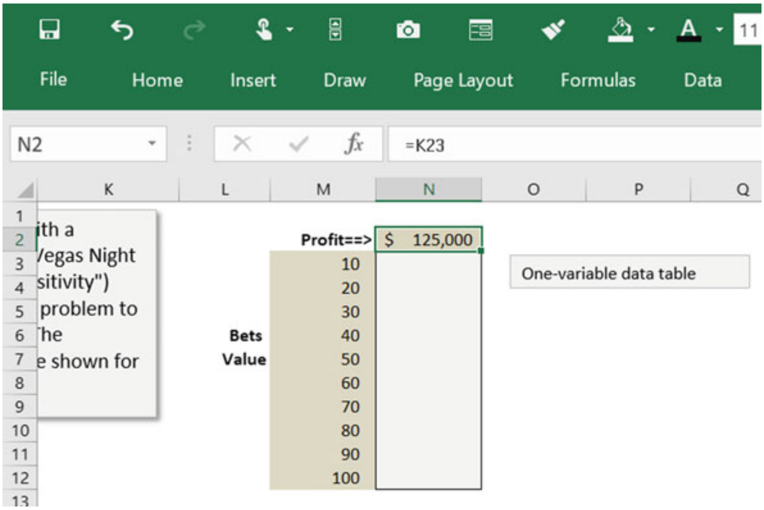


Fig. 7.12 One-variable data table

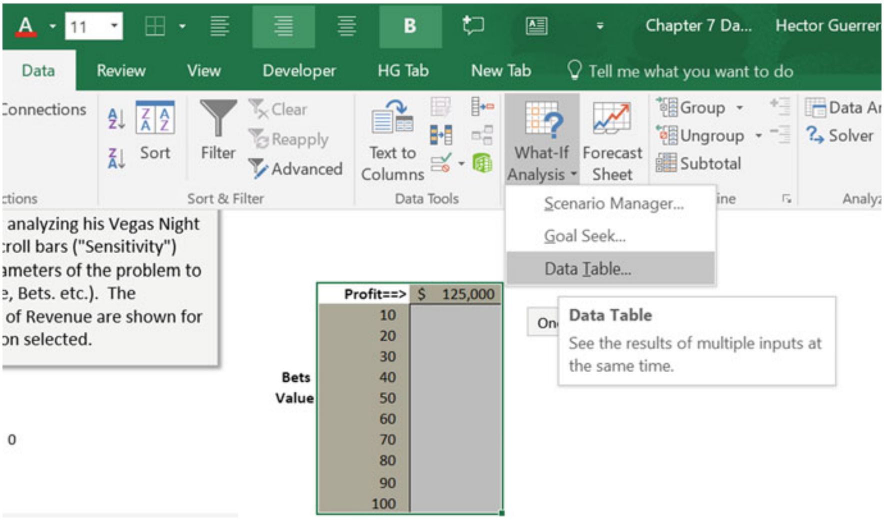


Fig. 7.13 Data table tool in data ribbon

Once the external structure of a table is constructed (the variable values and the formula cell) we select the table range: M2:N12. See Fig. 7.13. This step simply identifies the range that will contain the formula calculations and the value of the variable to use. Next, you will find in the Data ribbon the Data Table tool in the What-If Analysis. This step utilizes a wizard that requests the location of the Row input cell and Column input cell. See Fig. 7.14. In the case of the one variable table,

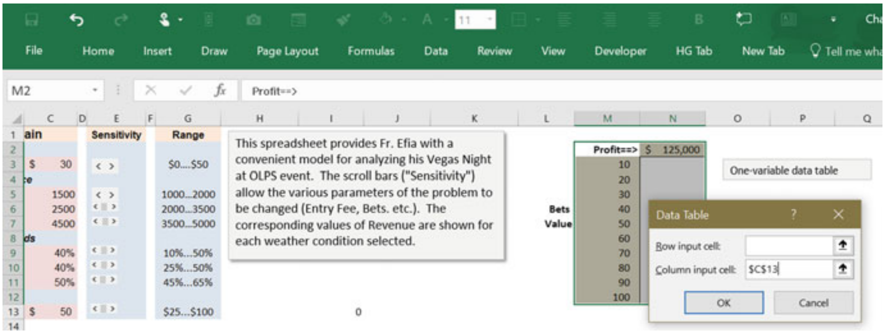


Fig. 7.14 Table wizard for one-variable data table

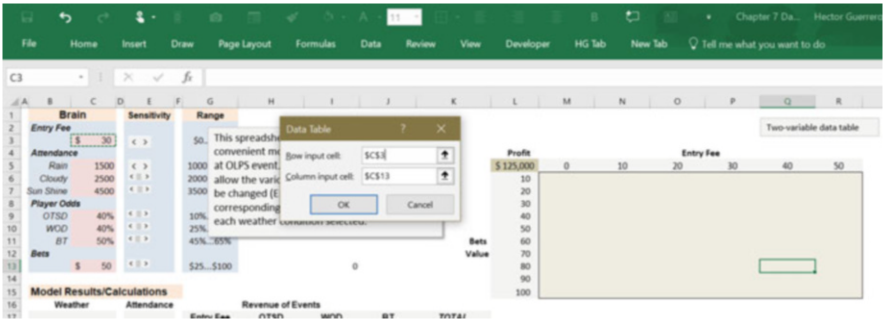


Fig. 7.15 Table wizard for two-variable data table

in vertical orientation, the relevant choice is the Column input cell because our variable values appear in *column* M. This wizard input identifies where the variable is located that will be used by the formula. For the one-variable table, the wizard input is cell C13; it has a current value of $50. The Data Table is being told where to make the changes in the formula. In Fig. 7.15, we see the two-variable Data Table with Row input cell as C3 and Column input cell as C13, the cell location of the formula input for *Entry Fee* and *Bets*, respectively. Results for the one-variable and the two-variable Data Table are shown in the previously introduced Fig. 7.11.

Of course, in more complex problems, the possible combination of variables for sensitivity analysis could be numerous. Even in our simple problem, there are 8 possible variables that we can examine individually or in combination (2, 3, 4, . . ., and 8 at a time); thus, there are literally thousands of possible sensitivity analysis scenarios we can study.

The spreadsheet model has permitted an in-depth analysis of Fr. Efia's event. It has met his initial goal of providing a model that allows him to analyze the revenues generated by the event. Additionally, he can ask a number of important what-if questions by varying individual values of model inputs. Finally, the formal use of sensitivity analysis through the Data Table tool provides a systematic approach to variable changes. All that is left is to examine some of the convenient devices that he has employed to control the model inputs—Scroll Bars and Option Buttons from the Forms Control.

7.5.3 Controls from the Forms Control Tools

Now, let us consider the devices that we have used for input control in Fr. Efia's spreadsheet model. These devices make analysis and collaboration with spreadsheets convenient and simple. We learned above that sensitivity analysis is one of the primary reasons we build spreadsheet models. In this section, we consider two simple tools that aid sensitivity analysis: (1) one to change a variable through incremental change control, and (2) a switching device to select a particular model condition or input. Why do we need such devices? Variable changes can be handled directly by selecting a cell and keying in new information, but this can be very tedious, especially if there are many changes to be made. So how will we implement these activities to efficiently perform sensitivity analysis and what tools will we use? The answer is the *Form Controls*. This is an often-neglected tool that can enhance spreadsheet control and turn a pedestrian spreadsheet model into a user friendly and powerful analytical tool.

The Form Controls provides several functions that are easily inserted in a spreadsheet. They are based on a set of instructions that are bundled into a **Macro**. Macros as you will recall are a set of programming instructions that can be used to perform tasks by executing the macro. To execute a macro, it must be assigned a name, keystroke, or symbol. For example, macros can be represented by *buttons* (a symbol) that launch their instructions. Consider the need to copy a column of numbers on a worksheet, perform some manipulation of the column, and move the results to another worksheet in the workbook. You can perform this task manually, but if this task has to be repeated many times, it could easily be automated as a macro and attached to a button. By depressing the macro button, we can execute multiple instructions with a single key stroke and a minimum of effort. Additionally, macros can serve as a method of control for the types of interactions a user may perform. It is often very desirable to control user interaction, and thereby the potential errors and the misuse of a model that can result.

To fully understand Excel Macros, we need to understand the programming language used to create them, Microsoft Visual Basic for Applications (**VBA**). Although this language can be learned through disciplined effort, Excel has anticipated that most users will *not* be interested, or need, to make this effort. Incidentally, the VBA language is also available in MS Word and MS Project, making it very attractive to use across programming platforms. Excel has provided a shortcut that permits some of the important uses of macros, without the need to learn VBA. Some of these shortcuts are found in the Form Controls. In Excel 2003, the Forms Control menu was found by engaging the pull-down menu View and selecting Toolbars. In

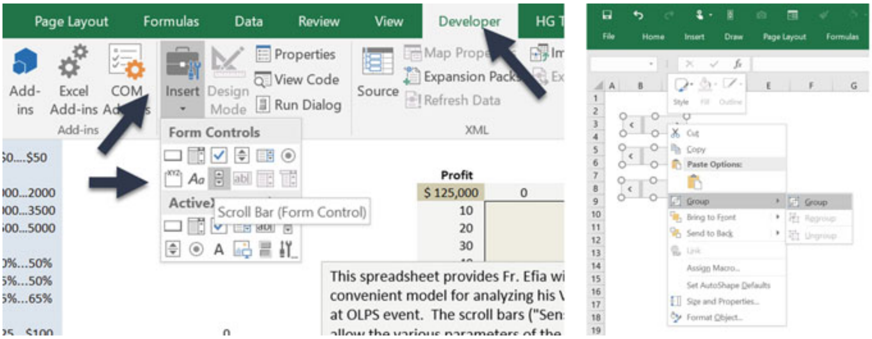


Fig. 7.16 Use of forms tools for control in developer Ribbon

Excel 2007 and beyond, Form Controls are not available in the standard ribbons, but can be placed into the Quick Access Toolbar. At the bottom of the Excel button or the File ribbon, depending on which version of Excel you are using, there is an Options menu. Upon entering the options, you can select Customize to add tools to the Quick Access Toolbar. One of the options is *Commands Not in the Ribbon*. You can also use the Customize Ribbon option. This is where Spin Button (Form Controls) and Scroll Bar (Form Controls) can be found. The arrow in Fig. 7.16 shows a number of icons: (1) List Box where an item can be selected from a list, (2) the Scroll Bar which looks like a divided rectangle containing two opposing arrow heads, (3) Spin Button which looks quite similar to the Scroll Bar, (4) the Option Button which looks like a circle containing a large black dot, and (5) Group Box for grouping buttons. The Active X controls are related to Macros that *you* must create with VBA; thus, do not use these unless this is your intent.

7.5.4 Option Buttons

Let us begin with the Option Button. Our first task is to consider how many buttons we want to use. The number of buttons will depend on the number of options you will make available in a spreadsheet model. For weather in the OLPS model, we have three options to consider and each option triggers several calculations. Additionally, for the sake of clarity, a single option will be made visible at a time. For example, in Fig. 7.7 the cloudy option is shown, and all others are hidden. The following are the detailed steps necessary to create a *group* of three options for which only one option will be displayed:

1. Creating a **Group Box**—We select the Group Box, the Form Controls icon designated as a box with XYZ. See Fig. 7.16. Drag-and-drop the Group Box onto the worksheet. See Fig. 7.17 for an example. Once located, a right click will allow you to move a Group Box. The grouping of Option Buttons in a Group Box alerts Excel that any Option Buttons placed in the box will be connected or associated with each other. Thus, by placing three buttons in the box, each button

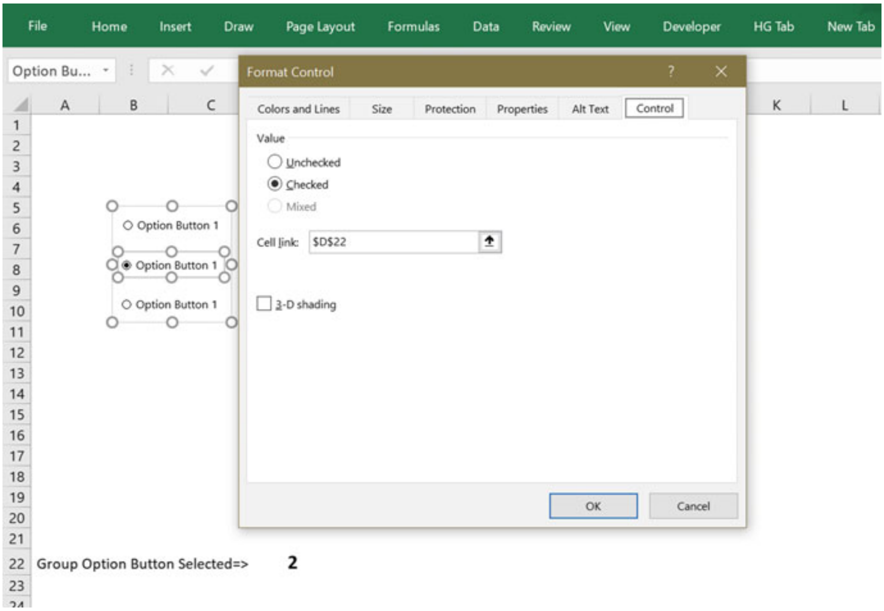


Fig. 7.17 Assigning a cell link to grouped buttons

will be assigned a specific output value (1, 2, or 3), and those values can be assigned to a cell of your choice on the worksheet. You can then use this value in a logical function to indicate the option selected. (If four buttons are used then the values will be 1, 2, 3, and 4.)

2. Creating Option Buttons—Drag-and-drop the Option in Form Controls into the Group Box. When you click on the Option Button and move your cursor to the worksheet, a cross will appear. Left click your cursor and drag the box that appears into the size you desire. This box becomes an Option Button. Repeat the process in the Group Box for the number of buttons needed. A right click will allow you to reposition the button and text can be added to identify the button.
3. Connecting button output to functions—Now we must assign a location to the buttons that will indicate which of the buttons is selected. Remember, that only one button can be selected at a time. Place the cursor on any button, and right click. A menu will appear, and the submenu of interest is Format Control. See Fig. 7.17. Select the submenu and then select the Control tab. At the bottom you will see a dialogue box requesting a **Cell Link**. In this box place the cell location where the buttons will record their unique identifier to indicate the single button that is selected. In this example, D22 is the cell chosen, and by choosing this location for one button, *all* grouped buttons are assigned the same cell link. Now, the cell can be used to perform worksheet tasks.
4. Using the Cell Link value—In Fr. Efia's spreadsheet model, the cell link values are used to display or hide calculations. For example, in Fig. 7.18 cell E21, the Attendance for the Cloudy scenario, contains: = IF(B29 = 2, C6,0). cell C6 is

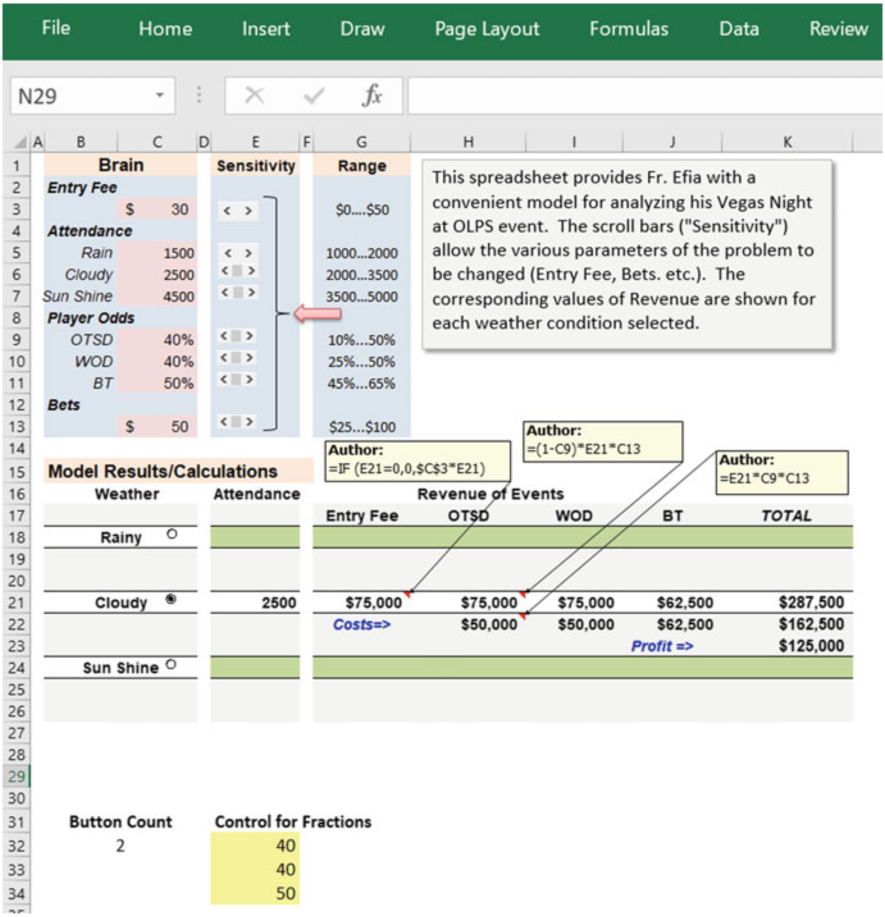


Fig. 7.18 Button counter logic and calculations

currently set to 2500. This logical function examines the value of B29, the cell link that has been identified in step 3. If B29 is equal to 2, it returns 2500 as the value cell E21. If it is not equal to 2, then a zero is returned. A similar calculation is performed for the Entry Fee in cell G21 by using the following cell function: = IF(E21 = 0,0,C3 ∗ E21). In this case, cell E21, the Attendance for Cloudy, is examined, and if it is found to be zero, then a zero is returned in E21. If it is *not* zero, then a calculation is performed to determine the Entry Fee revenue (2500 ∗ 30 = 75,000). The *Revenues of Events* are calculated in a similar manner. Note that it also is possible to show all values for all scenarios (Sunshine, Cloudy, and Rainy), and eliminate the logical aspect of the cell functions. Then the Option Buttons would not be needed. The buttons allow us to focus strictly on a single scenario. This makes sense since only one weather condition will prevail for a particular day of Vegas Night at OLPS. Of course, these choices are often a matter of taste for a specific application.

7.5.5 Scroll Bars

In the Brain, we also have installed eight Scroll Bars. See Fig. 7.18. They control the level of the variable in the column C. For example, the Scroll Bar in cell E5 controls C5 (1500), the attendance on a rainy day. These bars can be operated by clicking the left and right arrows, or by grabbing and moving the center bar. Scroll Bars can be designed to make incremental changes of a specific amount. For example, we can design an arrow click to result in an increase (or decrease) of five units, and the range of the bars must also be set to a specific maximum and minimum. Like the Option Button, a cell link needs to be provided to identify where cell values will be changed. Although Scroll Bars provide great flexibility and functionality, changes are restricted to be integer valued, for example, 1, 2, 3, etc. This will require additional minor effort if we are interested in producing a cell change that employs fractional values, for example percentages.

Consider the Scroll Bar located on E5 in Fig. 7.19. This bar, as indicated in the formula bar above, controls C5. By right clicking the bar and selecting the Format Control tab, one can see the various important controls available for the bar: Minimum value (1000), Maximum value (2000), Incremental change (50-the change due to clicking the arrows), and Page change (10-the change due to clicking between the bar and arrow). Additionally, the cell link must also be provided, and in this case it is C5. Once the link is entered, a right click of the button will show the cell link in the formula bar, C5 in this case.

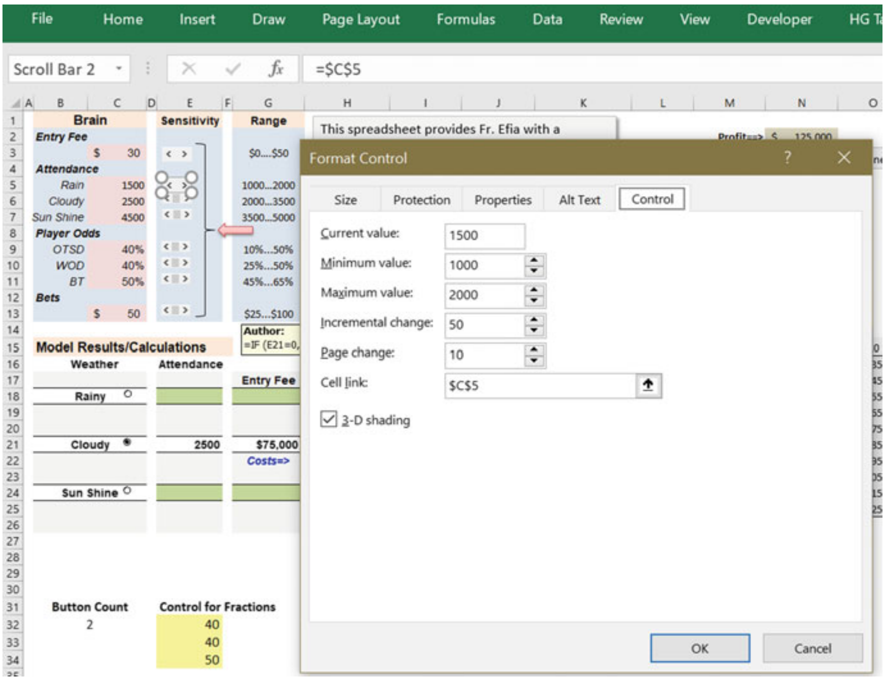


Fig. 7.19 Assigning a cell link to a scroll bar

Now, consider how we might use a Scroll Bar to control fractional values in Fig. 7.20. As mentioned above, since only integer values are permitted in Scroll Bar control, we will designate the cell link as cell E32. You can see the cell formula for C9 is E32/100. Dividing E32 by 100 will produce a fractional value, which we can use as a percentage. Thus, we can suggest the range of the Scroll Bar in G9 to range between 10 and 50, and this will result in a percentage in cell C9 from 10% to 50% for the Player Odds for OTSD.[8] You can also see that the other fractional odds are linked to Scroll Bars in cells E30 and E31. This inability to directly assign fractional values to a Scroll Bar is a minor inconvenience that can be easily managed.

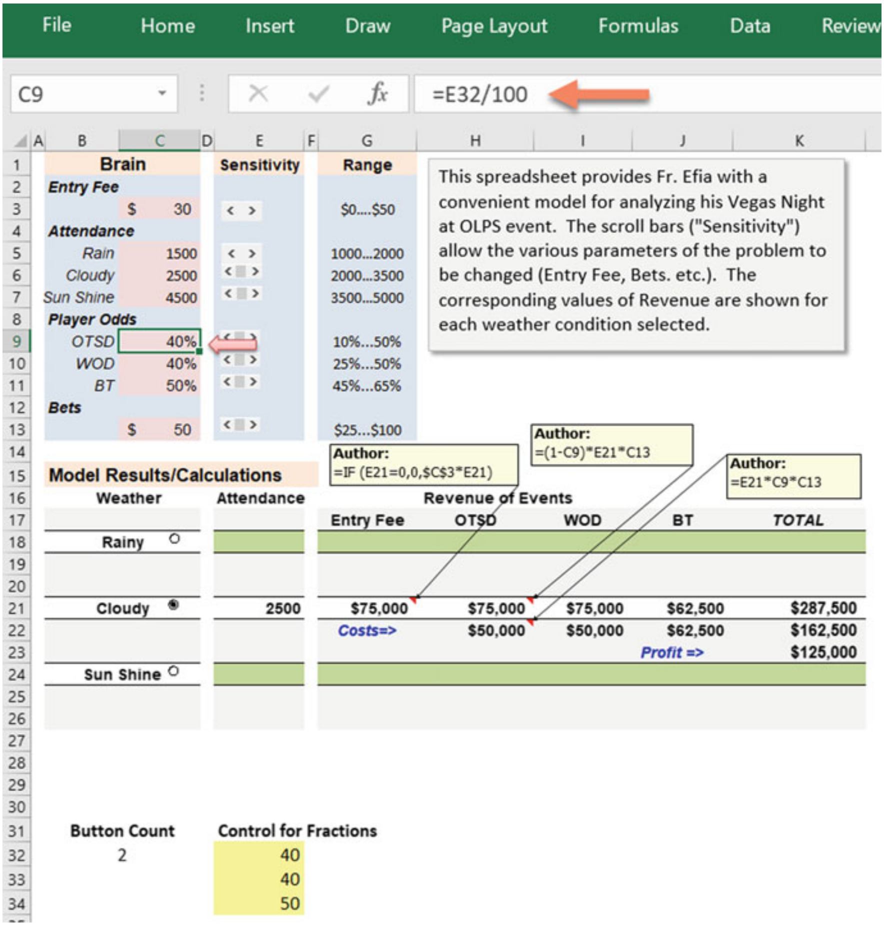


Fig. 7.20 Using a scroll bar for fractional values

[8]E29/100 = 40/100 = 0.40 or 40% . . . currently the value of the OTSD Player Odds.

7.6 Summary

This chapter has provided a foundation for modeling complex business problems. Although modeling contains an element of art, a substantial part of it is science. We have concentrated on the important steps for constructing sound and informative models to meet a modeler's goals for analysis. It is important that however simple a problem might appear, a rigorous set of development steps must be followed to insure a successful model outcome. Just as problem definition is the most important step in problem solving, model conceptualization is the most important step in modeling. In fact, I suggest that problem definition and model conceptualization are essentially the same.

In early sections of this chapter we discussed the concept of models and their uses. We learned the importance of classifying models to develop appropriate strategies for their construction, and we explored tools (Flow and Influence Diagrams) that aid in arriving at a preliminary concept for design.

All of our work in this chapter has been related to deterministic models. Although Fr. Efia's model contained probabilistic elements (game odds, uncertainty in weather conditions, etc.), we did not explicitly model these uncertain events. We accepted their deterministic equivalents: *expected value* of gambling outcomes. Thus, we have yet to explore the world of probabilistic simulation, in particular, Monte Carlo simulation. Monte Carlo simulation will be the focus of Chap. 8. It is a powerful tool that deals explicitly with uncertain events. The process of building Monte Carlo simulations will require us to exercise our current knowledge and understanding of probabilistic events. This may be daunting, but as we learned in previous chapters on potentially difficult topics, special care will be taken to provide thorough explanations. Also, I will present numerous examples to guide your learning.

This chapter has provided us with the essentials of creating effective models. In Chap. 8, we will rely on what we have learned in Chaps. 1–7 to create complex probabilistic models, and we will analyze the results of our model experimentation.

Key Terms

Data rich	Negative influence
Data poor	Mutually exclusive
Physical model	Collectively exhaustive
Analog model	Decision trees
Symbolic model	Expected value
Risk profiles	Sensitivity analysis
Deterministic	Quick access Toolbar8
Probabilistic	Scroll bars
PMT() function	Spinners
Model or problem definition phase	Combo boxes

(continued)

Data rich	Negative influence
Process flow map	Option buttons
Complexity	Developer ribbon
Pre-modeling or design phase	Data table
Modeling phase	Macro
Analysis phase	VBA
Final acceptance phase	Group box
Influence diagram (IFD)	Cell link
Positive influence	

Problems and Exercises

1. Data from a rich data environment is very expensive to collect—T or F?
2. What type of model is a site map that is associated with a website?
3. The x-axis of a risk profile is associated with probabilistic outcomes—T of F?
4. Deterministic is to probabilistic as point estimate is to range—T or F?
5. What is a single annual payment for the PMT() function for the following data: 6.75% annual interest rate; 360 months term; and $100,000 principal?
6. Draw a Process Flow Map of your preparation to leave your home, dormitory, or apartment in the morning. Use a rectangle to represent process steps like, *brush teeth*, and diamonds to represent decisions, like *wear warm weather clothes (?)*.
7. Create a diagram of a complex decision or process of your choice by using the structure of an influence diagram.
8. An investor has three product choices in a yearlong investment with forecasted outcomes—bank account (2.1% guaranteed); a *bond* mutual fund (0.35 probability of a 4.5% return; 0.65 probability of 7%), and a *growth* mutual fund (0.25 probability of −3.5% return, 50% probability of 4.5%, and remaining probability of 10.0%).

   (a) Draw the decision tree and calculate the expected value of the three investment choices. You decide that the maximum expected value is how you will choose an investment. What is your investment choice?
   (b) What will the guaranteed return for the bank deposit have to be to change your decision in favor of the bank deposit?
   (c) Create a spreadsheet that permits you to perform the following sensitivity analysis: What must the value of the largest return (currently 7%) for the bond fund be, for the expected value of the bond fund to be equal to the expected value of the growth fund?

9. For Fr. Efia's OLPS problem perform the following changes:
   (a) Introduce a 4th weather condition, *Absolutely Miserable*, where the number of alumni attending is a point estimate of only 750.
   (b) Perform all the financial calculations in a separate area below the others.
   (c) Add a scroll bar (range of 500–900) and an option button associated with the new weather condition, such that the spreadsheet look is consistent.
   (d) What will the entry fee for the new weather condition have to be for the profit to equal that in Fig. 7.7?
   (e) Find a different combination of Player odds that leads to the same Profit ($125,000) in Fig. 7.10.
   (f) Create a two-variable Data Table for cloudy weather, where the variables are Bet Value ($10 to $100 in $10 increments) and OTSD player odds (10–80% in 10% increments).

10. Create a set of four buttons that when a specific button is depressed (X) it provides the following message in a cell: *Button X is Selected* (X can take on values 1–4). Also, add conditional formatting for the cell that changes the color of the cell for each button that is depressed.
11. Create a calculator that asks a person their weight and permits them to choose, with a scroll bar, one of five % reductions 5, 10, . . .25%. The calculator should take the value of the percentage reduction and calculate their *desired* weight.
12. For the same calculator in 11, create a one-variable Data Table that permits the calculation of the desired weight for weight reduction from 1 to 25% in 1% increments.
13. *Advanced Problem*—Income statements are excellent mechanisms for modeling the financial feasibility of projects. Modelers often choose a level of revenue, a percent of the revenue as COGS (Cost of Goods Sold), and a percent of revenue as variable costs.
   (a) Create a deterministic model of a simple income statement for the data elements shown below (d-i)–(d-iv). The model should permit selection of various data elements using option buttons and scroll bars, as needed.
   (b) Produce a risk profile of the numerous combinations of data elements assuming that all data element combinations are of equal probability. (Recall, the vertical axis of a risk profile is the probability of occurrence of the outcomes on the horizontal axis, and in this case all probabilities are equal).
   (c) Also, provide summary data for all the resulting profit combinations for the problem—average, max, min, and standard deviation.
   (d) Data elements for the problem:
      (i) Revenue $100 k and $190 k (Option Button)
      (ii) COGS % of Revenue with outcomes of 24% and 47% (Option Button)
      (iii) Variable costs % of Revenue with outcome 35% and 45% (Option Button)
      (iv) Fixed costs $20 k to $30 k in increments of $5 k (Scroll bar)

(e) Create a Data Table that will permit you to change (with a scroll bar) the fixed cost in increments of $1 k that will result in instantaneous changes in the graph of the risk profile. Hint: combine (d-ii) and (d-iii) as a single variable and as a single dimension of a two variable Data Table, while using revenue as the second dimension. Fixed cost will act as a third dimension in the sensitivity analysis, but will not appear on the borders of the two variable Data Table.